

1 **Tempo and timing of ecological trait divergence in bird speciation**

2

3 Jay P. McEntee^{1,2}, Joseph A. Tobias³, Catherine Sheard⁴, & J. Gordon Burleigh¹

4

5 **Affiliations:**

6 ¹Biology Department, University of Florida, PO Box 118525, 220 Bartram Hall, Gainesville, FL 32611-
7 8525, USA

8 ²Ecology and Evolutionary Biology Department, University of Arizona, PO Box 210088, Biological
9 Sciences West Room 310, 1041 E. Lowell St., Tucson, Arizona 85721, USA

10 ³Department of Life Sciences, Imperial College London, Silwood Park, Ascot, SL5 7PY, UK

11 ⁴Department of Archaeology and Anthropology, University of Bristol, 43 Woodland Rd, Bristol, BS8 1UU,
12 UK

13

14

15

16 Number of words in the summary paragraph (241), in the manuscript as a whole, excluding summary
17 paragraph, references, methods, and captions (2,009)

18 Number of words in Methods for online version of the article (2,163)

19 Number of references in main text (28)

20 Number of figures (4) and tables (0)

21 Name and complete mailing address of the person to whom correspondence should be sent:

22 Jay McEntee, Biology Department, University of Florida, PO Box 118525, 220 Bartram Hall, Gainesville,

23 FL 32611-8525, USA Email: jaymcentee@ufl.edu

24

25 **Summary paragraph**

26 Organismal traits may evolve either gradually or in rapid pulses followed by periods of stasis, but the
27 relative importance of these evolutionary models in generating biodiversity has proven difficult to
28 resolve^{1,2}. In addition, while it is often assumed that pulses of trait evolution are associated with speciation
29 events, few studies have explicitly examined how the tempo of trait divergence varies with respect to
30 different geographical phases of speciation. Thus, we still know little about the trajectories of trait
31 divergence over timescales relevant to speciation, or the extent to which these trajectories are shaped by
32 variation in geographical isolation and overlap (sympatry) among incipient species. Here, we combine
33 divergence time estimates, trait measurements, and geographic range data for avian sister species pairs
34 worldwide to examine the tempo and timing of trait divergence during allopatric speciation. We show that
35 divergence in two important ecological traits—body mass and beak morphology—is best explained by a
36 model including pulses of divergence and periods of relative stasis. We also infer that trait divergence
37 pulses often precede sympatry, and that pulses leading to greater trait disparity are associated with earlier
38 transitions to sympatry. These findings suggest that early pulses of trait divergence promote subsequent
39 transitions to sympatry, rather than such pulses occurring after sympatry has been established, for example
40 via character displacement³. Incorporating pulsed divergence models into allopatric speciation theory helps
41 to resolve some apparently contradictory observations, including widespread instances of both rapid
42 sympatry and prolonged geographical exclusion⁴⁻⁶.

44 **TEXT**

45 Speciation in vertebrates may proceed over long and variable periods⁷. The entire process from onset to
46 completion is often subdivided into three stages, beginning with a phase of geographic isolation (allopatry),
47 followed by secondary contact, and finally the transition to coexistence in overlapping geographical ranges
48 (sympatry; see Fig. 1)⁷⁻¹⁰. In standard forms of this model, the third stage is delayed by competitive
49 interactions¹¹ or incomplete reproductive isolation⁹, and thus only occurs when traits are sufficiently
50 divergent to facilitate sympatry⁶. However, while the pattern of increased trait divergence in sympatric
51 lineages is widespread among animal taxa⁴, the timing and geographical context of the process of trait
52 divergence is often unclear. In particular, trait divergence could arise primarily by the accumulation of
53 differences prior to secondary contact¹² or alternatively after sympatry is established³. Furthermore, the
54 tempo and mode of trait divergence during speciation is also controversial, with some studies describing
55 divergence as slow or gradual throughout the process^{5,12}, while others provide evidence for abrupt, pulse-
56 like changes^{13,14} occurring either early in allopatry^{15,16} or later in sympatry^{3,6} (Fig. 1). It can even be argued

57 that ecological (local) adaptation in allopatry, followed by species interactions in sympatry, provide the
58 context for multiple pulses of trait divergence over time¹⁵.

59 Prolonged speciation creates problems when modeling pulsed trait evolution at a macroevolutionary
60 scale, and hence for most studies comparing evidence for pulsed versus gradual evolution. Pulsed models
61 typically assume that evolutionary change is concentrated at speciation², which is modelled as a single
62 instantaneous event, potentially leading to misinterpretations about trait evolution. Moreover, the
63 contribution of pulsed trait evolution may become obscured because stasis (or bounded evolution¹⁷) with
64 intermittent pulses can resemble gradualism at these coarse macroevolutionary scales². Micro-evolutionary
65 (population-level) studies conducted over short time periods confirm that stasis is prevalent¹⁸ and also that
66 trait evolution pulses sometimes occur^{4,17}. Observed pulse events, however, may represent brief departures
67 of trait values from their long-term means, followed by reversion, and may therefore not contribute
68 strongly to patterns of trait variation at macroevolutionary scales^{4,17,18}. Thus, although previous studies
69 have found support for both gradual and pulsed modes of evolution^{19,20}, their relative importance in
70 generating biological diversity remains unknown². An added complication is that trait divergence during
71 the allopatric phase of speciation may be ephemeral if gene pools merge during secondary contact, whereas
72 greater levels of divergence may lead to reproductive isolation and ultimately sympatry, a possibility that
73 could accentuate patterns of pulsed evolution in phylogenetic approaches and the fossil record, even when
74 divergence itself is gradual^{16,21}.

