Speciation history shapes patterns of assemblage species richness in birds

Bouwe R. Reijenga[1,2]*, Rampal S. Etienne[3], David J. Murrell[2], Alex L. Pigot[2]

[1]Department of Earth Sciences, University of Oxford, UK

[2]Department of Genetics, Environment and Evolution, University College London, UK

[3]Groningen Institute for Evolutionary Life Sciences, University of Groningen, The Netherlands

*Corresponding author: bouwe.reijenga@earth.ox.ac.uk

Abstract

Speciation is the ultimate source of biodiversity but how it impacts patterns of species richness remains unclear. While increasing total richness, co-occurrence need not if speciation generates spatially non-overlapping lineages. Here we apply a dynamic model of speciation and community assembly to phylogenetic and geographic data for 40 passerine bird families to estimate the time required for species to assemble locally and test to which extent this explains variation in assemblage richness. While variable across clades, we estimate an average lag-time from speciation to colonisation of ~8Myr, similar to time between speciation events. Concomitantly, the maximum proportion of co-occurrence is highest in old, species-poor clades, having had more time to accumulate locally. Phylogenetic patterns, previously interpreted as ecological limits to local richness, are consistent with slow, constant colonisation. Our results highlight how time-for-colonisation limits the formation of species assemblages and the strong legacy of speciation history on current biodiversity patterns.

Key words: Species richness, sympatry, coexistence, colonisation, ecological limits., community assembly, time-for-speciation

Introduction

It has long been recognised that the number of species found at a particular place on the Earth's surface likely reflects not only current environmental conditions but also the history of the environment and the lineages that occur there (Francis & Currie 2003; Jetz *et al.* 2004; Mayfield & Levine 2010; Ricklefs & Latham 1993; Webb 2000). This effect of history on the richness of species assemblages operates over a range of spatial and temporal scales, from the macroevolutionary dynamics of speciation and extinction (Kozak & Wiens 2012; Miller *et al.* 2018; Stephens & Wiens 2003; Wiens 2011), to the dispersal of lineages and their adaptation to novel environments (Cracraft 1994; Mayfield & Levine 2010; Patterson 1999; Ricklefs 1987; Webb 2000), to the order in which species colonise a local site (i.e., priority effects) (Reijenga *et al.* 2021; Stroud *et al.* 2024). In some cases, the recency of major disturbance events allows the effect of history on community assembly to be observed directly, such as the recolonisation of plants and animals following the eruption of Krakatau in 1883 (Bush & Whittaker 1991) or the retreat of northern hemisphere ice sheets at the end of the Pleistocene (Blois *et al.* 2010). Over deeper timescales, where community assembly must be inferred, phylogenetic approaches have been used to model the dynamics of species diversification and dispersal within and among biomes, regions and islands (Jetz & Fine 2012; Kozak & Wiens 2012; Miller *et al.* 2018; Quintero *et al.* 2023; Stephens & Wiens 2003; Swiston & Landis 2025; Valente *et al.* 2020; Wiens 2011). While such studies have advanced our understanding of what limits the accumulation of species within large or geographically isolated regions, they do not model community assembly within the smaller spatial grains (e.g., ~25km to ~100km grid cells) at which broad-scale gradients in assemblage richness are typically mapped (Orme *et al.* 2005; Rahbek & Graves 2001). Understanding how history shapes the accumulation of species richness at these finer

spatial grains requires approaches that consider both the origin of new species and how these lineages then come to assemble at a given location.

The origin of most new species involves some form of spatial separation (allopatry or parapatry) that reduces gene flow between populations (Coyne & Orr 2004; Phillimore *et al.* 2008). As a result, cladogenesis becomes increasingly unlikely in geographic areas that are small relative to the spatial scale of organism dispersal (Coyne & Price 2000; Kisel & Barraclough 2010). In birds, for example, cladogenetic speciation is extremely rare or debatable in continuous areas smaller than Madagascar (1600 x 350 km) (Coyne & Price 2000; Phillimore *et al.* 2008). For local assemblages—here defined as the set of species co-occurring at a spatial grain below the limit of cladogenesis—community assembly is thus dependent on the colonisation of species from the surrounding regional pool, as well the counteracting effect of local extinction (MacArthur & Wilson 1967; Mittelbach & Schemske 2015; Pigot & Etienne 2015). While speciation does not add directly to the richness of local assemblages, the history of speciation is still expected to shape patterns of species richness indirectly via its effect on the pool of species available for colonisation. First, differences in the cumulative number of speciation events across regions or clades, whether determined by differences in age (Kozak & Wiens 2012), rates of speciation or ecological limits (Rabosky 2012) will influence the size of the species pools from which local assemblages can be formed (Cornell & Lawton 1992). Second, and more rarely addressed, the timing of past speciation events will determine the age of the species in the pool and thus the time available for colonisation (Stephens & Wiens 2003; Wiens *et al.* 2011).

Previous studies analysing the transition of sister species to secondary sympatry (i.e. geographic range overlap) following speciation have shown that although there are

instances of sympatry among young lineages, most sister species continue to have geographically non-overlapping distributions for millions of years after the initiation of speciation (Alencar & Quental 2023; Pigot & Tobias 2015; Price 2010; Weir & Price 2011). This lag in the attainment of secondary sympatry has been attributed to a variety of intrinsic and extrinsic mechanisms, including limited dispersal across geologically persistent geographic barriers, reproductive interference, and competition for ecological resources (Pigot & Tobias 2015; Price 2010; Sexton *et al.* 2009; Weber & Strauss 2016). Regardless of the mechanism, these studies suggest that an important constraint on the build-up of species within local assemblages is the time available for colonisation following speciation. Where speciation events have occurred deeper in time, lineages will have had more time to overcome the intrinsic and extrinsic barriers to sympatry leading to higher assemblage richness. In contrast, where speciation events have occurred more recently, most lineages will still be in a state of allopatry (or parapatry) and the proportion of species that co-occur locally will be lower (Stephens & Wiens 2003). Given the strong gradients in the rates and temporal dynamics of speciation across geographic space and clades (Harvey *et al.* 2020; Jetz *et al.* 2012; Rabosky *et al.* 2018; Schluter & Pennell 2017), differences in the timing of past speciation events could leave an important imprint on present-day species richness. Yet, a focus on modelling pairwise patterns of range overlap means that the effect of such a 'speciation legacy' in limiting the species richness of entire assemblages has yet to be quantified.

Although directly testing the effect of speciation history on assemblage diversity is challenging, previous studies exploring the links between assemblage species richness and the phylogenetic properties of clades has revealed several intriguing patterns. For example, Rabosky *et al.* (2018) showed that average grid cell rates of speciation in marine fish are fastest at high latitudes where grid cell species richness is lowest. This

same pattern has also been reported for Suboscine birds (Harvey *et al.* 2020), suggesting that a negative relationship between rates of speciation and assemblage richness could be a general trend. Weir (2006) found that the maximum grid cell richness attained by Amazonian bird genera was higher in clades where the rate of diversification inferred from the reconstructed phylogeny had declined over time (but see (Pie *et al.* 2023)). Kennedy *et al.* (2017) reported a strong positive association between total clade richness and the maximum grid cell richness attained across 10 large clades of New World passerine birds. While the causes of these patterns remain unclear, they have generally been interpreted as evidence that local species richness is strongly regulated by the current environment, and that as this ecological limit is reached, this feeds back to dampen further speciation. In other words, speciation rates are fastest where local richness is low and has yet to reach an ecological carrying capacity (e.g., in young clades (Weir 2006) or recently deglaciated high latitudes (Rabosky *et al.* 2018; Harvey *et al.* 2020) and is higher in clades that have access to more niche space (Kennedy *et al.* 2017). Potential support for this idea comes from Price *et al.* (2014), who found that the accumulation of phylogenetic branches representing songbird species co-occurring in a single Himalayan grid cell appears to slow down towards the present, as could be expected if local species richness has reached an ecological limit and further colonisation is prevented. Although an explanation based on ecological limits is plausible, the absence of a process-based model for how species richness is assembled within grid cells, has hindered interpretation of these phylogenetic patterns.

