

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
7 32611-8525, USA

8 ²Ecology and Evolutionary Biology Department, University of Arizona, PO Box 210088,
9 Biological 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 ⁴School of Biology, University of St. Andrews, Harold Mitchell Building, St Andrews Fife,
12 KY16 9TJ, United Kingdom

13 *corresponding author

14

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

16 Jay McEntee, Biology Department, University of Florida, PO Box 118525, 220 Bartram Hall,
17 Gainesville, FL 32611-8525, USA Email: jaymcentee@ufl.edu

18

19 **Abstract**

20 Organismal traits may evolve either gradually or in rapid pulses, but the relative importance of
21 these modes in the generation of species differences is unclear. Additionally, while pulsed
22 evolution is frequently assumed to be associated with speciation events, few studies have
23 explicitly examined how the tempo of trait divergence varies with respect to different
24 geographical phases of speciation, starting with geographic isolation and ending, in many cases,
25 with spatial overlap (sympatry). Here we address these issues by combining divergence time
26 estimates, trait measurements, and geographic range data for 952 avian sister species pairs
27 worldwide to examine the tempo and timing of trait divergence in recent speciation events. We
28 show that patterns of divergence in key ecological traits are not gradual, but instead appear to
29 follow a pattern of relative stasis interspersed with evolutionary pulses of varying magnitude. We
30 also find evidence that evolutionary pulses generally precede sympatry, and that major pulses
31 leading to greater trait disparity are associated with sympatry. These findings suggest that early

32 pulses of trait divergence promote subsequent transitions to sympatry, rather than occurring after
33 sympatry has been established. Incorporating models with evolutionary pulses of varying
34 magnitude into speciation theory may explain why some species pairs achieve rapid sympatry
35 whereas others undergo prolonged geographical exclusion.

36
37

38 Speciation in vertebrates may proceed over long and variable periods¹⁻⁶. From onset to
39 completion, the process is often viewed as a cycle with three stages, beginning with geographic
40 isolation (allopatry), followed by secondary contact initiated at range edges, and finally
41 prolonged spatial coexistence in overlapping geographical ranges (sympatry; see Fig. 1)^{1, 5-7}.
42 The process does not always proceed to the second or third stages, but the establishment of
43 sympatry has been viewed by some as the only way to be certain that speciation is complete². In
44 some forms of this model, here termed “geographic speciation,” sympatry can be delayed after
45 secondary contact has been made, either by competitive interactions⁸ or incomplete reproductive
46 isolation⁵. Thus sympatry may only occur when species traits are sufficiently divergent⁹, while
47 recently diverged species with similar traits fail to overlap, and instead exhibit abutting ranges
48 (termed parapatry). However, while the pattern of increased trait divergence in sympatric versus
49 non-sympatric sister lineages is widespread among animal taxa^{10, 11}, the timing and geographical
50 context of trait divergence is difficult to determine. In particular, trait divergence could arise
51 primarily by the accumulation of differences either prior to sympatry¹² or after sympatry is
52 established (e.g. via character displacement)¹³.

53 The tempo and mode of ecological trait divergence during the geographic speciation
54 cycle are also debated^{14, 15}, with some studies describing divergence as slow or gradual
55 throughout the process^{4, 12}, while others describe abrupt, pulse-like changes occurring against a
56 background of stasis¹⁶. Pulsed changes could occur either early, in allopatry or parapatry^{17, 18}, or
57 later, in sympatry^{9, 13} (Fig. 1). It is further possible that ecological (local) adaptation in
58 allopatry/parapatry, followed by subsequent species interactions in sympatry provide the context
59 for multiple pulses of trait divergence over the course of geographic speciation cycles¹⁷.

60 Assessing the relative contributions of gradual versus pulse-and-stasis evolution to
61 species differences involves challenges associated with temporal scale¹⁹ and analytical
62 approach¹⁴. On short timescales, microevolutionary studies indicate that stasis generally
63 dominates²⁰, with abrupt pulses of divergence occasionally detected²¹. It is unclear whether these

64 pulses ultimately contribute to species-level differences or simply represent brief departures of
65 trait values from longer-term static, or gradually changing, means^{10, 14, 20, 22}. Fossil time series,
66 which allow analyses of complex mixtures of modes over timescales relevant to geographic
67 speciation (10^4 to 10^7 years), show evidence for gradual change and stasis within lineages^{23, 24}.
68 Taxonomic issues and patchy geographic sampling of the fossil record, however, limit the extent
69 to which the fossil record can be used to assess phenotypic divergence as a component of the
70 geographic speciation cycle¹⁴.

71 These challenges have led to an increased use of phylogenetic approaches to assess the
72 tempo and mode of phenotypic change associated with speciation. These methods generally use
73 data from extant species to test for correlations between trait evolution and speciation²⁵ or to
74 partition trait evolution between cladogenetic and anagenetic change²⁶. These approaches
75 suggest that pulse-like speciation (cladogenetic) processes may contribute to trait divergence.
76 However, such analyses are often conducted across large clades, potentially overlooking the role
77 of extinction in generating observed patterns in extant diversity²⁷⁻²⁹. Moreover, the assumption in
78 most phylogenetic approaches that evolutionary pulses occur only at speciation¹⁵, which is
79 modelled as a single instantaneous event^{30, 31}, limits how much these approaches can tell us about
80 the roles of pulsed and gradual evolution over the extended course of geographic speciation
81 cycles.

82 Disentangling alternative divergence pathways at different stages of speciation is a key
83 step to understand general patterns of trait evolution and predict the fate of nascent species when
84 changing environments redistribute geographic ranges^{2, 18, 32}. For example, minor trait divergence
85 during the allopatric phase of speciation may be ephemeral if gene pools merge during secondary
86 contact, whereas greater levels of divergence may lead to reproductive isolation and ultimately
87 sympatry. This latter possibility could accentuate patterns of pulsed evolution in phylogenies and
88 the fossil record, even when divergence itself is gradual^{18, 32}. Framed by these uncertainties, our
89 understanding of the rates and timing of trait divergence in vertebrate speciation remains highly
90 incomplete, not least because the data required are often patchy for large samples of species. In
91 particular, for most large taxonomic groups the information on ecological trait divergence and
92 geographic distributions is not sufficiently resolved to explore divergence pathways in the
93 context of geographical phases of speciation.

94 To address these issues, we examine phenotypic divergence, geographic relationships,
95 and divergence times among 952 pairs of avian sister species. By restricting the analysis to sister
96 species, we explicitly focus on ecological trait divergence associated with relatively recent
97 speciation events, which reduces the impact of species extinction on our inferences relative to
98 clade-wide approaches^{27,33}. For each pair of species, we estimated trait divergence and
99 determined geographical relationships using a geospatial database of ~178 million species
100 observation records^{34,35} and a set of standard range polygons. Combining detailed phenotypic,
101 temporal and spatial information, we consider (1) whether pulse-and-stasis or gradual evolution
102 predominate in ecological trait divergence over the extended duration of bird speciation, (2)
103 whether divergence in ecological traits is associated with sympatry establishment, and (3) how
104 the estimated timing of divergence relates to the geographic phases of speciation.

105

106 **Results**

107 Using estimated divergence times for species pairs generated from phylogenetic data³⁶⁻³⁸, along
108 with estimates of trait disparity for each pair of species, we fitted a set of evolutionary
109 divergence models²² designed to span microevolutionary and macroevolutionary processes. We
110 assessed relative support for four stochastic models: a Brownian motion (“gradual”) divergence
111 model, a “single pulse” divergence model, a “multiple pulse” divergence model, and a time-
112 independent (“white noise”) divergence model. The first three models (gradual, single pulse,
113 multiple pulse) incorporate an additional white noise component to represent bounded evolution
114 processes at shorter timescales and to help to account for measurement error. In the single and
115 multiple pulse models, the pulses are modeled as instantaneous jumps in trait values, which are
116 representative of brief periods of high-rate directional evolution²² as expected in rapid evolution
117 toward a new adaptive optimum³⁹ and not as saltational jumps. We interpret the single pulse
118 model as a “pulse-and-stasis” model, as the bounds of the bounded evolution component are
119 narrow⁴⁰ (Supplementary Tables 1 through 4). For two important ecological traits -- body mass
120 and beak morphology (see Methods) -- the single pulse model has strong support relative to the
121 other three models. Sensitivity analyses (Supplementary Tables 1 - 4) indicated that these results
122 were robust to an alternate approach to phylogenetic reconstruction and divergence time
123 estimation³⁸.

124 The difference in support between the single pulse and gradual models is consistent with
125 phylogenetic studies that have found evidence for a contribution of speciation pulses in the
126 accumulation of phenotypic diversity²⁶: Δ AIC between the single pulse and Brownian motion
127 models is 899 for body mass, Δ AICs between the single pulse and Brownian motion models are
128 808, 920, and 967 for culmen length, beak depth, and beak width, respectively (see Methods and
129 Supplementary Tables 1 through 4). Our analyses suggest, however, that such pulses need not be
130 completely coincident with speciation, i.e. as reconstructed as instantaneous events (nodes) on
131 phylogenetic trees. Instead, such pulses may take place at some point in the course of a
132 protracted geographic speciation process³⁰, as indicated by non-zero estimated waiting times to
133 pulses.