75 Disentangling these alternative divergence pathways is a key step in resolving general patterns of
76 trait evolution and predicting which nascent species ultimately succeed, leaving daughter species, and
77 which will perish or fail to remain distinct when changing environments re-organize geographic
78 ranges^{16,21,22}. However, our understanding of the rates and timing of trait divergence in relation to stages of
79 the speciation process in vertebrates remains highly incomplete, not least because the data required to test
80 these ideas are often lacking. In particular, although broad-scale information on ecological traits and
81 phylogenetic history is available for some large vertebrate clades²⁰, the accompanying information on
82 geographic ranges is not sufficiently resolved to explore divergence pathways in the context of
83 geographical phases of speciation.

84 To address this issue, we examined phenotypic divergence, geographic relationships, and estimated
85 divergence times²³⁻²⁵ among 952 pairs of avian sister species. As we were interested in the tempo and
86 timing of divergence during the speciation process, we applied a set of evolutionary models designed to
87 span both microevolutionary and macroevolutionary processes¹⁷ to estimated trait disparities and
88 divergence times. We used this approach because species pairs may vary in the extent to which their trait

89 divergence is better characterized by microevolutionary or macroevolutionary processes. Focusing on two
90 important ecological traits—body mass and beak morphology—we assessed relative support for four
91 stochastic trait divergence models: “gradual”, “single pulse”, “multiple pulse”, and time-independent
92 (“white noise”) models (see Methods). The first three models (i.e. all except the white noise model)
93 incorporate a bounded evolution component to represent processes at shorter timescales. We found the
94 strongest support for the single pulse model, in which the bounded evolution component is relatively
95 narrow²⁶ (Supplementary Tables 10-13). Support for the single pulse model was much stronger than for the
96 gradual model (Δ AIC 899 for body mass, Δ AICs 862, 1004, and 968 for beak PC1, PC2, and PC3,
97 respectively; see Methods and Supplementary Tables 10-13), in agreement with phylogenetic studies
98 reporting a strong contribution of pulses in the accumulation of phenotypic diversity^{19,27}.

99 Moreover, these results raise two further questions that we address here: 1) how does the estimated
100 timing of divergence pulses compare to typical progressions through the geographic speciation process
101 (allopatry, secondary contact, sympatry)? and 2) does pulsed ecological trait divergence impact transitions
102 through this process? With respect to the first question, the expected waiting time to a pulse of divergence
103 in the single pulse model was ~670,000 years (95% CI from likelihood profile: 280,000 years to 1.13 My)
104 for body mass (see also Figs. 3 and 4, Supplementary Table 10). The expected waiting times to a pulse in
105 beak morphology divergence in single pulse models were ~560,000 years for PC1 (95% CI: 200,000 years
106 to 1.0 My), ~170,000 years for PC2 (95% CI: 0 to 410,000 years), and ~90,000 years for PC3 (95% CI: 0
107 to 280,000 years; see Supplementary Tables 11-13). To compare these estimates with progression to
108 secondary contact and sympatry, we used a fine-grained geospatial database of ~178 million species
109 observation records and standard geographical range polygons, respectively (see Methods). Secondary
110 contact often occurs so rapidly that the signature of allopatry is difficult to detect in our analyses of local
111 co-occurrence (contact) and divergence time (Extended Data Fig. 4), suggesting that pulses may occur
112 following secondary contact or even during parapatric speciation (see below, Methods and Supplementary
113 Information). However, based on the relative timescales of trait divergence pulses and sympatry
114 establishment shown in (Fig. 4), we also conclude that divergence pulses typically precede the
115 establishment of sympatry, and are thus unlikely to be driven by character displacement processes³.

116 Early pulses of ecological trait divergence theoretically reduce both competition and reproductive
117 interference among incipient species¹⁵, potentially overcoming constraints on sympatry^{10,22}. To assess
118 whether such pulses influence rates of transition through geographical stages of the speciation process, we
119 tested whether variation in body mass and beak morphology predicted which species pairs are parapatric or
120 sympatric. Focusing on all species pairs found to locally co-occur ($n = 441$, see Methods), and accounting

121 for the effects of divergence time, dispersal ability, and latitude, we found strong evidence that sympatry is
122 associated with greater divergence in body mass, and, to a lesser extent, beak morphology (Fig. 2,
123 Extended Data Fig. 1, Supplementary Tables 4-5, 8-9, 18-19; see Methods). The relationship between large
124 body mass differences and increased likelihood of sympatry was highly consistent across sensitivity
125 analyses (Fig. 2, Extended Data Fig. 1, Supplementary Tables 4-5, 8-9, 18-19). These results are largely in
126 agreement with previous studies showing that the transition from secondary contact to sympatry is
127 facilitated by divergence in body mass and beak morphology^{10,12}, and further suggest that divergence in
128 body mass is a more critical factor.

129 We have shown that secondary contact occurs earlier in the speciation process than generally
130 assumed under classic models of allopatric speciation (Extended Data Figs 2, 4 and 7), suggesting a
131 potentially wider role for parapatric speciation (speciation with no stage a in Fig. 1)⁵. Under this model of
132 speciation, divergence pulses occur despite contact, and thus the potential for gene flow, between incipient
133 species¹⁵. Our analyses indicate that this scenario may be widespread in bird speciation. However, we also
134 found that the observed pattern of breeding co-occurrence and divergence times could also result from
135 purely allopatric speciation with rapid rates of transition to secondary contact (speciation with reduced
136 stage a in Fig. 1). Stochastic models indicate that an approximate minimum rate of 0.3 transitions to
137 secondary contact per million years is sufficient to explain the pattern of breeding co-occurrences among
138 avian sister species pairs (Extended Data Fig. 5, Supplementary Information). Thus, trait divergence pulses
139 may take place either during periods of allopatry or parapatry, with cases of both probably widespread.