We hypothesise that these patterns can potentially be explained without invoking environmental controls on species richness, if instead local co-occurrence is constrained by the timing of past speciation events and thus opportunities for colonisation. In other words, if there is a delay between speciation and the subsequent expansion of species

ranges leading to secondary sympatry, and if this delay operates on a macroevolutionary timescale (i.e., on the same order as the time between speciation events), then variation across clades in the maximum number of co-occurring species should be predictable based on a clades macroevolutionary history (Mittelbach & Schemske 2015; Pigot & Etienne 2015).

Here we develop an approach to test the effect of speciation history in limiting the build-up of species within local assemblages and, as a case study, apply this approach to predict differences in assemblage richness across 40 family level clades of primarily North and South American passerine birds. First, using phylogenetic data on species branching times and geographic range data to infer species presence and absence within ~50x50km grid cells, we use a maximum likelihood-based framework to infer the dynamics of community assembly leading to present day species assemblages. Our approach accounts for the dynamic nature of species pools and how these have changed over time as each clade has diversified and enables us to estimate the average waiting time to colonise local assemblages following speciation, as well as the rate of local extinction. Second, using our estimates for these key parameters we then conduct a computational experiment, simulating the historical assembly of communities, and test the extent to which differences in the timing of past speciation events can explain observed differences across clades in assemblage species richness. Because we are interested in the extent to which speciation history places an upper limit on the number of co-occurring species we focus on explaining variation in maximum assemblage richness across clades (Kennedy *et al.* 2017, 2018; Kozak & Wiens 2012; Weir 2006). Furthermore, because the richness of local assemblages is constrained by the total number of species in each clade, and this effect of pool size has been widely studied (Figure 1b, Figure S1), here we focus on explaining the maximum proportion of species

148 in each clade that co-occur. The passerine radiations we use are ideal in this respect
149 because they differ widely in both their phylogenetic branching patterns (i.e. speciation
150 history), as well as the maximum proportion of species that co-occur—ranging from
151 12.8 to 83.3% of species for Passerellidae and the Onychorhynchidae respectively
152 (Figure 1). Throughout, we compare the explanatory power of a 'historical' assembly
153 model in which the rate of colonisation (and local extinction) is estimated to that of a
154 'non-historical' null model in which there is no lag-time to colonisation following
155 speciation. Finally, we test the extent to which this 'historical' assembly model can
156 explain patterns in assemblage richness and phylogenetic structure previously
157 interpreted as evidence for ecological limits on community assembly but that we
158 suspected could arise because of speciation history.

159 **Materials and Methods**

160 *Phylogenetic and geographic range data*

161 We compiled a dataset consisting of oscine and suboscine passerines predominantly
162 endemic to the Americas. These clades were chosen because the high overall species
163 richness (n = 2081 species) provides sufficient power to infer the historical dynamics of
164 community assembly, and comprehensive phylogenetic and geographic data is available
165 for almost all species (BirdLife International and Handbook of the Birds of the World
166 2020). The evolutionary relationships between species and divergence times were
167 obtained from two recently published time-calibrated phylogenies containing >95%
168 (oscines) and >98% (suboscines) of described extant species (Barker *et al.* 2015; Harvey
169 *et al.* 2020). Our analysis focuses on how the different speciation histories of clades
170 impact present-day assemblage richness. Consequentially, the two main phylogenies
171 were divided into family level clades with at least five species resulting in 7 oscine and

172 18 suboscine clades. The use of families provided enough clades to analyse variation in
173 patterns of assemblage richness but also large enough clades to reliably calculate
174 phylogenetic metrics.

175 To identify the maximum number and identity of species occurring in each family
176 we used expert delineated range maps (BirdLife International and Handbook of the
177 Birds of the World 2020). In contrast to field inventories which are rarely complete or
178 sampled systematically, expert range maps provide a comprehensive characterisation of
179 species richness patterns albeit at a limited spatial resolution. The phylogenies and
180 range maps we used are based on different taxonomies and so we aligned these by using
181 the taxonomy provided by the phylogenetic trees and merging or splitting species
182 geographic ranges accordingly. Range maps were then extracted onto an equal area grid
183 of hexagonal cells where each cell covers ~2600 km^2 and where adjacent cell centroids
184 are located ~55km from one another (Barnes & Sahr 2017). This is a grain size
185 comparable to those used in macroecological analyses (Jetz & Fine 2012; Pigot *et al.*
186 2016; Rahbek 2005) and is much smaller than the minimum area required for *in situ*
187 speciation to occur in birds (Coyne & Price 2000; Kisel & Barraclough 2010). Thus,
188 species occurring co-occurring in a grid cell must have arrived at that state through
189 colonisation. To ensure our results were not dependent on the particular grain size used,
190 especially in regions of steep environmental and species turnover, we repeated our
191 analyses using a finer grid where adjacent cell centroids were ~10km apart and each
192 cell covers ~96 km^2. For each family-level clade, we identified the grid cell with the
193 highest species richness and recorded the identities of the species present in that cell.
194 Note, that the grid cell with the highest richness could vary across clades. Although we
195 were unable to attain expert range maps for a small number of species (n = 5, equally
196 distributed among clades of > 53 species each) we were still able to include these in our

197 analysis, because auxiliary information (e.g., occurrence records) showed their absence
198 from the grid cell of maximum richness.

199 *Estimating the dynamics of community assembly*

200 To test if speciation legacies can explain differences in assemblage richness across
201 clades we applied and modified the Dynamic Assembly Model of Colonisation, Local
202 Extinction and Speciation, 'DAMOCLES' (Pigot & Etienne 2015). DAMOCLES was
203 originally developed as a null model to evaluate patterns of phylogenetic structure for a
204 single community. DAMOCLES models the assembly of a single community from a
205 dynamically evolving regional pool via the processes of colonisation (γ) and local
206 extinction (μ) (Figure 1). These rates are assumed to be constant over time and across
207 lineages, with the composition of the regional pool at time t provided by the lineages in
208 the reconstructed phylogeny that are extant at time t. Thus, the timing of speciation
209 events through which the regional pool evolves are not modelled directly but are instead
210 obtained from the reconstructed phylogeny. DAMOCLES models the effect of speciation
211 on assemblage composition by assuming that for a parent lineage present in the
212 assemblage, cladogenesis results in one daughter lineage being present and the other
213 being absent from this assemblage. For a lineage absent from the assemblage, both
214 daughter lineages would also be absent (Figure 1b). Given this model, the phylogeny and
215 a vector of the presence (1) and absence (0) of each species in the grid cell of maximum
216 richness we estimated γ and μ using maximum-likelihood.

217 *Testing the effect of speciation history on assemblage richness*

218 To test the effect of speciation history in explaining differences in maximum assemblage
219 richness across clades we explored three different model scenarios. First, in a 'historical
220 global rate' model we estimated a single rate of γ and μ across all clades by maximising

221 the cumulative likelihood. In the 'historical global rate' model, differences in
222 (proportional) assemblage richness across clades arise entirely from differences in their
223 speciation histories and thus time for colonisation, rather than differences in rates of γ
224 and μ. Second, in a 'historical variable rate' model we estimated γ and μ separately for
225 each family, thus relaxing the assumption that all clades share identical rates of
226 colonisation and local extinction. For both scenarios, we condition the likelihood on
227 assemblages with at least 1 species. To avoid local optima in likelihood space, we
228 explored fitting our models using different starting parameter values, optimisers, and
229 ODE solvers (Table S1). Finally, we explored a third 'non-historical' null model, in which
230 we fixed γ at a very high rate ($\gamma = 1000$) and only estimated μ. A value $\gamma = 1000$ equates
231 to an average waiting time from speciation to colonisation of 1000 years. Although
232 arbitrary, this is sufficiently rapid compared the average branch lengths in the
233 phylogenetic trees (2.57 million years), that it captures a scenario in which there is
234 effectively no lag time to colonisation after speciation. Failure to reject this null model
235 would indicate that the richness of species assemblages is simply determined by the
236 ratio between γ and μ and there is no detectable effect of speciation history.