134 To estimate the timing of these divergence pulses, we used the rate parameter in the
135 single pulse model preferred in our analyses. For body mass, the expected waiting time to a pulse
136 was ~670,000 years (95% CI from likelihood profile: 275,000 years to 1.13 My; see also Figs. 2
137 and 3, Supplementary Table 1). The expected waiting times to a pulse in beak morphology
138 divergence in single pulse models were ~440,000 years for culmen length (95% CI: 200,000 to
139 930,000 years), ~360,000 years for beak depth (95% CI: 100,000 to 720,000 years), and
140 ~420,000 years for beak width (95% CI: 130,000 to 820,000 years; see Supplementary Tables 2 -
141 4).

142 To compare the accumulation of trait divergence with transitions from allopatry to
143 sympatry, we calculated and visualized the cumulative distribution function (CDF) for
144 undergoing a pulse of divergence over time (Fig. 3, see also Supplementary Figure 1). The pulse
145 magnitudes in the trait divergence models are variable (specifically they are drawn from a
146 normal distribution with a mean of 0), such that we only expect large-magnitude pulses for a
147 small fraction of species pairs. Most species pairs in this model accumulate modest trait
148 divergence. Under this model, however, large-magnitude divergence is possible not only for
149 older species pairs, but also for younger species pairs.

150 To compare our estimates of pulse timing and trait divergence accumulation with the
151 progression to secondary contact (inclusive of both parapatry and sympatry) and sympatry, we
152 obtained information on contact from a geospatial database of ~178 million species observation
153 records^{34, 35} and on sympatry from standard geographical range polygons⁴¹ (Fig. 4). Comparing
154 the inferred timescales of trait divergence and sympatry suggests that pulses of trait divergence

155 largely precede the establishment of sympatry (Fig. 3), which occur predominantly on the order
156 of 10^6 years. This result is contingent on the assumption that transitions out of sympatry are
157 infrequent⁷. The relative timescales of inferred trait divergence pulses and sympatry
158 establishment (Fig. 3) suggest that many divergence pulses occur in the absence of sympatric
159 interactions that could cause character displacement¹³. By contrast, we found that secondary
160 contact occurs early enough and frequently enough among species pairs that the signature of
161 initial allopatry is absent in our analyses of secondary contact and divergence time
162 (Supplementary Figure 3; Supplementary Information: Timing of secondary contact and
163 sympatry), leaving open the possibility that trait divergence often follows secondary contact, or
164 that parapatric speciation (without an initial allopatric stage) occurs with regularity among birds
165 (see below, Methods and Supplementary Information).

166 Pulses of ecological trait divergence theoretically reduce competition and reproductive
167 interference among incipient species¹⁷, potentially overcoming constraints on sympatry
168 establishment^{2,7}. The possibility that divergence by pulse-and-stasis evolution is widespread
169 during speciation raises the question of whether divergence pulses are consequential for
170 transitions through the geographic speciation process. Thus we tested whether variation in body
171 mass and beak morphology predicted which species pairs are parapatric (with abutting
172 distributions) or sympatric (with overlapping distributions). Focusing on all species pairs in
173 secondary contact ($n = 440$, see Methods), and accounting for the effects of divergence time,
174 dispersal ability, and latitude, we found evidence that sympatry is associated with greater
175 divergence in ecological traits (Fig. 4, Supplementary Tables 8 - 9). In different analyses, either
176 body mass divergence or aspects of beak morphology divergence had more support as predictors
177 of sympatry (Supplementary Tables 8 - 9). Relationships between increased probability of
178 sympatry and ecological trait divergence were present across sensitivity analyses accounting for
179 a different approach to phylogenetic reconstruction and divergence time estimation
180 (Supplementary Tables 12 - 13), and an alternate threshold for delineating sympatry versus
181 parapatry (Supplementary Tables 14 - 15). These results are largely in agreement with previous
182 studies showing that the transition from secondary contact to sympatry is associated with
183 divergence in body mass and/or beak morphology^{11,12}.

184 The hand-wing index, a morphological proxy of dispersal ability^{42,43}, was also a
185 consistent predictor of sympatry for species pairs in contact across analyses (Fig. 5,

186 Supplementary Information). This effect could be observed if sympatric populations tend to be
187 sinks supported by non-sympatric sources: species with elevated dispersal capacity may maintain
188 sympatric sink populations whereas those with weak dispersal may fail to. Additionally,
189 increased dispersal capacity may be associated with elevated founder population sizes or
190 densities during colonization of the sister species' range. These elevated founder population sizes
191 may facilitate colonization within the ranges of sister species by surmounting the challenges of
192 population establishment at low density or number^{44, 45}.

193 Lastly, we inferred that contact between members of species pairs often occurs early (i.e.
194 within the first two million years) in the geographic speciation cycle (Supplementary Fig. 3),
195 contrary to models of geographic speciation that specify that long periods of allopatry are
196 necessary for speciation². This suggests a potentially wider role for parapatric speciation
197 (speciation with no stage *a* in Fig. 1)⁴ than is often considered in bird speciation⁴⁶. In parapatric
198 speciation, divergence pulses prior to sympatry would necessarily occur despite contact, and thus
199 the potential for gene flow, between incipient species¹⁷. However, we also found that observed
200 patterns of contact and divergence time among bird sister species could be explained by
201 speciation models requiring some period of allopatry at the outset, but allowing for rapid rates of
202 transition to secondary contact (i.e. speciation with reduced duration of stage *a* in Fig. 1;
203 Supplementary Figs. 9 - 10, Supplementary Information). Stochastic modeling of species ranges
204 indicated that an approximate minimum rate of transition to secondary contact of 0.3 transitions
205 per million years is sufficient to explain the pattern of contact among sister pairs. That is, a 30%
206 probability of a species pair coming into contact within a million years is sufficient to yield a
207 combined set of divergence times for species pair contact approximating the empirical
208 observations. Thus, trait divergence pulses may occur during periods of either allopatry or
209 parapatry, with cases of both likely widespread. Our evidence for early contact during speciation
210 suggests that selection (e.g. local adaptation, immigrant inviability⁴⁷ or other immigrant
211 disadvantages, and/or selection against intermediate phenotypes) frequently plays a role in
212 driving or maintaining divergence in the early stages of speciation.

213 214 **Discussion**

215
216 A prevailing view on bird speciation is that ecological trait divergence sufficient to
217 enable sympatry requires long periods of allopatry^{9, 12}. Gradual evolution models, such as the

218 random walk models used in many phylogenetic comparative methods for continuous traits, are
219 consistent with this prevailing view^{7, 15}. However, gradual divergence models inadequately
220 account for strong divergence in young species pairs, unless they also incorporate brief bursts of
221 faster gradual divergence^{48, 49}, similar to pulses. Even the Early Burst model⁵⁰, used to represent
222 the explosively rapid trait divergence dynamics in extant adaptive radiations²⁸, poorly accounts
223 for highly divergent recent species pairs because it predicts comparatively slow evolutionary
224 rates at the tips of phylogenetic trees. Thus, while widely used gradual models may adequately
225 explain avian trait evolution at deeper macroevolutionary scales (as for comparing genera or
226 families²⁸), our findings indicate that patterns of divergence among sister pairs are better
227 captured by models incorporating pulsed divergence amid periods of stasis. In particular, the
228 single pulse model we test here receives relatively strong support, and can help to explain the
229 range of geographic speciation outcomes observed in nature, which includes many “old” species
230 pairs that are highly similar in ecological traits⁹ but also instances of abrupt ecological trait
231 divergence in young species pairs^{10, 17, 51, 52}.

232 The signal of trait divergence pulses during the geographic speciation cycle raises the
233 question of how such pulses may take place. We suggest that pulses are likely to result from the
234 intermittent discovery, by populations or lineages, of unoccupied adaptive peaks, as is expected
235 in niche-filling models of diversification^{6, 10}, or more generally from adaptive evolution on
236 heterogeneous fitness landscapes^{14, 53}. Unoccupied adaptive peaks may become occupied, via
237 pulses of adaptive directional evolution, during or immediately following range expansion^{16, 54},
238 for instance immediately following colonization of novel environments. Rapid phenotypic
239 divergence may also result from local adaptation along environmental gradients, with or without
240 gene flow⁵⁵. In such contexts, signals of pulsed divergence could plausibly arise from a
241 combination of local (clinal) adaptation and the subsequent extinction of phenotypically
242 intermediate populations^{16, 18, 56}. A key shared aspect of all these scenarios is that evolution on
243 adaptive landscapes allows for pulse-and-stasis modes of divergence¹⁴. Our data may also be
244 consistent with a role for character displacement⁵⁷ driving pulsed divergence in some instances.
245 Comparing the timescales of trait divergence and sympatry establishment (Fig. 3), however,
246 suggests that divergence pulses more frequently result from processes not requiring sympatry
247 than processes that require sympatry.