140 The standard evolutionary trajectory implied by our best-fitting model—an early, pulsed divergence
141 with constrained subsequent divergence—can potentially explain a variety of phenomena in bird
142 speciation. Under this single pulse model, pulses vary in magnitude across species, with a fraction of
143 species undergoing large evolutionary jumps early in the speciation process, and others incurring small-
144 magnitude divergence¹⁷. Thus, many sister species strongly resemble each other in ecological traits whether
145 they began to diverge recently or anciently, whereas a fraction of species pairs have undergone an early
146 pulse of rapid divergence, and remain highly divergent regardless of their age. One interpretation of this
147 pattern is that a species pair undergoing a small pulse of ecological trait divergence in the early stage of
148 speciation is unlikely to undergo large pulses at later stages in the absence of other speciation events, and
149 thus the species pair may be subject to extended periods of mutual exclusion via ecological competition,
150 perhaps in combination with reproductive interference^{5,9}.

151 A prevailing view is that strong divergence in ecological traits between lineages typically requires
152 long periods of time, i.e. slow-rate gradual divergence¹². Gradual divergence models can account for

153 prolonged mutual exclusion between highly similar species, a widespread phenomenon⁵, especially when
154 evolutionary rates of gradual evolution are low^{2,10}. However, gradual divergence models inadequately
155 account for highly divergent young species pairs, unless they also incorporate brief bursts of faster gradual
156 divergence (mimicking pulses)²⁷. While gradual models thus may adequately explain avian trait evolution
157 at macroevolutionary scales, our findings suggest that patterns of diversification among sister pairs are
158 better captured by models incorporating pulses of divergence, where the magnitude of pulses is
159 independent of divergence time. In particular, the single pulse model receives strong support and can help
160 to explain the full range of outcomes observed in nature, which not only include prolonged mutual
161 exclusion between ecologically similar sister species⁶ but also instances of rapid sympatry following abrupt
162 ecological trait divergence⁴.

163 The occurrence of early pulses of trait divergence raises the question of how such pulses arise. One
164 possibility is that they result from the intermittent discovery by populations of unoccupied adaptive peaks,
165 as is expected in niche-filling models of diversification⁸—a mechanism that may be particularly important
166 for large-magnitude pulses⁴. New adaptive peaks be discovered when local adaptation drives rapid
167 divergence after range expansion^{14,22}, for instance immediately following colonization of novel
168 environments. Rapid phenotypic divergence may also result from local adaptation along environmental
169 gradients, with or without gene flow²⁸. In such contexts, it is worth emphasizing that signals of pulsed
170 divergence may arise from a combination of gradual local (clinal) adaptation and subsequent extinction of
171 intermediate populations^{16,29}.

172 In combination, our results may help to resolve the longstanding question of why some nascent
173 species survive over evolutionary time while others are ephemeral. One of the major threats to young
174 lineages is the likelihood of fusion through swamping gene flow after secondary contact^{16,21}. On one hand,
175 we have shown this risk is widespread among nascent bird species because the lag time to secondary
176 contact is shorter than expected (Extended Data Figs 2, 4-5), supporting the view that gene flow routinely
177 becomes possible early in the speciation process. On the other hand, our findings suggest that species pairs
178 undergoing major early pulses of ecological trait divergence are more likely to transition rapidly to
179 sympatry, escaping both fusion and mutual exclusion, thereby extending their lifespan as independent
180 lineages. Conversely, if they meet at early stages in the speciation process, species pairs with minimally
181 divergent phenotypes may incur increased hybridization rates, or increased hybrid fitness, thereby reducing
182 their lifespan. Indeed, elevated rates of extinction in less divergent young lineages may increase the
183 signature of large early pulses in datasets compiled from extant species. Thus, differential extinction

184 coupled with pulses of early trait divergence may play a critical role in explaining broad-scale patterns in
185 the longevity and macroevolutionary diversity of species, as well as their geographical distributions.

187 **Acknowledgements**

188 We are grateful to numerous data collectors who contributed to eBird, GenBank, and the CRC bird body
189 mass data set (see Supplementary Information). We also thank Nico Alioravainen, Ed Braun, Samuel
190 Jones, Rebecca Kimball, Dan Ksepka, Monte Neate-Clegg, Alex Pigot, Aaron Ragsdale and Gleb Zhelezov
191 for data collection and technical assistance. This work was supported by the National Science Foundation
192 (DEB-1208428 to J.G.B.), the Natural Environment Research Council (NE/I028068/1 to J.A.T.), and the
193 Oxford Clarendon Fund and US-UK Fulbright Commission (to C.S.).

194
195 **Author contributions** J.G.B. and J.P.M. conceived the study; J.G.B, J.P.M. and J.A.T. designed the
196 conceptual framework and analyses; J.G.B. performed dating analyses and assembled phylogenetic,
197 occurrence, and body mass information; J.A.T. and C.S. provided morphometric data; J.P.M. integrated
198 data sets, and designed and performed statistical analyses with significant input from J.G.B.; J.P.M.
199 produced figures and tables; J.P.M. wrote the manuscript, with significant input from all authors.

200
201 **Author Information** Reprints and permissions information is available at www.nature.com/reprints. The
202 authors declare no competing financial interests. Readers are welcome to comment on the online version of
203 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to
204 jaymcentee@ufl.edu.

206 **METHODS**

207 **Sister pairs** We used the maximum likelihood topology of Burleigh *et al.*'s²³ avian supermatrix
208 phylogenetic tree (hereafter “Burleigh tree”), which contains 6,714 species of the ~10,500 bird species in
209 the world, to select all (n = 2,076) pairs of avian sister species (i.e. each other's closest relatives). The
210 inclusion of some pairs of non-sister lineages would not invalidate our analyses but we tried to minimize
211 this issue by excluding pairs that were unlikely to represent true sister species. Specifically, we excluded
212 763 pairs belonging to genera with <75% species-level sampling, and another 62 pairs that were deemed
213 either unlikely to be true sister species based on molecular evidence from other studies or presented
214 taxonomic problems. We further removed 299 species pairs for which we could not adequately score co-
215 occurrence (Supplementary Dataset 6). This generated a manageable sample (n = 952) for downstream data

216 quality checks. Different analyses use subsets of these 952 pairs depending on data availability and quality
217 (see Supplementary Information Datasets 1-6).