237 We compared the fit of the 'historical global rate', 'historical variable rate' and
238 'non-historical' model using AIC values as well as the ability to reproduce the observed
239 patterns of assemblage richness across clades. Specifically, for each model we simulated
240 the colonisation and local extinction of lineages in our focal assemblages for each family
241 according to the estimated values of γ and μ. Simulations were implemented via a
242 Gillespie algorithm (Gillespie 1977), and modelled from the root of the family
243 phylogenetic tree at $t = 0$ until the present. Under this algorithm the transition between
244 state 0 (locally absent) and state 1 (locally present) was modelled via the occurrence of
245 colonisation ($0 \rightarrow 1$) and local extinction ($1 \rightarrow 0$) events. The root state of 0 and 1 was

246 randomly assigned with equal probability, and the waiting time till the next event (δ)
247 was drawn from an exponential distribution with a mean equal to the sum of all per-
248 lineage rates. At time $t + \delta$, either a colonisation or local extinction event occurred to a
249 randomly drawn lineage with a probability equal to the sum of the respective per-
250 lineage colonisation and local extinction rates. This simulation framework was already
251 incorporated in DAMOCLES, but we rewrote the functions in *Rcpp* (Eddelbuettel &
252 Balamuta 2018). The simulation was repeated 2500 times per clade from which we
253 calculated the 95% confidence interval (95% CI) in the expected assemblage richness,
254 and through parametric bootstrapping, the 95% CI for the parameter estimates (Table
255 S1).

256 *Testing the association between assemblage richness and phylogenetic metrics*

257 Assemblage richness has previously been shown to correlate with various phylogenetic
258 metrics. To test if speciation history could explain these associations we quantified the
259 following metrics for each family: (*i*) the crown age (Myr), (*ii*) Colless' index of
260 imbalance standardised to the 'proportional to distinguishable arrangements' model in
261 order to be independent of tree size and richness, and where higher values indicate
262 greater imbalance (Bortolussi *et al.* 2006; Mooers & Heard 1997), (*iii*) the relative rate of
263 diversification (speciation – extinction) between the first and second half of the
264 phylogeny according to the ρ metric which provides an index of the strength of
265 slowdown ($\rho < 0$) or speed-up ($\rho > 0$) in the rate of diversification over time (Etienne &
266 Rosindell 2012; Janzen & Etienne 2024; Pigot *et al.* 2010), and (*iv*) the mean terminal
267 branch length (mbl) used as a measure of the average age of extant species. We tested
268 the effect of each of these phylogenetic metrics in predicting differences in
269 (proportional) assemblage richness across families using a generalised linear model

270 with a quasibinomial error distribution and a logit link function (Douma & Weedon
271 2019). We did this using both the empirical data and the mean expected (proportional)
272 assemblage richness obtained across simulation runs under our global and variable rate
273 models. Through this we were able to test whether the observed relationships can be
274 explained by speciation history and how potential differences in rates of colonisation
275 and local extinction across clades modulate these patterns. Because calculating
276 phylogenetic properties for very smaller trees becomes less informative, we only
277 considered families with more than five species.

278 Lineage-through-time (LTT) plots indicate the accumulation of branches in a
279 phylogenetic tree towards the present. Deviations in LTT shape from a constant rate null
280 model of species diversification can provide evidence of non-random changes in
281 diversification rate over time. For example, a non-random slowdown in diversification
282 rate could indicate a process of adaptive radiation, where an initial burst of speciation is
283 followed by a slowdown in speciation rate as ecological niches are filled (Etienne *et al.*
284 2012; Etienne & Haegeman 2012). Similar logic has been applied to the LTT plot
285 constructed from only those lineages that are currently sympatric in an assemblage
286 (Price *et al.* 2014; Weir 2006). The appearance of a slowdown in the assemblage-LTT
287 plot could indicate that rates of colonisation have slowed down over time, for example
288 due to the local filling of niche space. However, our expectation was that such a pattern
289 would also be expected due to speciation history: if there is a constant but slow rate of
290 colonisation following speciation then few closely related lineages (and thus branches)
291 will co-occur. To test this, for each family we pruned the phylogeny to only those species
292 that are present in the grid cell of maximum sympatric richness for each clade, and from
293 this constructed a lineage-through-time plot for each focal assemblage. We construct
294 95% confidence intervals via a rank envelope to deal with statistical non-independence

295 (Mrkvička *et al.* 2017; Murrell 2018). Additionally, we calculated ρ for these assemblage-
296 phylogenies, while repeating this for the empirical data, the global rate, variable rate,
297 and non-historical null model. The aim of this analysis was to test whether phylogenetic
298 properties of assemblages previously interpreted as evidence for ecological limits could
299 also be explained by a model accounting for speciation history but in which ecological
300 limits are absent.

301 **Results**

302 *The dynamics of community assembly*

303 Both the 'historical global rate' and 'historical variable rate' models indicate that the lag-
304 time between speciation and the colonisation of assemblage is substantial. For the
305 'historical global rate' model, the estimated rate of colonisation is $\gamma = 0.120$ Myr^{-1},
306 equating to a mean lag time from speciation to colonisation of 8.3 Myr. The estimated
307 local extinction rate is $\mu = 0.178$ Myr^{-1}. The 95% confidence interval for γ is broad
308 ranging from 0.003 to 9 Myr mean lag time, potentially suggesting substantial variation
309 in γ across lineages. Indeed, in the 'historical variable rate' model, we estimate that γ
310 varies between 0.04 to 100 Myr^{-1} and μ from 0 to 187 Myr^{-1} across different families.
311 Much of the variation in rates is driven by species poor clades and the mean lag-time to
312 colonisation weighted for clade richness under the 'historical variable rate' model (7.9
313 Myr) is similar to that of the 'historical global rate' model (8.3 Myr) (Table S1). As such,
314 the 'historical global rate' model ($AIC_{\gamma\mu} = 2333.398$), assuming a single γ and μ across all
315 clades (n = 2 parameters), is statistically supported compared to the 'historical variable
316 rate' model ($AIC_V = 2350.29$) which contains a much larger number of parameters (n =
317 50 parameters, corresponding to a separate γ and μ for each clade). Both the 'historical
318 global rate' ($AIC_{\gamma\mu} = 2333.398$) and 'historical variable rate' model ($AIC_V = 2350.29$)

319 provide a substantially better fit to the data than a 'non-historical' model in which the
320 colonisation occurs rapidly following speciation and where the effect of speciation
321 history is therefore absent (AIC_{NH} = 2379.482).

322 *The effect of speciation history on assemblage richness*

323 Slow rates of colonisation imply an important role of speciation history, and this is
324 supported by posterior simulations using the estimated parameters to predict the
325 maximum proportion of co-occurring species in each clade. For the 'historical' models,
326 the observed proportional assemblage richness fell within the 95% CI of the expected
327 proportions in 21 ('historical global rate') and 24 ('historical variable rate') of the 25
328 clades. In contrast, the 'non-historical' model explains the observed patterns less well,
329 with only 16 out of 25 clades having proportional assemblage richness within the 95%
330 CI of the null model (Figure 3c). However, these simulations also show that much of the
331 variation in assemblage richness across clades remains unexplained. For example, under
332 the 'historical global rate' model, assemblage richness is underestimated for
333 Dendrocolaptidae, and overestimated for Furnariidae, Passerellidae, Rhinocryptidae
334 (Figure 3b).