248 Our finding that greater divergence in ecological traits is positively associated with the
249 probability of sympatry accords well with what has been found previously in many other studies
250 of vertebrates^{58, 59}, including birds^{11, 12}, using a variety of different methodological approaches.
251 The focus of such studies is commonly character displacement^{11, 58, 60}, yet this association can
252 also be explained by the spatial sorting of species ranges by competition^{12, 61} or because these
253 trait differences may be associated with stronger reproductive isolation^{5, 62}. Tobias et al.¹² found
254 that the association of greater divergence in size and beak morphology of sympatric versus non-
255 sympatric ovenbird species could be explained by a third factor, time since divergence, which
256 they interpreted as evidence for spatial sorting of ecologically diverged lineages. Our results
257 provide indirect support for this sorting hypothesis in a sample of species pairs that spans the
258 diversity of birds, in that the timescale of ecological trait divergence under the single pulse
259 model is faster than the relatively slow accumulation of sympatry sister species pairs (Fig. 3)⁶².
260 These findings suggest that pulses of trait divergence promote subsequent transitions to
261 sympatry, rather than such pulses primarily occurring through character displacement after
262 sympatry has been established. This interpretation is contingent on the assumption that the
263 movement of species ranges over time results in low transition rates of sympatry to non-
264 sympatry in sister pairs, i.e. that the curve in Fig. 3 is mostly representative of the initial
265 achievement of secondary sympatry and not of an equilibrium between relatively even transition
266 rates into and out of sympatry^{7, 63}.

267 In combination, our results may help to resolve the longstanding question of why only
268 some nascent species survive over evolutionary time. One of the major threats to young lineages
269 is fusion through swamping gene flow upon contact². We have demonstrated that this risk is
270 widespread among nascent bird species because the lag time to secondary contact is generally
271 short (Supplementary Figs. 3 and 6 - 8), supporting the view that gene flow routinely becomes
272 possible early in the speciation process. In particular, species pairs with minimally divergent
273 phenotypes may incur increased hybridization rates or increased hybrid fitness relative to more
274 diverged pairs, potentially leading to the extinction of one member of the species pair.
275 Importantly, though, our findings also suggest that species pairs undergoing major early pulses
276 of ecological trait divergence are more likely to rapidly transition to sympatry, escaping both
277 fusion and mutual exclusion, and possibly extending their lifespan as independent lineages^{56, 64-}
278 ⁶⁶. Elevated rates of extinction in less-divergent young lineages may increase the signature of

279 large early pulses in datasets compiled from extant species, such as that presented here. Thus,
280 differential extinction coupled with pulses of early trait divergence may play a critical role in
281 explaining broad-scale patterns in the longevity, macroevolutionary diversity, and geographical
282 distribution of species.

283

284 **METHODS**

285 **Sister pairs** We selected all ($n = 2,076$) pairs of avian sister species (i.e. each other's closest
286 relatives) from the maximum likelihood topology from Burleigh *et al.*'s³⁶ avian supermatrix
287 phylogenetic analysis (hereafter "Burleigh tree"), which contains 6,714 species of the ~10,500
288 bird species in the world. The inclusion of pairs of non-sister lineages would not invalidate our
289 analyses, but for the purposes of interpretation we excluded pairs that may not represent true
290 sister species because one taxon may have a closer relative. Specifically, we excluded 763 pairs
291 belonging to genera with <75% species-level sampling, and another 62 pairs that were deemed
292 unlikely to be true sister species based on either molecular evidence from other studies or
293 taxonomic problems. For most analyses, we further removed 299 species pairs for which we
294 could not adequately score the presence or absence of contact (Supplementary Dataset 6).
295 Different analyses used different subsets of sister pairs, depending on data availability and
296 quality (see Supplementary Datasets 1 - 6). In using the sister pair data set to make inferences
297 about the speciation process, we leveraged the broad variation in divergence times and stages of
298 the geographic speciation cycle represented among sister pairs, including pairs that are in the
299 early stages of divergence.

300 **Divergence times** Divergence time estimates were obtained from an ultrametric tree inferred
301 from the maximum likelihood tree from Burleigh *et al.*³⁶ using penalized likelihood analysis
302 implemented in r8s^{37, 67}. The r8s analysis used 20 fossil calibrations and constrained the root of
303 the tree to a maximum age of 110 mya. However, applying a maximum age constraint to the root
304 node had little effect on estimates of sister pair divergence times (Supplementary Fig. 5). A list
305 of the fossil calibrations (Supplementary Information Dataset 7, see also⁶⁸) and a command
306 block for the r8s analyses are available in the Supplementary Information. We performed
307 sensitivity analyses using alternate sets of divergence time estimates, both from bootstrap
308 analysis of the Burleigh tree, and from an independent phylogenetic and dating analysis³⁸ (see
309 Supplementary Information).

310 **Ecological trait measurements** *Body mass* Divergence in body size may be a strong contributor
311 to ecological divergence, potentially reducing interspecific competition^{10, 69} or reproductive
312 interference⁵. We compiled data on body mass (a proxy for body size) for species pairs from
313 updated global datasets⁷⁰⁻⁷². When multiple body mass values were reported, we took the mean;
314 when male and female body masses were reported separately, we calculated an average of the
315 two sex-specific means. We estimated body mass disparity as the difference between species in
316 natural log of mean body mass. This estimate is a unitless measurement representing
317 proportional change in body mass²².

318 *Beak morphology* Species with similar body mass may partition niches according to diet. Thus,
319 to quantify differences in foraging ecology among sister species, we collected three beak
320 measurements (culmen length, beak depth, beak width) associated with food item selection and
321 manipulation^{7, 73, 74}. Culmen length was measured as the distance from the distal part of the
322 nostril to the beak tip. Beak depth and beak width were both measured at the distal edge of the
323 nostril. All beak measurements were made on wild birds or museum specimens using calipers, to
324 the nearest 0.1mm (n \geq 4 individuals sampled per species, \geq 2 males and \geq 2 females, where
325 possible; see Supplementary Information for further details about the beak morphology data set).
326 For evolutionary mode analyses, we estimated proportional differences between species'
327 phenotypes by taking the difference of the natural log of the mean body mass of each species²².
328 To calculate multivariate species differences, we first fit separate phylogenetic generalized linear
329 regressions of the log species means of each of the beak measurements on species log mean body
330 mass (with Ornstein-Uhlenbeck errors⁷⁵, see Supplementary Information section 5: Predictors of
331 contact and sympatry). We then calculated the Mahalanobis distance between the three beak
332 measurement residuals from the phylogenetic generalized linear regressions for each species pair
333 to estimate multivariate divergence in beak morphology.

334 *Dispersal* Highly vagile taxa with greater dispersal capacities should undergo faster range
335 expansions, potentially leading to earlier secondary contact in nascent species⁴³. This can be
336 associated with faster transition rates to sympatry⁶³, but when secondary contact is very early, it
337 may also slow or reverse the speciation process by promoting gene flow, leading to merged gene
338 pools rather than continued divergence². Because of the importance of dispersal in geographic
339 speciation models, we assessed how dispersal capacity influences transitions from allopatry to
340 secondary contact, and from secondary contact to sympatry, respectively. As it is difficult to

341 measure dispersal capacity directly, we instead used the hand-wing index (HWI), an index of
342 wing shape related to the aspect ratio of the wing⁴³ and a proxy for flight performance and
343 dispersal capacity in birds^{42, 43, 63, 76}. Using measurements (to the nearest 0.5 mm) taken from
344 wild birds and museum specimens, we calculated this index as

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

346
347 where WL (wing length) is the length of the closed wing from carpal joint to wing tip, and SL
348 (secondary length) is the distance from the carpal joint to the tip of the first secondary feather. As
349 a secondary index of dispersal, we also used range maps⁴¹ to assess migratory behavior. If either
350 member of a pair was illustrated as migratory to any degree, the species pair was scored as
351 migratory.

352

353 **Geographical phases of speciation**

354 *Secondary contact* We examined evidence for contact in species pairs using ~178 million bird
355 species observation records stored in the eBird observational record database^{34, 77}. We
356 downloaded the ebird Basic Dataset (EBD_reIOct-2013) from the eBird website
357 (<http://ebird.org>). For a given species pair, contact was defined by evidence for the co-occurrence
358 of both species on the same day at the same reported locality. Evidence that such co-occurrence
359 takes place was interpreted as evidence for contact between species ranges. We wrote a Perl
360 script that, for all sister species pairs, identified all instances when both species were listed at the
361 exact same latitude and longitude on the same day. However, the spatial extent of eBird
362 checklists varies when it is reported (Supplementary Figure 4), such that some apparent co-
363 occurrence records might come from checklists made for large areas or long transects.
364 Furthermore, some co-occurrence records may result from human-aided introductions to areas
365 extrinsic to native ranges or from ambiguity in different taxonomic treatments. To address these
366 issues, we compared our contact scoring with what would result from contact scoring from range
367 maps⁴¹ and examined eBird records where there was any discrepancy (see additional details in
368 Supplementary Information: *General supplementary methods*). These data quality checks
369 resulted in the re-scoring of 15 species pairs (from having contact to having no contact) and the
370 exclusion of an additional 29 species pairs from the analysis where contact could not be scored
371 with confidence (Supplementary Dataset 6).