218 **Divergence times** Divergence time estimates were obtained from a penalized likelihood analysis
219 implemented in r8s³⁰, using the maximum likelihood topology and molecular branch lengths from the
220 Burleigh tree. We used 20 carefully-vetted fossil calibrations²⁴ and constrained the root of the tree to a
221 maximum age of 110 mya; applying an age constraint made little difference to estimated sister pair
222 divergence times (see Extended Data Fig. 6). A list of the fossil calibrations (Supplementary Information
223 Dataset 6) and a command block for the r8s analyses are available in the Supplementary Information. We
224 performed sensitivity analyses using alternate sets of divergence time estimates, both from bootstrap
225 analysis of the Burleigh tree, and from an independent phylogenetic and dating analysis²⁵ (see
226 Supplementary Information).

227 **Ecological trait measurements** *Body mass* Divergence in body size may be a strong contributor to
228 ecological divergence, potentially reducing interspecific competition^{4,31} or reproductive interference⁹. We
229 compiled data on body mass (a proxy for body size) from updated global datasets³²⁻³⁴. When multiple body
230 mass values were reported, we took the mean; when male and female body masses were reported
231 separately, we calculated an average of the two sex-specific means. We estimated body mass divergence as
232 the difference between species in natural log of mean body mass¹⁷.

233 *Beak morphology* Species with similar body mass may partition niches according to diet. Thus, to quantify
234 differences in foraging ecology among sister species, we collected three beak measurements (culmen
235 length, beak depth, beak width) associated with food item selection and manipulation^{10,35,36}. Culmen length
236 was measured as the distance from the distal part of the nostril to the beak tip. Beak depth and beak width
237 were both measured at the distal edge of the nostril. All beak measurements were made on wild birds or
238 museum specimens using calipers to the nearest 0.1mm ($n \geq 4$ individuals sampled per species, 2 males and
239 2 females, where possible). See Supplementary Information for further details and rationale. To account for
240 colinearity between beak measurements, we performed a phylogenetic Principal Components Analysis
241 (phylogenetic PCA³⁷; see Supplementary Table 1 for PC loadings).

242 *Dispersal* Highly vagile taxa with greater dispersal capacities should undergo faster range expansions,
243 leading to earlier secondary contact in nascent species³⁸. This can be associated with faster transition rates
244 to sympatry³⁹, but when secondary contact is very early, it may also slow or reverse the speciation process
245 by promoting gene flow, leading to merged gene pools rather than coexistence²². Because of the
246 importance of dispersal in allopatric speciation models, we assess how dispersal capacity influences
247 transitions from allopatry to secondary contact, and from secondary contact to sympatry, respectively. As it

248 is difficult to measure dispersal capacity directly, we instead used the hand-wing index (HWI), an index of
249 wing shape related to the aspect ratio of the wing³⁸ and a proxy for flight performance⁴⁰. Using
250 measurements (to the nearest mm) taken from wild birds and museum specimens, we calculated this index
251 as

$$HWI = 100 \times \frac{WL - SL}{WL}$$

252
253
254 where WL (wing length) is the length of the closed wing from carpal joint to wing tip, and SL (secondary
255 length) is the distance from the carpal joint to the tip of the first secondary feather. As a secondary index of
256 dispersal, we also used range maps⁴¹ to assess migratory behaviour. If either member of a pair was
257 illustrated as migratory to any degree, the species pair was scored as migratory.

258

259 **Geographical phases of speciation**

260 *Secondary contact* We estimated local co-occurrence using ~178 million bird species observation records
261 stored in the eBird observational record database^{42,43}. For a given species pair, local co-occurrence was
262 defined as the occurrence of both species on the same day at the same reported locality. We also produced
263 a narrower dataset of breeding range local co-occurrence by checking the dates and localities of co-
264 occurrence records against breeding range maps⁴¹ and breeding phenology³². To qualify as evidence of
265 breeding range local co-occurrence, species had to be reported on the same day and in the same locality
266 during known breeding seasons of both species³² and within the known breeding range of one of the two
267 species⁴¹. All sister species pairs for which breeding local co-occurrence has been documented were
268 considered to have established secondary contact for the purposes of downstream analyses.

269 Because co-occurrence is unlikely to be reported for species with very few observations, we
270 excluded sister pairs where at least one species had fewer than 10 eBird sightings reported. Our co-
271 occurrence scores likely underestimate the true extent of co-occurrence among species pairs, as even after
272 this filtering process, the minimum number of observations strongly predicts the probability of species pair
273 local co-occurrence in our data set (GLM with the log of the minimum observations as sole predictor:
274 coefficient estimate = $3.8 \times 10^{-4} \pm 8.5 \times 10^{-5}$ SE; see Supplementary Information). Consequently, we
275 conducted sensitivity analyses adopting minima of 20 and 50 observations (see Extended Data Fig. 9;
276 Supplementary Tables 14-17). We also checked observational evidence for co-occurrence, discounting
277 cases likely attributable to anthropogenic introductions and excluding cases potentially based on
278 misidentifications or taxonomic confusion (see included and excluded species pairs in SI Datasets 1 and 6).

279 *Sympatry* To examine the transition from secondary contact to wider coexistence (sympatry), we calculated
280 percent breeding range overlap from geographic range polygons⁴¹ with a custom R script, using the R
281 libraries *rgdal*, *rgeos*, *maptools*, and *raster*. A small subset of species pairs (n = 17 of 441 species pairs)
282 could not be scored using our automated routine, and their range overlap was estimated visually. Species
283 pairs with a range overlap >20% of the smaller range were scored as sympatric, while those with breeding
284 co-occurrence but with ≤20% range overlap were scored as parapatric (having abutting ranges)³⁹. Because
285 some species pairs might best be considered sympatric even when their range overlap is less than 20%, we
286 performed sensitivity analyses using an overlap of >10% scored as sympatric (Supplementary Tables 18-
287 19).