335 These results were consistent when assemblages were constructed at a smaller
336 grain size of 96 km^2 (Figure S3). For the 'historical global rate' model, the mean lag-time
337 to colonisation was 11.5 Myr (γ = 0.087 Myr^{-1}, μ = 0.119 Myr^{-1}), with this model
338 outperforming (AIC = 2182.164), both the 'historical variable rate' model (mean lag-time
339 = 10 Myr, AIC = 2187.684) and the 'non-historical' null model (AIC = 2236.614)
340 according to both AIC and the ability to predict variation in proportional sympatric
341 richness across clades. This is not necessarily surprising, because even though the
342 spatial scale is substantially smaller, assemblage species richness across spatial scales is

343 comparable (Figure S2). Similar results were also obtained when we used grid cells of
344 mean assemblage richness rather than the cells with the highest species richness (Figure
345 S5).

346 *The association between assemblage richness and phylogenetic metrics*

347 Across avian families, the maximum proportion of species that co-occur increases with
348 clade age and mean terminal branch length and declines with clade richness (Figure 4,
349 Figure S4, Figure S6). There was no effect of phylogenetic imbalance nor family-level ρ
350 on the maximum proportion of co-occurring species. These patterns are consistent with
351 those predicted by both the 'historical global rate' and 'historical variable rate' models.
352 In contrast, the 'non-historical' null model predicts little variation in proportional
353 assemblage richness (Figure 3c) nor the observed correlations with phylogenetic
354 metrics (Figure 4). The predictions of the models fitted to a grid cell with mean-richness
355 are comparable to using the grid cell of maximum richness, but empirical variation in
356 assemblage richness is constrained (Figure S6).

357 We found that assemblage-level phylogenetic trees (i.e., pruned to only consist of
358 those species occurring in the cell of maximum richness), show strong temporal
359 slowdowns in branching rate over time (Figure 5, Figure S7). We found that such strong
360 slowdowns are expected under the 'non-historical' null model in 22/25 clades (except
361 Cardinalidae, Emberizidae and Tyrannidae), suggesting that in most cases they can be
362 explained by the pruning of the phylogenetic tree as this method is more likely to
363 remove more recent branching events (Figure S7). For the 'historical global rate' and
364 'historical variable rate' model, no clades or only a single clade (Calyptomenidae) show
365 stronger slowdowns than expected respectively. Indeed, some clades (Emberizidae,
366 Icteridae, and Pittidae) show weaker slowdowns than expected under the best fitting

367 'historical global rate' model. Thus, we find the slowdown in branching rate observed
368 across assemblage level trees can be explained by a combination of (1) the pruning of
369 the phylogeny that occurs when sampling only the subset of lineages that co-occur and
370 (2) the lag time between speciation and colonisation of the community.

371 **Discussion**

372 The evolutionary history of speciation, extinction and colonisation has been argued to be
373 of fundamental importance in explaining the current richness of species assemblages
374 (Ricklefs 1987) (Mittelbach & Schemske 2015). Here, by expanding upon a dynamic
375 model of community assembly that incorporates the processes of speciation,
376 colonisation, and local extinction, we show that the history of past speciation events has
377 left a clear lecacy in the number of species co-occurring within local assemblages of
378 passerine birds.

379 The model of community assembly that we develop incorporates the historical
380 legacy of speciation by assuming that upon speciation daughter lineages do not co-occur
381 and only attain sympatry later, via colonisation of the same locations. This assumption is
382 supported by the many studies demonstrating the pre-ponderance of allopatric (or
383 parapatric) speciation in animals and plants and that sympatric speciation happens only
384 rarely (Coyne & Orr 2004; Coyne & Price 2000; Olivares *et al.* 2025; Phillimore *et al.*
385 2008). Using this model, we were able to estimate the rate of colonisation using
386 maximum likelihood and show that the average lag time (i.e., $1/\gamma$) between the
387 speciation of a lineage and its colonisation of the focal assemblage within a clade is ~8
388 Myr. This long waiting time to colonisation is comparable to the average per-lineage
389 waiting time between speciation events: the colonisation rate across our sample of
390 passerine families is $\gamma = 0.12$ Myr^{-1} whereas the mean rate of species diversification

391 estimated by Jetz *et al.* (2012) across the global bird phylogeny is 0.16 species Myr^{-1}.
392 While stochastic variation in the timing of colonisation will mean that by chance some
393 species rapidly co-occur following speciation, the limited time elapsed since past
394 speciation events thus places a strong constraint on the build-up of species richness
395 within assemblages. We demonstrate that this model incorporating the historical legacy
396 of speciation outperforms a 'non-historical' null model in which there is no effect of
397 speciation history on current assemblage richness.

398 Explanations of species richness patterns have often categorised as either driven
399 by the current environment or accidents of history (Buckley & Jetz 2007; Hawkins *et al.*
400 2003). While the current environment is expected to lead to predictable patterns of
401 richness that are repeated across independent evolutionary arenas, historical factors are
402 generally regarded as degrading this predictability leading to diversity deviations and
403 anomalies (Jetz & Fine 2012; Ricklefs & He 2016). Here we show that the history of
404 speciation leads to predictable variation in assemblage richness across clades with
405 different phylogenetic properties, with these effects driven by differences in species age.
406 In particular, the maximum proportion of species that co-occur increases with clade age
407 because older clades contain older species providing more time for colonisation. In
408 contrast, while larger clades have more speciose assemblages, the proportion of species
409 that co-occur is lower than in smaller clades. This is because, large clades will tend to be
410 those that have experienced many recent speciation events, leading to younger species
411 and less time for colonisation.

412 The finding that species age and 'time for colonisation' is a strong predictor of
413 assemblage species richness bears both similarities and differences to what has been
414 termed the 'time for speciation' effect (Hutter *et al.* 2013; Stephens & Wiens 2003).

415 According to the 'time for speciation' effect, clades which earlier colonise a region have
416 more time to speciate and thus accumulate more species. In other words, while the time
417 for speciation effect focusses on explaining the size of the species pool, our model
418 reveals the importance of time in allowing species within these pools to assemble
419 locally. Specifically, our results show that it is the timing of speciation events within a
420 clade, rather than clade age itself, which constrains assemblage richness.

421 Macroecological studies have variously used phylogenetic properties,
422 diversification rates, and diversity measures to infer ecological processes such as
423 competition from biogeographic patterns. For example, Weir (2006) showed that clades
424 of Amazonian birds showing stronger slowdowns in diversification rate over time
425 measured by the γ-statistic (Pybus & Harvey 2000) had higher absolute levels of
426 assemblage richness, and interpreted this as evidence that niche space within
427 assemblages was saturated, inhibiting further range expansions and speciation (Weir
428 2006). We find no relationship between clade slowdowns and the proportion of species
429 that co-occur locally in either the empirical or the simulated data (Figure 4). We argue
430 that this distinction is caused by the inherent positive relationship between clade and
431 assemblage richness (Figure S1), and the dependence of the γ-statistic on tree size (Pigot
432 *et al.* 2010). Our results further support this by finding negative relationships between
433 the γ-statistic and assemblage richness for empirical and simulated data without
434 invoking ecological limits (Figure S8). This does not rule out the possibility that
435 ecological limits have caused a slowdown in diversification rate (Rabosky & Hurlbert
436 2015), but shows that relationships between phylogenetic metrics and local assemblage
437 richness should be evaluated against models incorporating the geographic context of

438 speciation before being interpreted as evidence for or against ecological limits to species
439 richness.