372 Because contact can take place outside of the breeding range in cases where either
373 member of a species pair is migratory, we scored breeding contact for such species pairs
374 specifically by examining whether co-occurrence took place in the breeding range and breeding
375 season, using breeding range maps⁴¹ and breeding phenology descriptions⁷⁰ respectively. To
376 qualify as evidence of breeding contact, species co-occurrences in species pairs with any
377 migratory behavior had to be reported on the same day and in the same locality during the known
378 breeding seasons of both species⁷⁰ and within the known breeding range of at least one of the
379 two species⁴¹. For species pairs without known migratory behavior, we assumed evidence of
380 contact was indicative of breeding contact (see also Supplementary Information).

381 Because co-occurrence is unlikely to be reported for species with very few observations,
382 we excluded sister pairs where at least one species had fewer than 10 eBird sightings reported.
383 Our contact scores likely underestimate the true extent of contact among species pairs, as even
384 after this filtering process, the minimum number of observations strongly predicts the probability
385 of species pair contact in our data set (GLM with the log of the minimum observations as sole
386 predictor: coefficient estimate = $3.8 \times 10^{-4} \pm 8.5 \times 10^{-5}$ SE; Supplementary Table 5).
387 Consequently, we conducted sensitivity analyses adopting an alternate minimum threshold, 20 or
388 50 instead of 10, for observations (Supplementary Table 5 and Supplementary Fig. 8). We also
389 checked observational evidence for contact, discounting co-occurrence cases likely attributable
390 to anthropogenic introductions, and excluding cases potentially based on misidentifications or
391 taxonomic confusion (see included and excluded species pairs in Supplementary Datasets 1 and
392 6).

393 *Sympatry* To examine the transition from contact to broad range overlap (sympatry), we
394 calculated percent breeding range overlap from geographic range polygons⁴¹ with a custom R
395 script, using the R libraries *rgdal*, *rgeos*, *maptools*, and *raster*. A small subset of species pairs (n
396 = 17 of 440 species pairs) could not be scored using our automated routine, and their range
397 overlap was estimated visually. Species pairs with a range overlap > 20% of the smaller range⁶³,
398 ^{78, 79} were scored as sympatric, while those with contact but with $\leq 20\%$ range overlap were
399 scored as parapatric (having abutting ranges)⁶³. This range overlap threshold may seem high
400 based on the definition of parapatry as the circumstance where ranges abut but are separate⁸⁰.
401 Range maps are coarse-grained, however, such that ranges that do not substantially overlap in
402 reality appear to have overlap in range maps. As this 20% threshold, like any threshold, was

403 somewhat arbitrary, we performed additional analyses using an overlap of > 10% scored as
404 sympatric (Supplementary Tables 14-15).

405 **Analyses**

406 *Tempo of ecological trait divergence*

407 To examine the tempo and timing of trait divergence, we evaluated the relative support for four
408 divergence models across the set of species pairs for which body mass or beak morphology data
409 were available (n = 869 species pairs for body mass, n = 945 species pairs for beak morphology).
410 For all four traits, we defined disparities as the between-species difference in the log mean²². We
411 fit one model of time-independent bounded evolution, as well as three different models that
412 comprise a bounded evolution component on shorter timescales and one of three additional
413 components for longer timescales²². These longer-timescale components are a gradual evolution
414 model (Brownian motion) and two forms of pulsed divergence: a single pulse model where a
415 single instantaneous displacement occurs following a waiting time sampled from an exponential
416 distribution and a multiple pulse model where the expected number of displacements for a given
417 divergence time is determined by a Poisson process. We examined relative support for these
418 models using AIC from likelihood calculations performed in R. We calculated confidence
419 intervals for the Poisson rate parameter λ , the inverse of which is taken as the expected waiting
420 time to a pulse, using likelihood profiling. To generate likelihood profiles, we calculated the
421 likelihood of single pulse models with a series of fixed λ values at increments of 0.01, over an
422 interval containing the maximum likelihood parameter estimate. We assume the log likelihood
423 ratio is chi-square distributed, and define the 95% confidence interval for λ as the interval for
424 which the likelihood is within 1.92 units of the maximum likelihood model. To visualize the
425 cumulative probability of incurring a pulse for Fig. 3, we used the equation for the cumulative
426 distribution function for an exponential distribution, $1 - e^{-\lambda t}$.

427 *Secondary contact and sympatry* We examined the probability of contact versus non-contact, and
428 parapatry versus sympatry, using GLM with binomial error distributions, implemented in R⁸¹. In
429 analyses of contact and breeding contact for sister pairs, we began by predicting the probability
430 of contact, with divergence time as the only predictor (Extended Data Figs 3a-b and 7). We
431 subsequently performed a model generation and selection routine (using the genetic algorithm of
432 R package glmulti⁸², see Supplementary Information) to examine which among a set of
433 phenotypic measures best predict contact or sympatry while accounting for the effects of three

434 variables that may influence the timing of transitions from allopatry to sympatry: divergence
435 time¹², latitude⁵, and dispersal ability⁶³. The predictors of primary interest were between-species
436 disparity in two traits implicated in ecological and reproductive isolation: body mass⁸³ and beak
437 morphology⁸⁴. We incorporated disparity in beak morphology (i) as the Mahalanobis distance
438 between species residuals from phylogenetic linear regressions of beak measurements on body
439 mass (see above), or (ii) by including each of the three residuals (from phylogenetic linear
440 regressions of culmen length, beak width, and beak depth on body mass) as predictors. We used
441 disparities in mean body mass and beak morphology as predictors in these analyses instead of
442 sex-specific disparities as (i) many individuals in these datasets are not sexed, (ii) overall sexual
443 size dimorphism is minor, and (iii) species interactions involve both males and females, such that
444 the mean of these measurements is likely relevant.

445 To account for differences among sister species pairs in dispersal ability, we also
446 included the average log hand-wing index⁴³ and migratory status of the sister pair as predictors.
447 We further included divergence time and midpoint latitude (average of the two median
448 observational latitudes for each species from eBird⁷⁷). Our model generation routine permitted
449 all pairwise interactions between predictors to enter the model, under the constraint that all
450 models were marginal. We report support for all predictors entering the set of contact models
451 with $\Delta\text{AIC} < 2$ (Supplementary Tables 6 - 7). All continuous variables were scaled and centered,
452 such that estimated slope magnitudes for individual variables are meaningful in relation to one
453 another.

454 For GLM examining the probability of sympatry versus parapatry, we first limited the
455 sister species data set to those pairs with breeding contact. This restriction focuses the analysis
456 on taxa that have the opportunity to interact to some degree in the breeding season⁸⁵. The
457 response variable in GLM is the geographic configuration: parapatric (interacting but without
458 substantial range overlap) versus sympatric (having substantial range overlap: $> 20\%$ of the
459 smaller range in the analyses presented in the main text). We again used a genetic algorithm (see
460 Supplementary Information) to generate model variants and performed model selection using the
461 R package *glmulti*⁸².

462 To assess the sensitivity of our results to uncertainty in phylogenetic inference and
463 divergence time estimates, we repeated all GLM analyses using mean divergence times for our
464 species pairs from 100 samples of the pseudo-posterior distribution of trees from an alternative

465 Bayesian species-level phylogenetic analysis³⁸ (hereafter, the “Jetz tree”; Supplementary Tables
466 9 - 12). For analyses examining the probability of contact (and breeding contact) with divergence
467 time, we performed additional sensitivity analyses using divergence time estimates from 100
468 bootstraps of the Burleigh tree and for each of the 10,000 pseudo-posterior samples from the Jetz
469 tree (Supplementary Fig. 7).