288 **Analyses**

289 *Tempo of body mass divergence*

290 To examine the tempo and timing of body mass and beak PC1 divergence, we investigated the relative
291 support for four models of divergence for the species pairs from the full dataset for which body mass or
292 beak morphology data were available (n = 869 species pairs for body mass, n = 926 species pairs for beak
293 morphology). We fit one model of time-independent bounded evolution, and three different models that
294 comprise a bounded evolution component on shorter timescales and one of three additional components for
295 longer timescales¹⁷. These longer-timescale components are a gradual evolution model (Brownian motion)
296 and two forms of pulsed divergence: a single pulse model where a single instantaneous displacement
297 occurs following a waiting time sampled from an exponential distribution, and a multiple pulse model
298 where the expected number of displacements for a given divergence time is determined by a Poisson
299 process. We examined relative support for these models using AIC from likelihood calculations performed
300 in R. We calculated confidence intervals for the Poisson rate parameter λ , the inverse of which is taken as
301 the expected waiting time to a pulse, using likelihood profiling.

302 *Secondary contact and sympatry* We examined the probability of local co-occurrence, and parapatry versus
303 sympatry, using GLM with binomial error distributions, implemented in R⁴⁴. In analyses of local co-
304 occurrence and breeding range local co-occurrence for sister pairs, we began by predicting the probability
305 of co-occurrence with divergence time as the only predictor (Extended Data Figs 3a-b and 7). We
306 subsequently performed a model generation and selection routine (using the genetic algorithm of R
307 package *glmulti*⁴⁵, see Supplementary Information) to examine which among a set of phenotypic measures
308 best predict local co-occurrence or sympatry while accounting for the effects of three variables that may
309 influence the timing of transitions from allopatry to sympatry: divergence time¹², latitude⁹ and dispersal
310 ability³⁹. The predictors of primary interest were between-species disparity in two traits implicated in

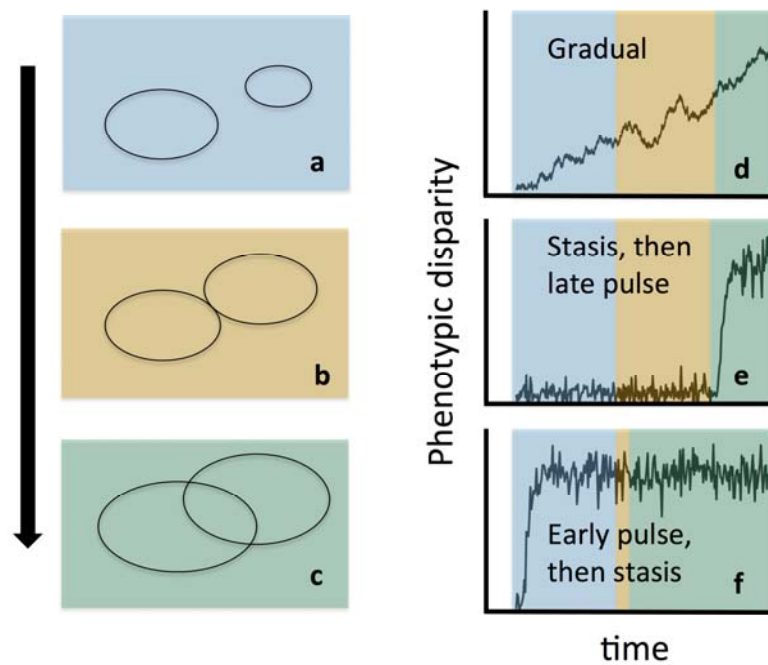
311 ecological and reproductive isolation: body mass⁴⁶ and beak morphology⁴⁷. We incorporated disparity in
312 beak morphology as (i) the Euclidean distance between species in PC space (following scaling of all PC's
313 to unit variance) or (ii) species differences along each PC axis as separate predictors. To account for
314 differences among sister species pairs in dispersal ability, we also included the average log hand-wing
315 index³⁸ and migratory status of the sister pair as predictors. We further included divergence time and
316 midpoint latitude (average of the two median observational latitudes for each species from eBird⁴³). Our
317 model generation routine permitted all pairwise interactions between predictors to enter the model, under
318 the constraint that all models were marginal. We report support for all predictors entering the set of local
319 co-occurrence models with $\Delta AIC < 2$ (Supplementary Tables 2-3). All continuous variables were scaled
320 and centered, such that estimated slope magnitudes for individual variables are meaningful in relation to
321 one another.

322 For GLM examining the probability of sympatry versus parapatry, we first limited the sister species
323 data set to those pairs that locally co-occur in breeding ranges. This restriction focuses the analysis on taxa
324 that have the opportunity to interact to some degree in the breeding season⁴⁸. The response variable in
325 GLM is the geographic configuration: parapatric (interacting but without substantial range overlap) versus
326 sympatric (having substantial range overlap: >20% of the smaller range in the analyses presented in the
327 main text). We again used a genetic algorithm (see Supplementary Information) to generate model variants
328 and performed model selection using the R package *glmulti*⁴⁵.

329 To assess the sensitivity of our results to uncertainty in phylogenetic inference and divergence time
330 estimates, we repeated all GLM analyses using mean divergence times for our species pairs from 100
331 samples of the pseudo-posterior distribution of trees from an alternative Bayesian species-level
332 phylogenetic analysis²⁵ (hereafter, the "Jetz tree"; Supplementary Information Tables S6-S9). For analyses
333 examining the probability of local co-occurrence (and breeding local co-occurrence) with divergence time,
334 we performed additional sensitivity analyses using divergence time estimates from 100 bootstraps of the
335 Burleigh tree, and for each of the 10,000 pseudo-posterior samples from the Jetz tree.