440 Another pattern where our model provides an alternative explanation regards
441 the appearance of temporal slowdowns in the rate of branching among co-occurring
442 species. While it is recognised that a slowdown in a local lineage through time plot could
443 potentially arise as a result of incomplete sampling of a phylogeny (Cusimano & Renner
444 2010), the appearance of a very strong slowdown could provide evidence that rates of
445 colonisation have slowed down over time as niches are filled locally (McPeek 2008;
446 Price *et al.* 2014). However, our results provide a further non-ecological explanation for
447 such strong slowdowns, showing that they are expected to arise even when rates of
448 colonisation are constant over time because of the geographic context of speciation.
449 Specifically, for a given number of species, allopatric speciation results in fewer closely
450 related lineages co-occurring than expected under random sampling, leading to the
451 appearance of a slowdown in how sympatric lineages have accumulated over time. This
452 slowdown is, however, illusory. While it was recognised more than a decade ago that the
453 appearance of slowdowns in diversification rates could arise because speciation takes
454 time (i.e. protracted speciation (Etienne & Rosindell 2012)), here we show that the
455 appearance of slowdowns in community phylogenies can arise because speciation
456 generates spatially isolated lineages and colonisation of the same communities takes
457 time.

458 While our results highlight the importance of time in limiting the build-up of
459 assemblage richness we caution against interpreting this as evidence that ecological
460 limits are not involved in regulating species richness (Rabosky 2012; Rabosky & Benson
461 2021; Stadler *et al.* 2014). Previous studies focussing on sister species have provided

evidence that in addition to dispersal limitation, competition for ecological resources and incomplete reproductive isolation are important rate limiting factors in the transition to secondary sympatry (Pigot *et al.* 2018; Weir & Price 2011). Thus, the slow colonisation of local assemblages could also reflect limited ecological opportunity. Our findings also clearly show that time for colonisation is not the only process shaping the richness of local assemblages. Assemblage richness of some clades are consistently over or under-predicted by our 'historical global rate' model, and accurately predicting assemblage richness requires a model allowing for different rates of colonisation and local extinction across clades ('historical variable rate' model). Finally, our study also does not explain the processes regulating the dynamics of speciation within clades nor the variation in the total species richness of different avian families, both of which could ultimately be driven by locally or regional imposed ecological limits to species richness (Rabosky 2013; Reijenga *et al.* 2021; Storch & Okie 2019). Given other sources of evidence for ecological controls on diversity (e.g., the stability of local species richness over geological time (Close *et al.* 2019; Close & Reijenga 2025)), we expect patterns of species richness to reflect a combination of ecological controls and historical legacies. A key next step will therefore be to develop a diversity-dependent model to evaluate if the fit to the empirical data improves significantly when assuming ecological limits (Valente *et al.* 2015).

Our results show that the timing of past speciation events leaves a strong and predictable signature in the proportion of species that co-occur within local assemblages. However, in addition to the difference in maximum assemblage richness studied here, local richness also varies substantially within the distribution of a single clade. Furthermore, our model assumes that all lineages are governed by identical rates of colonisation and local extinction and that only time since speciation (and chance)

matters for determining which and how many lineages are present in any assemblage. We explored the effects of relaxing this assumption by allowing different rates of colonisation and local extinction for each clade, but such a model contains many parameters and provides little insight into the potential cause of rate heterogeneity. Future studies could advance beyond this by modelling how rates of colonisation (and local extinction) within and between clades depend on the local environment (e.g., ecosystem productivity)(Pigot *et al.* 2016), the traits of species (e.g., dispersal ability), their ecological differences (Reijenga *et al.* 2023) and arrival order (Fukami 2015; Reijenga *et al.* 2021). While the simple models we have employed here enable the spatial legacy of speciation on assemblage richness to be isolated, a comprehensive understanding of the role of evolutionary history in community assembly will require more complex models that account for the interactions between species and the potential feedbacks between local assemblages and the dynamics of diversification and niche evolution within the species pool.

Acknowledgements

We thank James Rosindell and Luis Valente for comments and discussion on an earlier version of this work.

Funding

ALP and BRR thank the Royal Society for funding through a Royal Society Research Fellowship awarded to ALP and studentship to BRR.

Conflict of interest statement

The authors declare no conflict of interest.

References

Alencar, L.R.V. & Quental, T.B. (2023). Geographical and ecological drivers of coexistence dynamics in squamate reptiles. *Glob. Ecol. Biogeogr.*, 32, 1937–1951.

Barker, K., Burns, K.J., Klicka, J., Lanyon, S.M. & Lovette, I.J. (2015). New insights into New World biogeography: An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *Auk*, 132, 333–348.

Barnes, R. & Sahr, K. (2017). dggridR: Discrete Global Grids for R. R package version.

BirdLife International and Handbook of the Birds of the World. (2020). Bird species distribution maps of the world. Version 2020.1.

Blois, J.L., McGuire, J.L. & Hadly, E.A. (2010). Small mammal diversity loss in response to late-Pleistocene climatic change. *Nature*, 465, 771–774.

Bortolussi, N., Durand, E., Blum, M. & François, O. (2006). apTreeshape: Statistical analysis of phylogenetic tree shape. *Bioinformatics*, 22, 363–364.

Buckley, L.B. & Jetz, W. (2007). Environmental and historical constraints on global patterns of amphibian richness. *Proc. R. Soc. B Biol. Sci.*, 274, 1167–1173.

Bush, M.B. & Whittaker, R.J. (1991). Krakatau: Colonization Patterns and Hierarchies. *J. Biogeogr.*, 18, 341.

Close, R.A., Benson, R.B.J., Alroy, J., Behrensmeyer, A.K., Benito, J., Carrano, M.T., *et al.* (2019). Diversity dynamics of Phanerozoic terrestrial tetrapods at the local-community scale. *Nat. Ecol. Evol.*, 3, 590–597.

Close, R.A. & Reijenga, B.R. (2025). Tetrapod species–area relationships across the Cretaceous–Paleogene mass extinction. *Proc. Natl. Acad. Sci.*, 122, e2419052122.

Cornell, H.V. & Lawton, J.H. (1992). Species Interactions, Local and Regional Processes, and Limits to the Richness of Ecological Communities: A Theoretical Perspective. *J. Anim. Ecol.*, 61, 1.

Coyne, J.A. & Orr, H.A. (2004). *Speciation*. Sinauer Associates, Inc, Sunderland, MA.

Coyne, J.A. & Price, T.D. (2000). Little evidence for sympatric speciation in island birds. *Evolution*, 54, 2166–2171.

Cracraft, J. (1994). Species Diversity, Biogeography, and the Evolution of Biotas. *Am. Zool.*, 34, 33–47.

Cusimano, N. & Renner, S.S. (2010). Slowdowns in diversification rates from real phylogenies may not be real. *Syst. Biol.*, 59, 458–464.

Douma, J.C. & Weedon, J.T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods Ecol. Evol.*, 10, 1412–1430.

Eddelbuettel, D. & Balamuta, J.J. (2018). Extending R with C++: A Brief Introduction to Rcpp. *Am. Stat.*, 72, 28–36.

Etienne, R.S. & Haegeman, B. (2012). A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am. Nat.*, 180.

Etienne, R.S., Haegeman, B., Stadler, T., Aze, T., Pearson, P.N., Purvis, A., *et al.* (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. Biol. Sci.*, 279, 1300–9.

Etienne, R.S. & Rosindell, J. (2012). Prolonging the past counteracts the pull of the present: Protracted speciation can explain observed slowdowns in diversification. *Syst. Biol.*, 61, 204–213.

Francis, A.P. & Currie, D.J. (2003). A Globally Consistent Richness-Climate Relationship for Angiosperms. *Am. Nat.*, 161, 523–536.

Fukami, T. (2015). Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu. Rev. Ecol. Evol. Syst.*, 46, 1–23.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81, 2340–2361.

Harvey, M.G., Bravo, G.A., Claramunt, S., Cuervo, A.M., Derryberry, G.E., Battilana, J., *et al.* (2020). The evolution of a tropical biodiversity hotspot. *Science*, 370, 1343–1348.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.-F., Kaufman, D.M., *et al.* (2003). ENERGY, WATER, AND BROAD-SCALE GEOGRAPHIC PATTERNS OF SPECIES RICHNESS. *Ecology*, 84, 3105–3117.