470

471 *Simulations of range dynamics* To aid in the interpretation of our GLM predicting local co-
472 occurrence, we performed stochastic range dynamic simulations⁶³. We used these simulations to
473 place an approximate lower bound on the rate of secondary contact establishment from an
474 initially allopatric configuration. To perform this estimation, we simulated the establishment of
475 secondary contact using a simple model^{7,63}, in which sister pairs can be in one of two states: co-
476 occurring and not co-occurring. We simulated transitions into and out of contact over a set of
477 possible rates from 0.1 to 0.8 per million years, in which the forward rate (rate of transition from
478 isolation to contact, σ) is always greater than or equal to the reverse rate (rate of transition out of
479 contact, ϵ). The forward and reverse rates are constant⁶³, and the variation in rates among species
480 arises only from stochasticity. Reverse rates were simulated at 0.005, 0.01, 0.05, 0.1, 0.2, and 0.5
481 times each of the forward rates. We present the maximum intercept calculated across all reverse
482 rates (ϵ) for each simulated forward rate (σ) (Extended Data Figs 5, 12). To calculate the
483 approximate percentage of species pairs coming into secondary contact by given points in time
484 following divergence (100,000 years, 1 million years), we simulated range dynamics with $\sigma =$
485 0.3, and $\epsilon = 0.15$ (corresponding to the minimal σ that yielded intercept > 0.434 , and the value of
486 ϵ that yielded the highest intercept for $\sigma = 0.3$).

487 *Code availability* Computer code for analyses not found in the Supplementary Information can
488 be obtained from the authors upon request.

489

490 **References**

491

- 492 1. Mayr, E. in *Systematics and the Origin of Species* (Columbia University Press, 1942).
- 493 2. Mayr, E. in *Animal Species and Evolution* (Belknap Press of Harvard University Press,
494 Cambridge, Massachusetts, 1963).
- 495 3. Mayr, E. & Diamond, J. M. in *The birds of northern Melanesia: speciation, ecology &*
496 *biogeography* (Oxford University Press on Demand, 2001).
- 497 4. Price, T. in *Speciation in Birds* (Roberts and Company, Greenwood, Village, Colorado, 2008).

- 498 5. Weir, J. T. & Price, T. D. Limits to speciation inferred from times to secondary sympatry and
499 ages of hybridizing species along a latitudinal gradient. *Am. Nat.* **177**, 462-469 (2011).
- 500 6. Price, T. D. *et al.* Niche filling slows the diversification of Himalayan songbirds. *Nature* **509**,
501 222-225 (2014).
- 502 7. Pigot, A. L. & Tobias, J. A. Species interactions constrain geographic range expansion over
503 evolutionary time. *Ecol. Lett.* **16**, 330-338 (2013).
- 504 8. MacArthur, R. & Levins, R. The limiting similarity, convergence, and divergence of
505 coexisting species. *Am. Nat.*, **101**, 377-385 (1967).
- 506 9. Rundell, R. J. & Price, T. D. Adaptive radiation, nonadaptive radiation, ecological speciation
507 and nonecological speciation. *Trends Ecol. Evol.* **24**, 394-399 (2009).
- 508 10. Schluter, D. in *The Ecology of Adaptive Radiation* (Oxford University Press, Oxford, UK,
509 2000).
- 510 11. Bothwell, E., Montgomerie, R., Loughheed, S. C. & Martin, P. R. Closely related species of
511 birds differ more in body size when their ranges overlap-in warm, but not cool, climates.
512 *Evolution* **69**, 1701-12 (2015).
- 513 12. Tobias, J. A. *et al.* Species coexistence and the dynamics of phenotypic evolution in adaptive
514 radiation. *Nature* **506**, 359-363 (2014).
- 515 13. Pfennig, D. W. & Pfennig, K. S. Character displacement and the origins of diversity. *Am.*
516 *Nat.* **176 Suppl 1**, S26-44 (2010).
- 517 14. Hunt, G. & Rabosky, D. L. Phenotypic evolution in fossil species: pattern and process. *Annu.*
518 *Rev. Earth Planet. Sci.* **42**, 421-441 (2014).
- 519 15. Pennell, M. W., Harmon, L. J. & Uyeda, J. C. Is there room for punctuated equilibrium in
520 macroevolution? *Trends Ecol. Evol.* **29**, 23-32 (2014).
- 521 16. Mayr, E. in *Evolution as a process.* (eds. Huxley, J., Hardy, A.C., & Ford, E.B.) 157-180
522 (Allen & Unwin, 1954).
- 523 17. Rundle, H. D. & Nosil, P. Ecological speciation. *Ecol. Lett.* **8**, 336-352 (2005).
- 524 18. Futuyma, D. J. On the role of species in anagenesis. *Am. Nat.* **130**, 465-473 (1987).
- 525 19. Gingerich, P. Rates of evolution: effects of time and temporal scaling. *Science* **222**, 159-162
526 (1983).
- 527 20. Estes, S. & Arnold, S. J. Resolving the paradox of stasis: Models with stabilizing selection
528 explain evolutionary divergence on all timescales. *Am. Nat.* **169**, 227-244 (2007).
- 529 21. Boag, P. T. & Grant, P. R. Intense natural selection in a population of Darwin's finches
530 (Geospizinae) in the Galapagos. *Science*, **214**, 82-85 (1981).
- 531 22. Uyeda, J. C., Hansen, T. F., Arnold, S. J. & Pienaar, J. The million-year wait for
532 macroevolutionary bursts. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15908-15913 (2011).
- 533 23. Hunt, G., Hopkins, M. J. & Lidgard, S. Simple versus complex models of trait evolution and
534 stasis as a response to environmental change. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4885-4890
535 (2015).
- 536 24. Prothero, D. R. *et al.* Size and shape stasis in late Pleistocene mammals and birds from
537 Rancho La Brea during the Last Glacial-Interglacial cycle. *Quat. Sci. Rev.* **56**, 1-10 (2012).
- 538 25. Rabosky, D. L. & Adams, D. C. Rates of morphological evolution are correlated with species
539 richness in salamanders. *Evolution* **66**, 1807-1818 (2012).
- 540 26. Bokma, F. Time, species, and separating their effects on trait variance in clades. *Syst. Biol.*
541 **59**, 602-607 (2010).
- 542 27. Seddon, N. *et al.* Sexual selection accelerates signal evolution during speciation in birds.
543 *Proc. Roy. Soc. B* **280**, 20131065 (2013).

- 544 28. Cooney, C. R. *et al.* Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature*
545 **542**, 344-347 (2017).
- 546 29. Marshall, C. R. Five palaeobiological laws needed to understand the evolution of the living
547 biota. *Nat. Ecol. Evol.* **1**, 165-017-0165 (2017).
- 548 30. Etienne, R. S., Morlon, H. & Lambert, A. Estimating the duration of speciation from
549 phylogenies. *Evolution* **68**, 2430-2440 (2014).
- 550 31. Kuchta, S. R. & Wake, D. B. Wherefore and whither the ting species? *Copeia* **104**, 189-201
551 (2016).
- 552 32. Futuyma, D. J. in *Macroevolution* 29-85 (Springer, 2015).
- 553 33. Weir, J. T. & Wheatcroft, D. A latitudinal gradient in rates of evolution of avian syllable
554 diversity and song length. *Proc. Roy. Soc. B* **278**, 1713-1720 (2011).
- 555 34. Sullivan, B. L. *et al.* eBird: A citizen-based bird observation network in the biological
556 sciences. *Biol. Conserv.* **142**, 2282-2292 (2009).
- 557 35. Sullivan, B. L. *et al.* The eBird enterprise: An integrated approach to development and
558 application of citizen science. *Biol. Conserv.* **169**, 31-40 (2014).
- 559 36. Burleigh, J. G., Kimball, R. T. & Braun, E. L. Building the avian tree of life using a large-
560 scale, sparse supermatrix. *Mol. Phylogenet. Evol.* **84**, 53-63 (2015).
- 561 37. Baiser, B., Valle, D., Zelazny, Z. & Burleigh, J. G. Non-random patterns of invasion and
562 extinction reduce phylogenetic diversity in island bird assemblages. *Ecography* (2017).
- 563 38. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of
564 birds in space and time. *Nature* **491**, 444-448 (2012).
- 565 39. Hunt, G., Bell, M. A. & Travis, M. P. Evolution toward a new adaptive optimum: phenotypic
566 evolution in a fossil stickleback lineage. *Evolution* **62**, 700-710 (2008).
- 567 40. Arnold, S. J. Phenotypic evolution: the ongoing synthesis (American Society of Naturalists
568 address). *Am. Nat.* **183**, 729-746 (2014).
- 569 41. BirdLife International and NatureServe. in *Bird species distribution maps of the world*
570 (BirdLife International and NatureServe, Cambridge, UK and Arlington, USA, 2014).
- 571 42. Dawideit, B. A., Phillimore, A. B., Laube, I., Leisler, B. & Böhning-Gaese, K.
572 Ecomorphological predictors of natal dispersal distances in birds. *J. Anim. Ecol.* **78**, 388-395
573 (2009).
- 574 43. Claramunt, S., Derryberry, E. P., Remsen, J. V., Jr. & Brumfield, R. T. High dispersal ability
575 inhibits speciation in a continental radiation of passerine birds. *Proc. Roy. Soc. B* **279**, 1567-1574
576 (2012).
- 577 44. Stephens, P. A., Sutherland, W. J. & Freckleton, R. P. What is the Allee effect? *Oikos* **87**,
578 185-190 (1999).
- 579 45. Case, T. J., Holt, R. D., McPeck, M. A. & Keitt, T. H. The community context of species'
580 borders: ecological and evolutionary perspectives. *Oikos* **108**, 28-46 (2005).
- 581 46. Phillimore, A. B. *et al.* Sympatric speciation in birds is rare: Insights from range data and
582 simulations. *Am. Nat.* **171**, 646-657 (2008).
- 583 47. Nosil, P. in *Ecological speciation* (Oxford University Press, 2012).
- 584 48. Landis, M. J., Schraiber, J. G. & Liang, M. Phylogenetic analysis using Levy processes:
585 finding jumps in the evolution of continuous traits. *Syst. Biol.* **62**, 193-204 (2013).
- 586 49. Landis, M. J. & Schraiber, J. G. Pulsed evolution shaped modern vertebrate body sizes. *Proc.*
587 *Natl. Acad. Sci. U. S. A.* (2017).
- 588 50. Harmon, L. J. *et al.* Early Bursts of Body Size and Shape Evolution are Rare in Comparative
589 Data. *Evolution* **64**, 2385-2396 (2010).