336
337 *Simulations of range dynamics* To aid in the interpretation of our GLM predicting local co-occurrence, we
338 performed stochastic range dynamic simulations³⁹. We used these simulations to place an approximate
339 lower bound on the rate of secondary contact establishment from an initially allopatric configuration. To
340 perform this estimation, we simulated the establishment of secondary contact using a simple model^{10,39}, in
341 which sister pairs can be in one of two states: co-occurring and not co-occurring. We simulated transitions
342 into and out of contact over a set of possible rates from 0.1 to 0.8 per million years, in which the forward

343 rate (rate of transition from isolation to contact, σ) is always greater than or equal to the reverse rate (rate of
344 transition out of contact, ϵ). The forward and reverse rates are constant³⁹, and the variation in rates among
345 species arises only from stochasticity. Reverse rates were simulated at .005, .01, .05, .1, .2, and .5 times
346 each of the forward rates. We present the maximum intercept calculated across all reverse rates (ϵ) for each
347 simulated forward rate (σ) (Extended Data Figs 5, 12). To calculate the approximate percentage of species
348 pairs coming into secondary contact by given points in time following divergence (100,000 years, 1 million
349 years), we simulated range dynamics with $\sigma = 0.3$, and $\epsilon = .15$ (corresponding to the minimal σ that yielded
350 intercept >0.434 , and the value of ϵ that yielded the highest intercept for $\sigma = 0.3$).
351

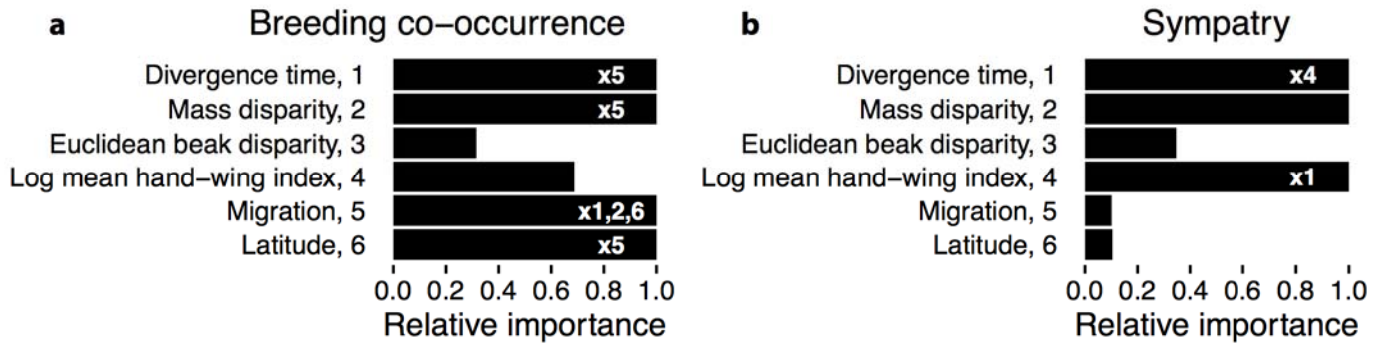


352
353
354
355
356
357
358
359
360

Figure 1 | The speciation cycle and phenotypic trait divergence

Bird speciation typically involves a sequence of geographical states, starting with an allopatric phase (a), followed by secondary contact (b), and finally sympatry (c). Phenotypic divergence may take different pathways during this cycle: gradual models predict no pulse of divergence at any point in the cycle (d), whereas punctuated models involve stasis punctuated by pulses, which can follow the onset of coexistence (e) or precede it (f). Note that secondary contact (orange) is extended when traits are similar (e), and reduced when traits have already substantially diverged in allopatry (f).

361
362
363
364
365

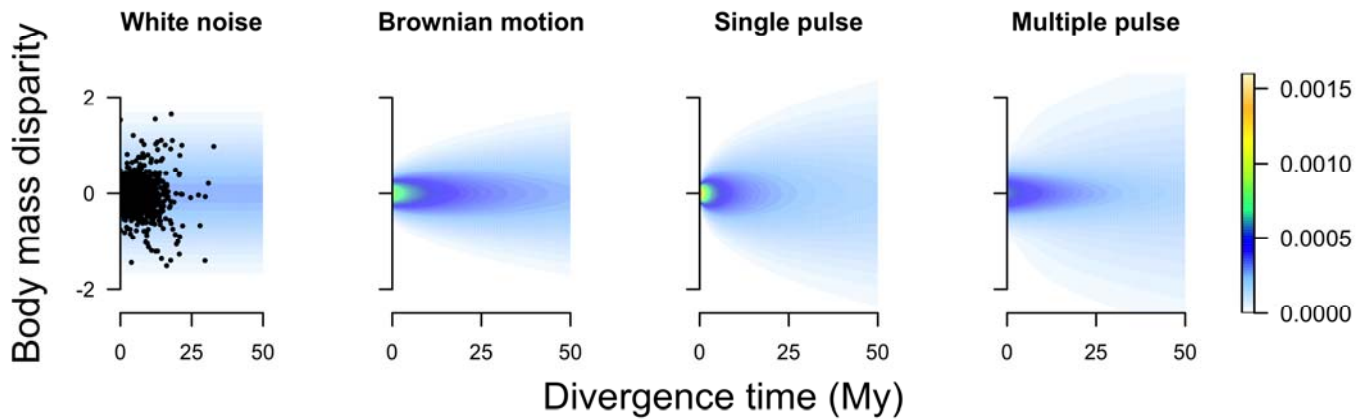


366
367

368 **Figure 2 | Factors associated with the establishment of secondary contact and sympatry in birds.**

369 Results of generalized linear models assessing the relative importance of predictors of breeding range local
370 co-occurrence (a) and sympatry (b) in 952 avian sister species. Pairs with breeding co-occurrence include
371 both parapatric and sympatric species pairs. Relative importance is estimated as the proportion of the
372 summed model weights for all models with $\Delta AIC < 2$, and indicates the extent to which each variable
373 predicts the probability of co-occurrence or sympatry. Pairwise interactions with relative importance > 0.6
374 are indicated by the numbers within the bar for each variable.