Hutter, C.R., Guayasamin, J.M. & Wiens, J.J. (2013). Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. *Ecol. Lett.*, 16, 1135–1144.

Janzen, T. & Etienne, R.S. (2024). Phylogenetic tree statistics: A systematic overview using the new R package 'treestats.' *Mol. Phylogenet. Evol.*, 200, 108168.

Jetz, W. & Fine, P.V.A. (2012). Global gradients in vertebrate diversity predicted by historical area-productivity dynamics and contemporary environment. *PLoS Biol.*, 10.

Jetz, W., Rahbek, C. & Colwell, R.K. (2004). The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecol. Lett.*, 7, 1180–1191.

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

Kennedy, J.D., Price, T.D., Fjeldså, J. & Rahbek, C. (2017). Historical limits on species co-occurrence determine variation in clade richness among New World passerine birds. *J. Biogeogr.*, 44, 736–747.

Kennedy, J.D., Rahbek, C., Borregaard, M.K., Marki, P.Z., Machac, A. & Fjeldså, J. (2018). Expansion in geographical and morphological space drives continued lineage diversification in a global passerine radiation. *Proc. R. Soc. B*, 285.

Kisel, Y. & Barraclough, T.G. (2010). Speciation Has a Spatial Scale That Depends on Levels of Gene Flow. *Am. Nat.*, 175, 316–334.

Kozak, K.H. & Wiens, J.J. (2012). Phylogeny, ecology, and the origins of climate–richness relationships. *Ecology*, 93, 167–181.

MacArthur, R.H. & Wilson, E.O. (1967). *The theory of island biogeography*. Princeton university press.

Mayfield, M.M. & Levine, J.M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol. Lett.*, 13, 1085–1093.

McPeek, M.A. (2008). The Ecological Dynamics of Clade Diversification and Community Assembly. *Am. Nat.*, 172, E270–E284.

Miller, E.C., Hayashi, K.T., Song, D. & Wiens, J.J. (2018). Explaining the ocean's richest biodiversity hotspot and global patterns of fish diversity. *Proc. R. Soc. B Biol. Sci.*, 285.

Mittelbach, G.G. & Schemske, D.W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends Ecol. Evol.*, 30, 241–247.

Mooers, A.O. & Heard, B. (1997). Inferring evolutionary process from phylogenetic tree shape, 72, 31–54.

Mrkvička, T., Myllymäki, M. & Hahn, U. (2017). Multiple Monte Carlo testing, with applications in spatial point processes. *Stat. Comput.*, 27, 1239–1255.

Murrell, D.J. (2018). A global envelope test to detect non-random bursts of trait evolution. *Methods Ecol. Evol.*, 9, 1739–1748.

Olivares, I., Faurby, S., Cámara-Leret, R. & Pigot, A.L. (2025). The likelihood of sympatric speciation and morphological divergence in plants. *Proc. Natl. Acad. Sci.*, 122, e2508958122.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., *et al.* (2005). Global hotspots of species richness are not congruent with endemism or threat. *Nature*, 436, 1016–1019.

Patterson, B.D. (1999). Contingency and Determinism in Mammalian Biogeography: The Role of History. *J. Mammal.*, 80, 345–360.

Phillimore, A.A.B., Orme, C.D.C.D.L., Thomas, G.G.H., Blackburn, T.T.M., Bennett, P.M.P., Gaston, K.K.J., *et al.* (2008). Sympatric Speciation in Birds Is Rare: Insights from Range Data and Simulations. *Am. Nat.*, 171, 646–657.

Pie, M.R., Divieso, R. & Caron, F.S. (2023). Clade density and the evolution of diversity-dependent diversification. *Nat. Commun.*, 14, 4576.

Pigot, A.L. & Etienne, R.S. (2015). A new dynamic null model for phylogenetic community structure. *Ecol. Lett.*, 18, 153–163.

Pigot, A.L., Jetz, W., Sheard, C. & Tobias, J.A. (2018). The macroecological dynamics of species coexistence in birds. *Nat. Ecol. Evol.*, 2, 1112–1119.

Pigot, A.L., Phillimore, A.B., Owens, I.P.F. & Orme, C.D.L. (2010). The shape and temporal dynamics of phylogenetic trees arising from geographic speciation. *Syst. Biol.*, 59, 660–673.
Pigot, A.L. & Tobias, J.A. (2015). Dispersal and the transition to sympatry in vertebrates. *Proc. R. Soc. B Biol. Sci.*, 282, 20141929.
Pigot, A.L., Tobias, J.A. & Jetz, W. (2016). Energetic Constraints on Species Coexistence in Birds. *PLoS Biol.*, 14.
Price, T.D. (2010). The roles of time and ecology in the continental radiation of the old world leaf warblers (Phylloscopus and seicercus). *Philos. Trans. R. Soc. B Biol. Sci.*, 365, 1749–1762.
Price, T.D., Hooper, D.M., Buchanan, C.D., Johansson, U.S., Tietze, D.T., Alström, P., *et al.* (2014). Niche filling slows the diversification of Himalayan songbirds. *Nature*, 509, 222–225.
Pybus, O.G. & Harvey, P.H. (2000). Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B Biol. Sci.*, 267, 2267–2272.
Quintero, I., Landis, M.J., Jetz, W. & Morlon, H. (2023). The build-up of the present-day tropical diversity of tetrapods. *Proc. Natl. Acad. Sci.*, 120, e2220672120.
Rabosky, D.L. (2012). Testing the time-for-speciation effect in the assembly of regional biotas. *Methods Ecol. Evol.*, 3, 224–233.
Rabosky, D.L. (2013). Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annu. Rev. Ecol. Evol. Syst.*, 44, 481–502.
Rabosky, D.L. & Benson, R.B.J. (2021). Ecological and biogeographic drivers of biodiversity cannot be resolved using clade age-richness data. *Nat. Commun.*, 12, 1–10.
Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., *et al.* (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559, 392–395.
Rabosky, D.L. & Hurlbert, A.H. (2015). Species richness at continental scales is dominated by ecological limits. *Am. Nat.*, 185, 572–583.
Rahbek, C. (2005). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecol. Lett.*, 8, 224–239.
Rahbek, C. & Graves, G.R. (2001). Multiscale assessment of patterns of avian species richness. *Proc. Natl. Acad. Sci.*, 98, 4534–4539.
Reijenga, B.R., Freeman, B.G., Murrell, D.J. & Pigot, A.L. (2023). Disentangling the historical routes to community assembly in the global epicentre of biodiversity. *Glob. Ecol. Biogeogr.*, geb.13728.
Reijenga, B.R., Murrell, D.J. & Pigot, A.L. (2021). Priority effects and the macroevolutionary dynamics of biodiversity. *Ecol. Lett.*, 24, 1455–1466.
Ricklefs, R.E. (1987). Community diversity: relative roles of local and regional processes. *Science*, 235, 167–171.
Ricklefs, R.E. & He, F. (2016). Region effects influence local tree species diversity. *Proc. Natl. Acad. Sci. U. S. A.*, 113, 674–679.
Ricklefs, R.E. & Latham, R.E. (1993). Global patterns of diversity in mangrove floras. In: *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds. Ricklefs, R.E. & Schluter, D.). pp. 215–229.
Schluter, D. & Pennell, M.W. (2017). Speciation gradients and the distribution of biodiversity. *Nature*, 546, 48–55.
Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and Ecology of Species Range Limits. *Annu. Rev. Ecol. Evol. Syst.*, 40, 415–436.
Stadler, T., Rabosky, D.L., Ricklefs, R.E. & Bokma, F. (2014). On age and species richness of higher taxa. *Am. Nat.*, 184, 447–455.
Stephens, P.R. & Wiens, J.J. (2003). Explaining species richness from continents to communities: The time-for-speciation effect in emydid turtles. *Am. Nat.*, 161, 112–128.
Storch, D. & Okie, J.G. (2019). The carrying capacity for species richness. *Glob. Ecol. Biogeogr.*, 28, 1519–1532.
Stroud, J.T., Delory, B.M., Barnes, E.M., Chase, J.M., De Meester, L., Dieskau, J., *et al.* (2024). Priority effects transcend scales and disciplines in biology. *Trends Ecol. Evol.*, 39, 677–688.