- 590 51. Benkman, C. W. Divergent selection drives the adaptive radiation of crossbills. *Evolution* **57**,
591 1176-1181 (2003).
- 592 52. Lamichhaney, S. *et al.* Evolution of Darwin's finches and their beaks revealed by genome
593 sequencing. *Nature* **518**, 371-375 (2015).
- 594 53. Gavrillets, S. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* **12**,
595 307-312 (1997).
- 596 54. Friis, G., Aleixandre, P., Rodríguez Estrella, R., Navarro Sigüenza, A. G. & Milá, B.
597 Rapid postglacial diversification and long-term stasis within the songbird genus Junco:
598 phylogeographic and phylogenomic evidence. *Mol. Ecol.* **25**, 6175-6195 (2016).
- 599 55. Endler, J. A. in *Geographic variation, speciation, and clines (Vol. 10)* (Princeton University
600 Press, 1977).
- 601 56. Futuyma, D. J. Evolutionary constraint and ecological consequences. *Evolution* **64**, 1865-
602 1884 (2010).
- 603 57. Brown, W. L. & Wilson, E. O. Character displacement. *Syst. Zool.* **5**, 49-64 (1956).
- 604 58. Dayan, T. & Simberloff, D. Ecological and community-wide character displacement: the
605 next generation. *Ecol. Lett.* **8**, 875-894 (2005).
- 606 59. Jonathan Davies, T., Meiri, S., Barraclough, T. G. & Gittleman, J. L. Species coexistence
607 and character divergence across carnivores. *Ecol. Lett.* **10**, 146-152 (2007).
- 608 60. Drury, J.P., Grether, G.F., Garland, T., & Morlon, Hélène. An assessment of phylogenetic
609 tools for analyzing the interplay between interspecific interactions and phenotypic evolution. *Sys.*
610 *Biol.* doi: <https://doi.org/10.1093/sysbio/syx079>. (2017).
- 611 61. Roughgarden, J. Competition and theory in community ecology. *Am. Nat.* **122**, 583-601
612 (1983).
- 613 62. Hudson, E. J. & Price, T. D. Pervasive reinforcement and the role of sexual selection in
614 biological speciation. *J. Hered.* **105**, 821-833 (2014).
- 615 63. Pigot, A. L. & Tobias, J. A. Dispersal and the transition to sympatry in vertebrates. *Proc.*
616 *Roy. Soc. B* **282**, 20141929 (2015).
- 617 64. Rosenzweig, M. L. in *Species diversity in space and time* (Cambridge University Press,
618 1995).
- 619 65. Rosenblum, E. B. *et al.* Goldilocks meets Santa Rosalia: an ephemeral speciation model
620 explains patterns of diversification across time scales. *J. Evol. Biol.* **39**, 255-261 (2012).
- 621 66. Dynesius, M. & Jansson, R. Persistence of within-species lineages: a neglected control of
622 speciation rates. *Evolution* **68**, 923-934 (2014).
- 623 67. Sanderson, M. R8s: Inferring absolute rates of molecular evolution and divergence times in
624 the absence of a molecular clock. *Bioinformatics* **19**, 301-302 (2003).
- 625 68. Ksepka, D. & Clarke, J. Phylogenetically vetted and stratigraphically constrained fossil
626 calibrations within Aves. *Palaeontol. Electron.* **18**, 1-25 (2015).
- 627 69. Wilson, D. S. The adequacy of body size as a niche difference. *Am. Nat.* **109**, 769-784
628 (1975).
- 629 70. Handbook of the Birds of the World Alive (eds. Del Hoyo, J., Elliott, A., Sargatal, J.,
630 Christie, D.A. & de Juana, E.). <http://www.hbw.com>. (Lynx Edicions, Barcelona, 2015).
631 Accessed 2015.
- 632 71. Dunning, J. B. in *Body Masses of Birds of the World* (Taylor and Francis Group, Boca Raton,
633 Florida, 2008).
- 634 72. Dunning, J.B. Body Masses of Birds of the World.
635 <https://ag.purdue.edu/fnr/Documents/WeightBookUpdate.pdf>. (2016).

- 636 73. Schoener, T. W. The evolution of bill size differences among sympatric congeneric species
637 of birds. *Evolution* **19**, 189-213 (1965).
- 638 74. Miles, D. B. & Ricklefs, R. E. The correlation between ecology and morphology in
639 deciduous forest passerine birds. *Ecology* **65**, 1629-1640 (1984).
- 640 75. Ho, L. & Ane, C. A linear-time algorithm for Gaussian and non-Gaussian trait evolution
641 models. *Syst. Biol.* **63**, 397-408 (2014).
- 642 76. Stoddard, M. C. *et al.* Avian egg shape: Form, function, and evolution. *Science* **356**, 1249-
643 1254 (2017).
- 644 77. Sullivan, B. L. *et al.* The eBird enterprise: An integrated approach to development and
645 application of citizen science. *Biol. Conserv.* **169**, 31-40 (2014).
- 646 78. Chesser, R. T. & Zink, R. M. Modes of speciation in birds: a test of Lynch's method.
647 *Evolution* **48**, 490-497 (1994).
- 648 79. Barraclough, T. G. & Vogler, A. P. Detecting the geographical pattern of speciation from
649 species-level phylogenies. *Am. Nat.* **155**, 419-434 (2000).
- 650 80. Bull, C. Ecology of Parapatric Distributions. *Annu. Rev. Ecol. Syst.* **22**, 19-36 (1991).
- 651 81. R Core Team. R: A language and environment for statistical computing. (2012).
- 652 82. Calcagno, V. & de Mazancourt, C. glmulti: an R package for easy automated model selection
653 with (generalized) linear models. *Journal of Statistical Software* **34**, 1-29 (2010).
- 654 83. Yasukawa, K. Male quality and female choice of mate in the red-winged blackbird (*Agelaius*
655 *phoeniceus*). *Ecology* **62**, 922-929 (1981).
- 656 84. Grant, P. R. & Grant, B. R. Hybridization, sexual imprinting, and mate choice. *Am. Nat.* **149**,
657 1-28 (1997).
- 658 85. Lovette, I. J. & Hochachka, W. M. Simultaneous effects of phylogenetic niche conservatism
659 and competition on avian community structure. *Ecology* **87**, S14-S28 (2006).
- 660

661 **Acknowledgements**

662 We are grateful to numerous data collectors who contributed to eBird, GenBank, and the CRC
663 bird body mass data set (see Supplementary Information). We also thank Nico Alioravainen, Ed
664 Braun, Samuel Jones, Rebecca Kimball, Dan Ksepka, Monte Neate-Clegg, Alex Pigot, Aaron
665 Ragsdale and Gleb Zhelezov for data collection and technical assistance. We thank three
666 anonymous reviewers and Vera Domingues for constructive comments that improved the
667 manuscript. This work was supported by the National Science Foundation (DEB-1208428 to
668 J.G.B.), the Natural Environment Research Council (NE/I028068/1 to J.A.T.), and the Oxford
669 Clarendon Fund and US-UK Fulbright Commission (to C.S.).

670

671 **Author contributions** J.G.B. and J.P.M. conceived the study; J.G.B, J.P.M. and J.A.T. designed
672 the conceptual framework and analyses; J.G.B. performed dating analyses and assembled
673 phylogenetic, occurrence, and body mass information; J.A.T. and C.S. provided morphometric
674 data; J.P.M. integrated data sets, and designed and performed statistical analyses with significant

675 input from J.G.B., J.A.T., and C.S.; J.P.M. and C.S. produced figures and tables; J.P.M. wrote
676 the manuscript, with significant input from all authors.

677

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

682

683 **Data Availability**

684 We provide the data sets used to perform the analyses in this manuscript in a Microsoft Excel file
685 (.xlsx) in the Supplementary Information.