375
376



377

378

379 **Figure 3 | Tempo of body mass divergence for avian sister species.** Stochastic pulsed models provide

380 better fits to patterns of body mass divergence and divergence time, with the best fit a single pulse model

381 (Δ AIC relative to the multiple pulse model: 797). Colors denote probability density. The probability

382 density for any time slice follows a normal distribution (most apparent in the white noise model where the

383 probability density distribution is independent of time). Relative probability density can be assessed within

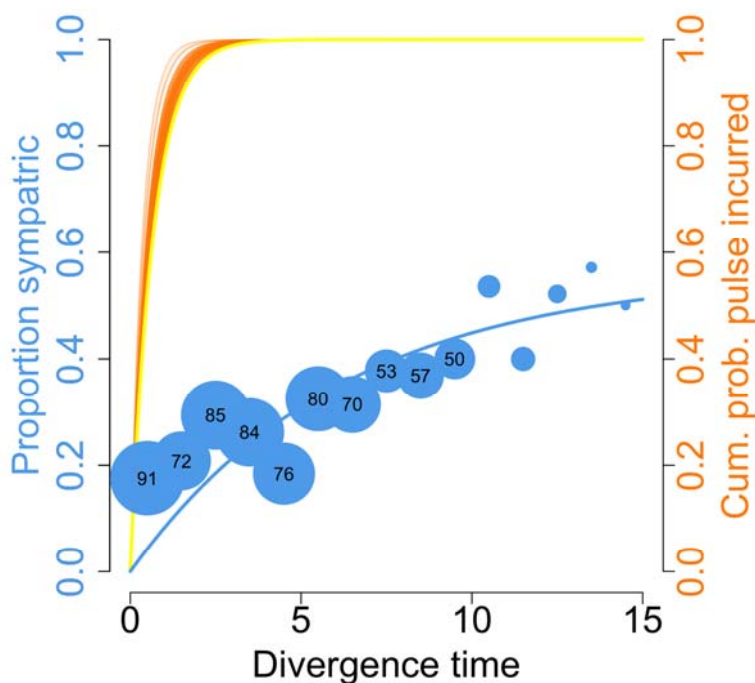
384 each time slice but not across time. For clarity, the empirical data points are plotted only on the white noise

385 model.

386

387

388



389

390

391 **Figure 4 | Timing of body mass divergence pulses and sympatry.** Comparison of timescales suggests
392 that mass divergence tends to precede sympatry among 952 avian sister species. Yellow and orange lines
393 are cumulative probability distributions of incurring a pulse under the single pulse model (yellow:
394 estimated from maximum likelihood phylogenetic tree²³; orange: from 100 bootstrap trees²³). Circles are
395 proportions of sympatric species pairs for 1-million year intervals of divergence time²³; circle sizes
396 represent sample sizes, numbered where ≥ 50 . For visual comparison, an exponential decay model has been
397 fitted to the proportion of sister pairs in sympatry (blue curve; assumes sympatry is secondary).

398

399

400

401

402

References

- 403 1. Eldredge, N. & Gould, S. in *Models in Paleobiology* (Freeman, Cooper, San Francisco, 1972).
- 404 2. Pennell, M. W., Harmon, L. J. & Uyeda, J. C. Is there room for punctuated equilibrium in
405 macroevolution? *Trends Ecol. Evol.* **29**, 23-32 (2014).
- 406 3. Pfennig, D. W. & Pfennig, K. S. Character displacement and the origins of diversity. *Am. Nat.* **176**
407 **Suppl 1**, S26-44 (2010).
- 408 4. Schluter, D. in *The Ecology of Adaptive Radiation* (Oxford University Press, Oxford, UK, 2000).
- 409 5. Price, T. in *Speciation in Birds* (Roberts and Company, Greenwood, Village, Colorado, 2008).