Swiston, S.K. & Landis, M.J. (2025). Testing Relationships between Multiple Regional Features and Biogeographic Processes of Speciation, Extinction, and Dispersal. *Syst. Biol.*, 74, 282–300.

Valente, L., Phillimore, A.B., Melo, M., Warren, B.H., Clegg, S.M., Havenstein, K., *et al.* (2020). A simple dynamic model explains the diversity of island birds worldwide. *Nature*, 579, 92–96.

Valente, L.M., Phillimore, A.B. & Etienne, R.S. (2015). Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecol. Lett.*, 18, 844–852.

Webb, C.O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *Am. Nat.*, 156, 145–155.

Weber, M.G. & Strauss, S.Y. (2016). Coexistence in Close Relatives: Beyond Competition and Reproductive Isolation in Sister Taxa. *Annu. Rev. Ecol. Evol. Syst.*, 47, 359–381.

Weir, J.T. (2006). Divergent Timing and Patterns of Species Accumulation in Lowland and Highland Neotropical Birds. *Evolution*, 60, 842.

Weir, J.T. & Price, T.D. (2011). Limits to Speciation Inferred from Times to Secondary Sympatry and Ages of Hybridizing Species along a Latitudinal Gradient. *Am. Nat.*, 177, 462–469.

Wiens, J.J. (2011). The causes of species richness patterns across space, time, and clades and the role of"ecological limits". *Q. Rev. Biol.*, 86, 75–96.

Wiens, J.J., Pyron, R.A. & Moen, D.S. (2011). Phylogenetic origins of local-scale diversity patterns and the causes of Amazonian megadiversity. *Ecol. Lett.*, 14, 643–652.

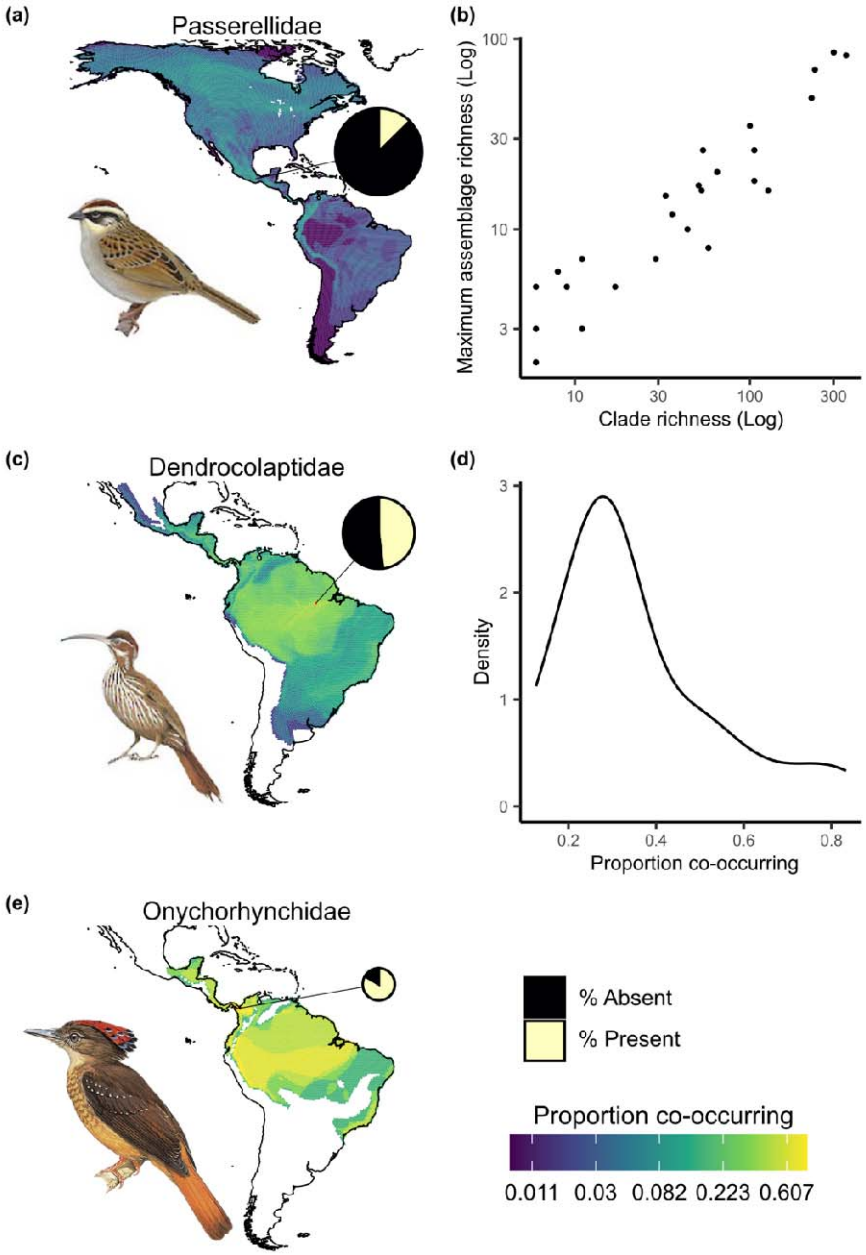


Figure 1 Species richness patterns of exemplary clades. (a, c, and d) assemblage species richness (as a proportion of clade richness) of three exemplar avian families showing low, intermediate, and high levels of co-occurrence. The size of the pie charts shows clade richness, and the black lines point towards the grid cell of highest richness. (b) the relationship between clade richness and maximum local species richness (n =40 families). An ordinary least squares regression estimates an intercept of -0.257 and slope of 0.761. (d) distribution of maximum (proportional) local species richness across clades.