686

687

688 **Figures and figure legends**

689

690 **Figure 1 | The speciation cycle and phenotypic trait divergence**

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

698

699 **Figure 2 | Tempo of body mass divergence for avian sister species.** Stochastic pulsed models
700 provide better fits to patterns of body mass divergence and divergence time in avian sister
701 species pairs ($n = 869$), with the best fit a single pulse model (ΔAIC relative to the multiple pulse
702 model: 797). Colors denote probability density. The probability density for any time slice
703 follows a normal distribution (most apparent in the white noise model where the probability
704 density distribution is independent of time). Relative probability density can be assessed within
705 each time slice but not across time. For clarity, the empirical data points are plotted only on the
706 white noise model.

707

708 **Figure 3 | Timing of body mass divergence pulses and sympatry.** Comparison of timescales
709 suggests that mass divergence tends to precede sympatry among 952 avian sister species. Yellow
710 and orange lines are cumulative probability distributions of incurring a pulse under the single
711 pulse model (yellow: estimated from maximum likelihood phylogenetic tree³⁶; orange: from 100
712 bootstrap trees). Circles are proportions of sympatric species pairs for 1-million year intervals of
713 divergence time; circle sizes represent sample sizes, numbered where ≥ 50 . For visual
714 comparison, an exponential decay model has been fitted to the proportion of sister pairs in
715 sympatry (blue curve; assumes all sympatry is secondary). If we assume that rates of transition
716 out of sympatry are low, the timing of divergence pulses is early compared to the generally
717 delayed attainment of sympatry (i.e. the yellow and orange curves are steeper than the blue
718 curve).

719

720 **Figure 4 | Phylogenetic patterns of contact and sympatry across avian sister species.**

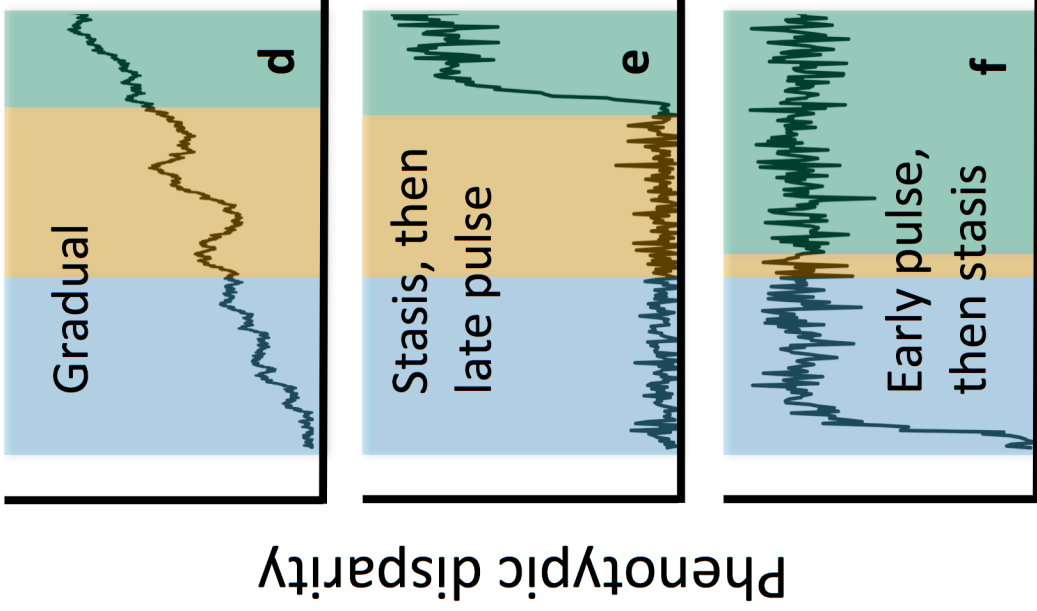
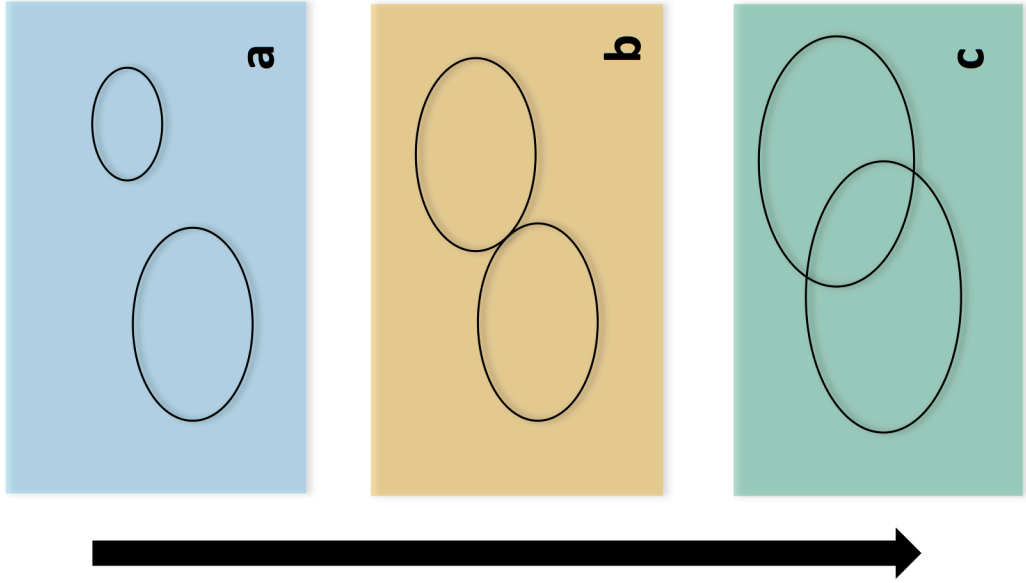
721 Evolutionary and geographic range relationships among 952 pairs of sister species scored for
722 breeding contact (orange) and sympatry (blue), showing that both are widespread in sister
723 species across phylogeny. Pairs with breeding contact include both parapatric and sympatric
724 species pairs. The tree is derived from the Burleigh tree³⁶, and has been pruned. Terminal
725 branches are further pruned such that tips represent the most recent common ancestors of sister
726 species pairs. There are 23 sister pairs that were scored as sympatric from range polygons, but
727 for which no breeding contact records existed in the eBird data set^{34, 76}, likely reflecting a
728 combination of sparse sampling in these sister species' ranges and limited syntopy despite broad
729 geographic overlap.

730

731 **Figure 5 | Factors associated with the establishment of secondary contact and sympatry in**

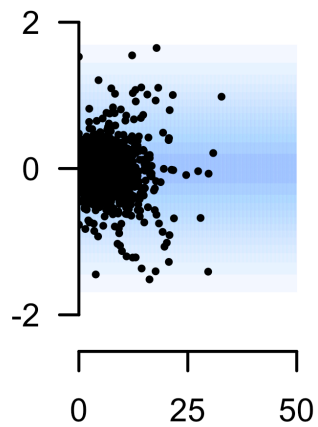
732 **birds.** Results of generalized linear models assessing the relative importance of predictors of
733 breeding contact (**a**, $n = 849$ sister pairs) and sympatry (**b**, $n = 440$ sister pairs with breeding co-
734 occurrence). Pairs with breeding contact include both parapatric and sympatric species pairs.
735 Relative importance is estimated as the proportion of the summed model weights for all models
736 with $\Delta AIC < 2$, and is indicative of the extent to which each variable is necessary to explain the

737 variation. Pairwise interactions with relative importance >0.6 are indicated by the numbers
738 within the bar for each variable.
739

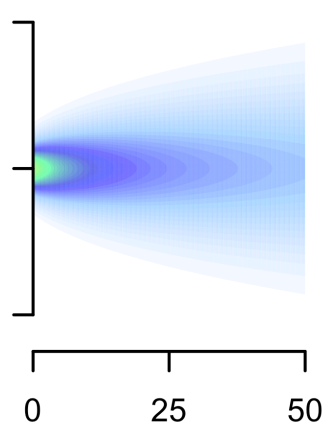


Body mass disparity

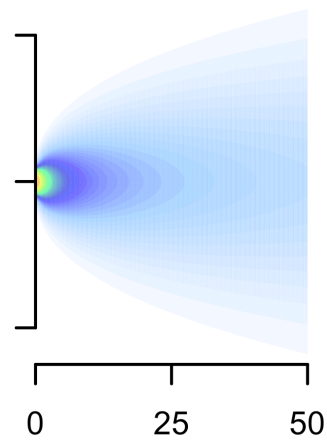
White noise



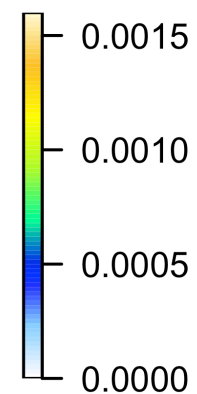
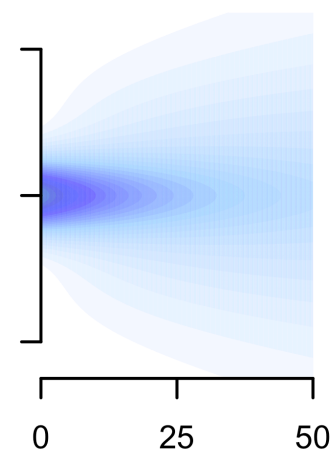
Brownian motion