- 410 6. Rundell, R. J. & Price, T. D. Adaptive radiation, nonadaptive radiation, ecological speciation and
411 nonecological speciation. *Trends Ecol. Evol.* **24**, 394-399 (2009).
- 412 7. Mayr, E. in *Systematics and the Origin of Species* (Columbia University Press, 1942).
- 413 8. Price, T. D. *et al.* Niche filling slows the diversification of Himalayan songbirds. *Nature* **509**, 222-
414 225 (2014).
- 415 9. Weir, J. T. & Price, T. D. Limits to speciation inferred from times to secondary sympatry and ages of
416 hybridizing species along a latitudinal gradient. *Am. Nat.* **177**, 462-469 (2011).
- 417 10. Pigot, A. L. & Tobias, J. A. Species interactions constrain geographic range expansion over
418 evolutionary time. *Ecol. Lett.* **16**, 330-338 (2013).
- 419 11. MacArthur, R. & Levins, R. The limiting similarity, convergence, and divergence of coexisting
420 species. *Am. Nat.*, 377-385 (1967).
- 421 12. Tobias, J. A. *et al.* Species coexistence and the dynamics of phenotypic evolution in adaptive
422 radiation. *Nature* **506**, 359-363 (2014).
- 423 13. Smith, J. W. & Benkman, C. W. A coevolutionary arms race causes ecological speciation in
424 crossbills. *Am. Nat.* **169**, 455-465 (2007).
- 425 14. Friis, G., Aleixandre, P., Rodríguez-Estrella, R., Navarro-Sigüenza, A. G. & Milá, B. Rapid postglacial
426 diversification and long-term stasis within the songbird genus Junco: phylogeographic and
427 phylogenomic evidence. *Mol. Ecol.* **25**, 6175-6195 (2016).
- 428 15. Rundle, H. D. & Nosil, P. Ecological speciation. *Ecol. Lett.* **8**, 336-352 (2005).
- 429 16. Futuyma, D. J. On the role of species in anagenesis. *Am. Nat.* **130**, 465-473 (1987).
- 430 17. Uyeda, J. C., Hansen, T. F., Arnold, S. J. & Pienaar, J. The million-year wait for macroevolutionary
431 bursts. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15908-15913 (2011).
- 432 18. Estes, S. & Arnold, S. J. Resolving the paradox of stasis: Models with stabilizing selection explain
433 evolutionary divergence on all timescales. *Am. Nat.* **169**, 227-244 (2007).
- 434 19. Bokma, F. Time, species, and separating their effects on trait variance in clades. *Syst. Biol.* **59**, 602-
435 607 (2010).
- 436 20. Rabosky, D. L. & Adams, D. C. Rates of morphological evolution are correlated with species
437 richness in salamanders. *Evolution* **66**, 1807-1818 (2012).
- 438 21. Futuyma, D. J. in *Macroevolution* 29-85 (Springer, 2015).
- 439 22. Mayr, E. in *Animal Species and Evolution* (Belknap Press of Harvard University Press, Cambridge,
440 Massachusetts, 1963).
- 441 23. Burleigh, J. G., Kimball, R. T. & Braun, E. L. Building the avian tree of life using a large-scale, sparse
442 supermatrix. *Mol. Phylogenet. Evol.* **84**, 53-63 (2015).
- 443 24. Baiser, B., Valle, D., Zelazny, Z. & Burleigh, J. G. Non-random patterns of invasion and extinction
444 reduce phylogenetic diversity in island bird assemblages. *Ecography* (2017).
- 445 25. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in
446 space and time. *Nature* **491**, 444-448 (2012).
- 447 26. Arnold, S. J. Phenotypic evolution: the ongoing synthesis (American Society of Naturalists
448 address). *Am. Nat.* **183**, 729-746 (2014).
- 449 27. Landis, M. J., Schraiber, J. G. & Liang, M. Phylogenetic analysis using Levy processes: finding jumps
450 in the evolution of continuous traits. *Syst. Biol.* **62**, 193-204 (2013).
- 451 28. Endler, J. A. in *Geographic Variation, Speciation, and Clines (Vol. 10)* (Princeton University Press,
452 1977).
- 453 29. Futuyma, D. J. Evolutionary constraint and ecological consequences. *Evolution* **64**, 1865-1884
454 (2010).
- 455 30. Sanderson, M. R8s: Inferring absolute rates of molecular evolution and divergence times in the
456 absence of a molecular clock. *Bioinformatics* **19**, 301-302 (2003).

- 457 31. Wilson, D. S. The adequacy of body size as a niche difference. *Am. Nat.* **109**, 769-784 (1975).
458 32. del Hoyo et al. (eds.). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona.
459 (retrieved from <http://www.hbw.com/> in 2015-16).
460 33. Dunning, J. B. in *Body Masses of Birds of the World* (Taylor and Francis Group, Boca Raton, Florida,
461 2008).
462 34. Dunning, J.B. Body Masses of Birds of the World.
463 <https://ag.purdue.edu/fnr/Documents/WeightBookUpdate.pdf>. (2016).
464 35. Schoener, T. W. The evolution of bill size differences among sympatric congeneric species of
465 birds. *Evolution* **19**, 189-213 (1965).
466 36. Miles, D. B. & Ricklefs, R. E. The correlation between ecology and morphology in deciduous forest
467 passerine birds. *Ecology* **65**, 1629-1640 (1984).
468 37. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things).
469 *Methods Ecol. Evol.* **3**, 217-223 (2012).
470 38. Claramunt, S., Derryberry, E. P., Remsen, J. V., Jr. & Brumfield, R. T. High dispersal ability inhibits
471 speciation in a continental radiation of passerine birds. *Proc. R. Soc. Lond. B* **279**, 1567-1574 (2012).
472 39. Pigot, A. L. & Tobias, J. A. Dispersal and the transition to sympatry in vertebrates. *Proc. R. Soc.*
473 *Lond. B* **282**, 20141929 (2015).
474 40. Dawideit, B. A., Phillimore, A. B., Laube, I., Leisler, B. & Böhning-Gaese, K. Ecomorphological
475 predictors of natal dispersal distances in birds. *J. Anim. Ecol.* **78**, 388-395 (2009).
476 41. BirdLife International and NatureServe. in *Bird species distribution maps of the world* (BirdLife
477 International and NatureServe, Cambridge, UK and Arlington, USA, 2014).
478 42. Sullivan, B. L. *et al.* eBird: A citizen-based bird observation network in the biological sciences.
479 *Biol. Conserv.* **142**, 2282-2292 (2009).
480 43. Sullivan, B. L. *et al.* The eBird enterprise: An integrated approach to development and application
481 of citizen science. *Biol. Conserv.* **169**, 31-40 (2014).
482 44. R Core Team. R: A language and environment for statistical computing. (2012).
483 45. Calcagno, V. & de Mazancourt, C. glmulti: an R package for easy automated model selection with
484 (generalized) linear models. *J. Stat. Softw.* **34**, 1-29 (2010).
485 46. Yasukawa, K. Male quality and female choice of mate in the red-winged blackbird (*Agelaius*
486 *phoeniceus*). *Ecology* **62**, 922-929 (1981).
487 47. Grant, P. R. & Grant, B. R. Hybridization, sexual imprinting, and mate choice. *Am. Nat.* **149**, 1-28
488 (1997).
489 48. Lovette, I. J. & Hochachka, W. M. Simultaneous effects of phylogenetic niche conservatism and
490 competition on avian community structure. *Ecology* **87**, S14-S28 (2006).
491