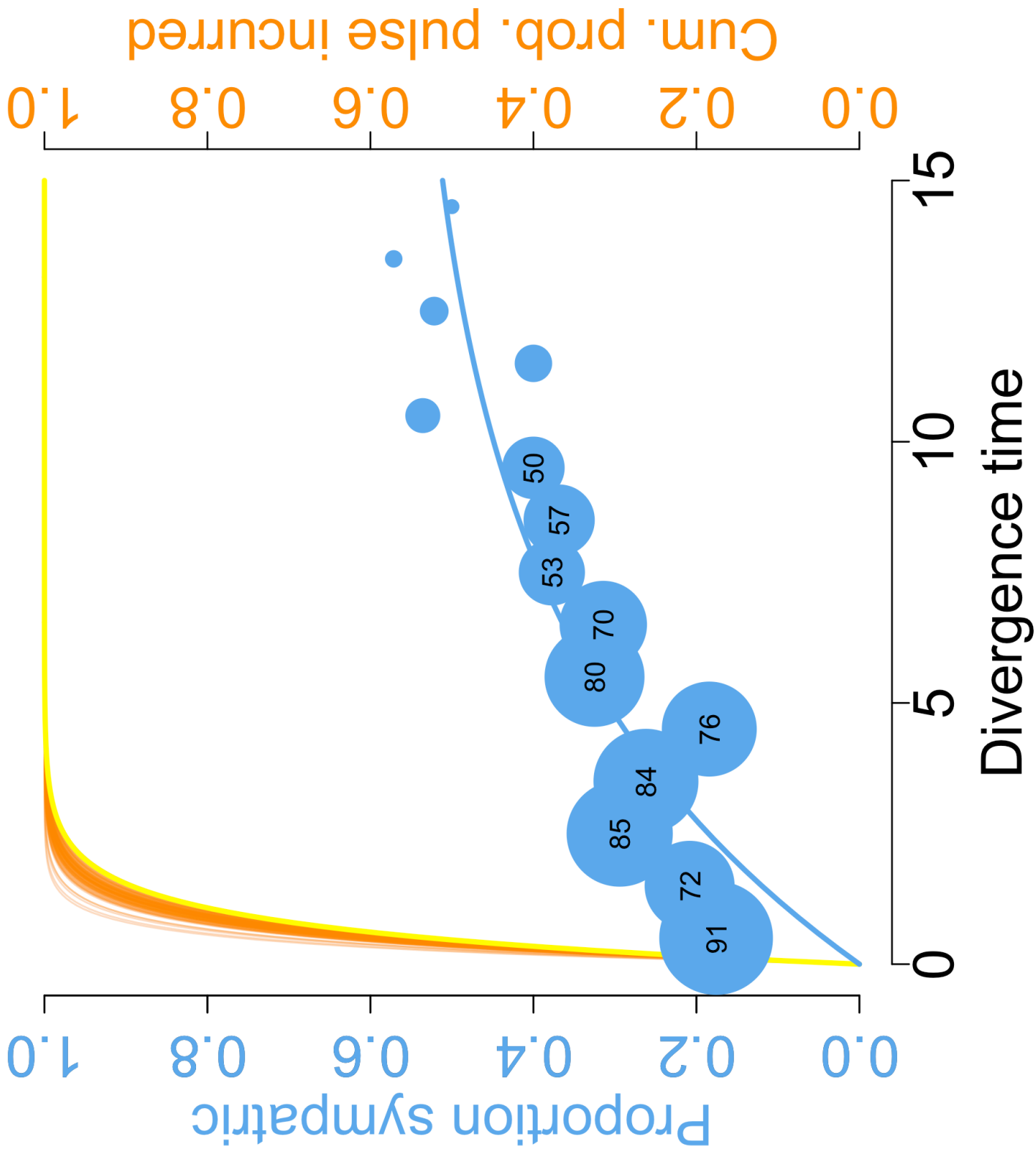
Single pulse



Multiple pulse



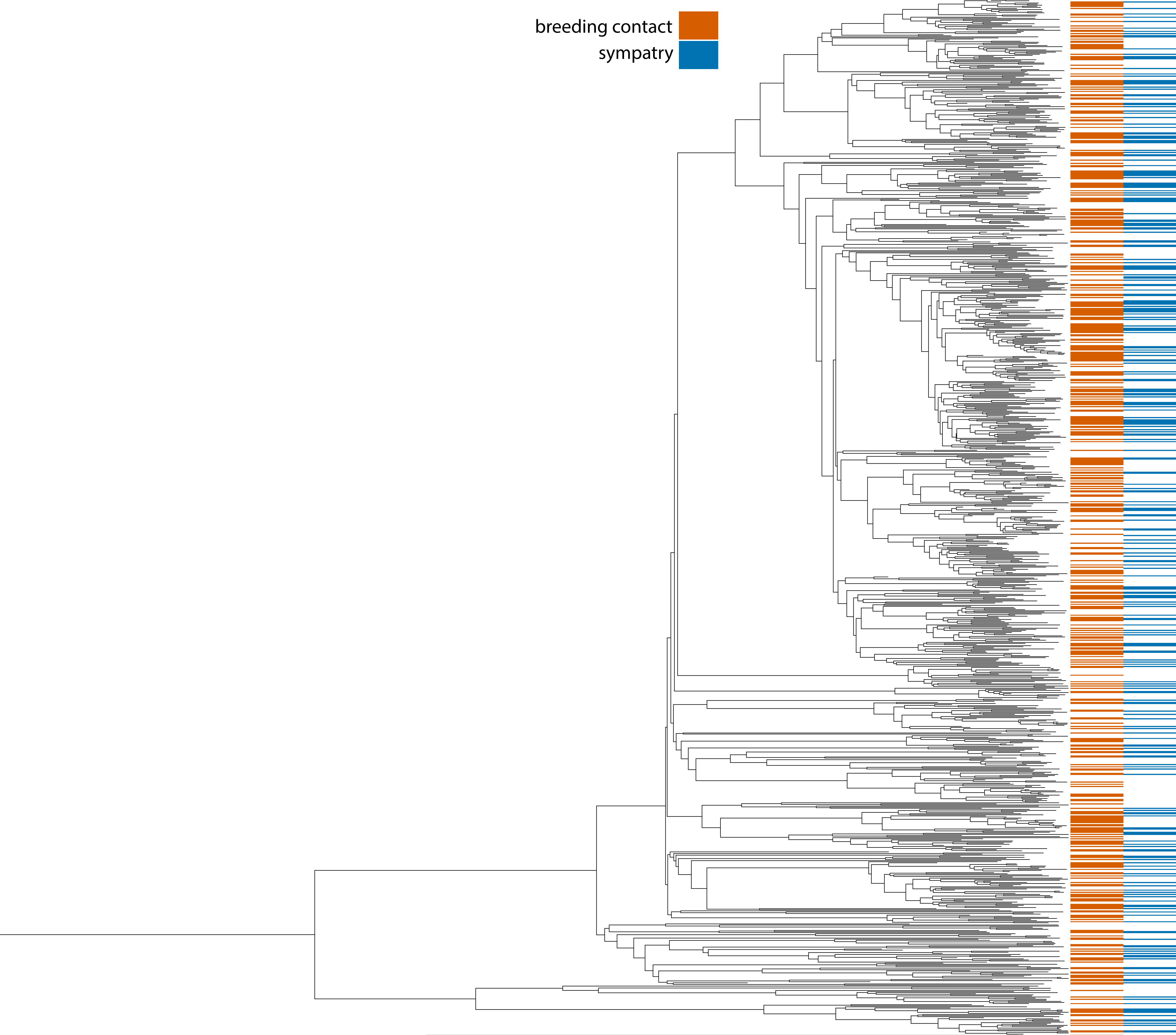
Divergence time (My)



breeding contact

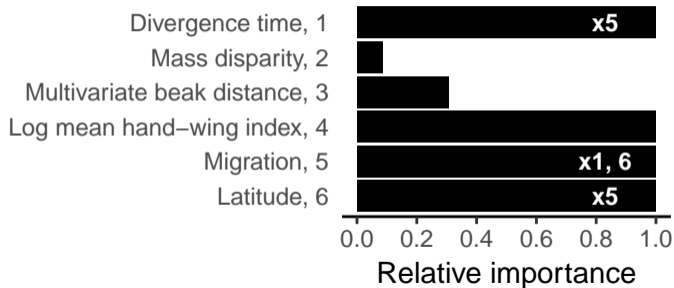


sympatry



20 my

Secondary contact



Sympatry