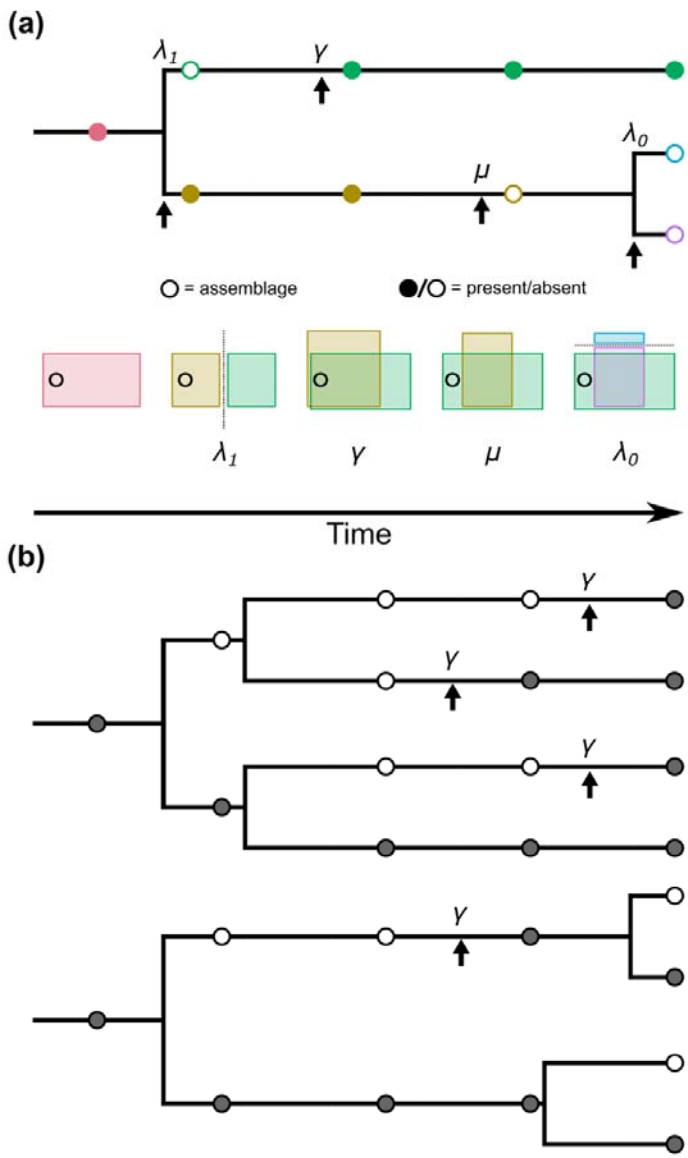
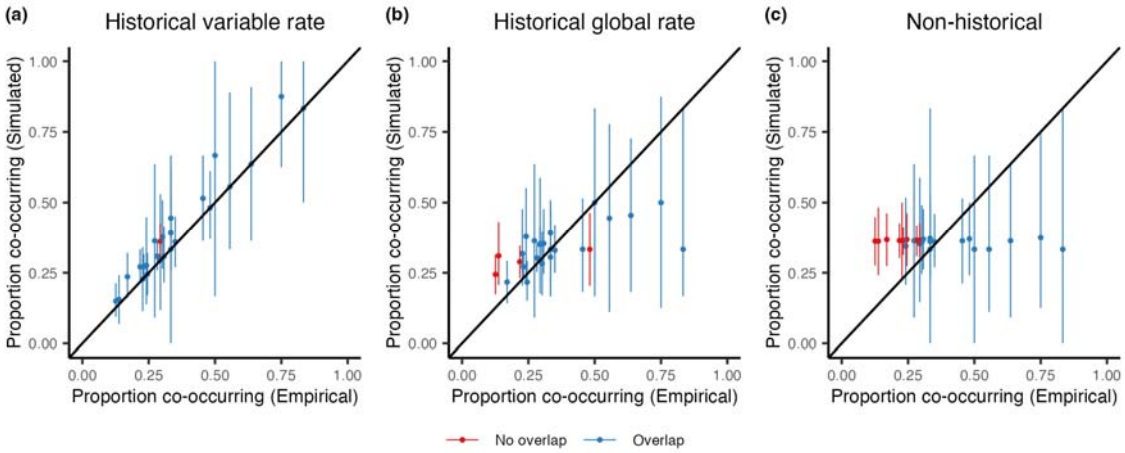


Figure 2 Visualisation of how speciation history influence assemblage richness. (a) conceptualisation of the link between allopatric speciation, range expansion and assemblage richness. Upward-pointing arrows show timing of events. When speciation occurs the ancestral geographic range (red rectangle) is divided into two (yellow and green), and at most one species will be present (λ_1) in the local assemblage (circle). If the ancestor is not present, neither of the descendants (blue and purple) will be present (λ_0) in the local assemblage. The build-up of local assemblage thus depends on the colonisation of species (γ), as well as local extinction (μ), as species undergo expansion (green) or contractions (yellow) of their geographic range. (b) two hypothetical clades, each with four species, with the top clade having older species on average than the bottom clade. Filled circles highlight the presence of a lineage in the local assemblage,

701 and arrows indicate colonisation events. For both clade the ancestor was present in the assemblage, but
702 the dynamics of allopatric speciation and colonisation result in a higher assemblage richness at present in
703 the clade with older species.



704

705 **Figure 3** Observed variation in assemblage richness of passerine clades compared to model predictions.
706 Assemblage richness is plotted as a % of the total number of species in each clade. Model predictions are
707 shown for (a) the 'historical variable rate', (b) 'historical global rate', and (c) the 'non-historical' model.
708 The 1:1 line shows where the empirical the predicted richness matches. Bars represent the 95%
709 confidence intervals of predicted assemblage richness from 2500 replicate simulations of each model.
710 Colours indicate if empirical assemblages fall within (blue) or outside (red) of model expectations.

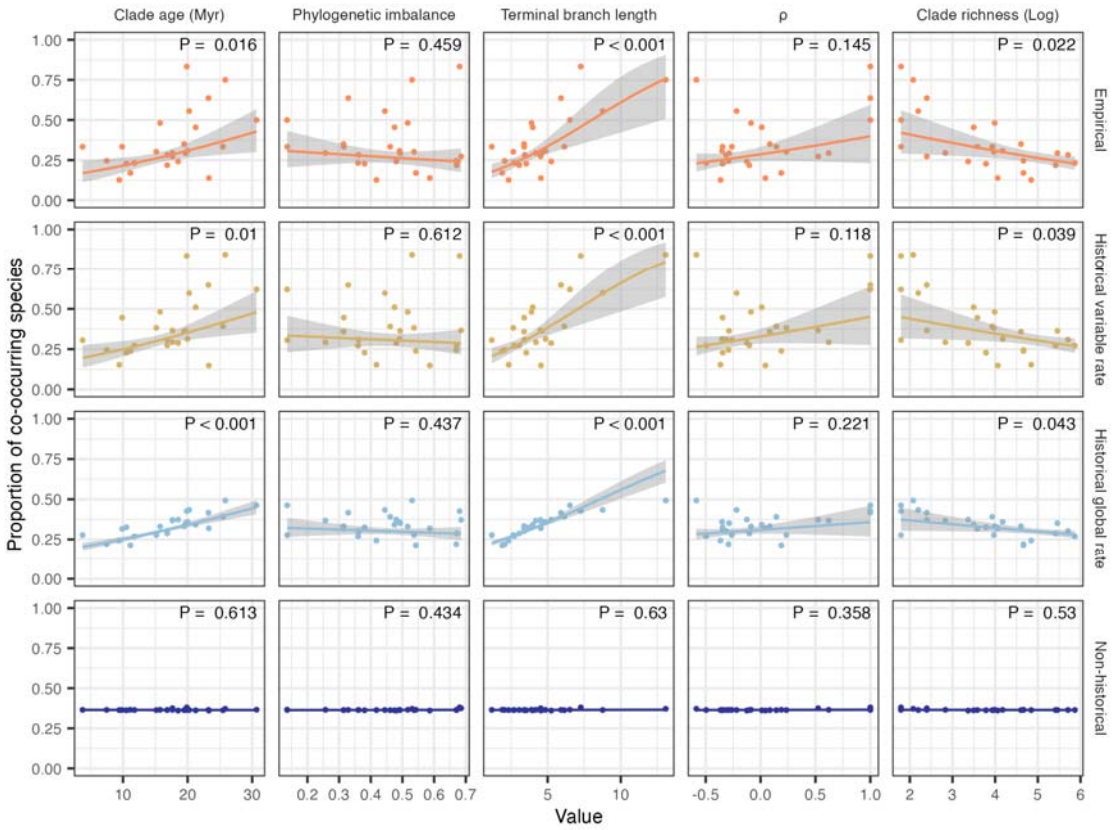


Figure 4 Relationship between phylogenetic metrics and assemblage species richness for real bird clades and model simulations. Colours and rows denote the empirical relationships for all clades (orange) and relationships simulated under the 'historical variable rate' (yellow), 'historical global rate' (blue) colonisation rate and 'non-historical' (dark blue) model. For the simulation scenarios the mean proportion of co-occurring species across 2500 replicated simulations are plotted. Columns represent the five phylogenetic metrics capturing different aspects of clade evolutionary history: The crown age of the clade, phylogenetic imbalance (Colless' index), mean branch length of the extant species, ρ a measure of temporal change in diversification rates through time where positive (negative) values indicate an increase (decrease), and log-transformed clade species richness. Shaded areas represent 95% confidence intervals of fitted generalised linear models with quasibinomial error distribution and logit link function weighted for clade richness.

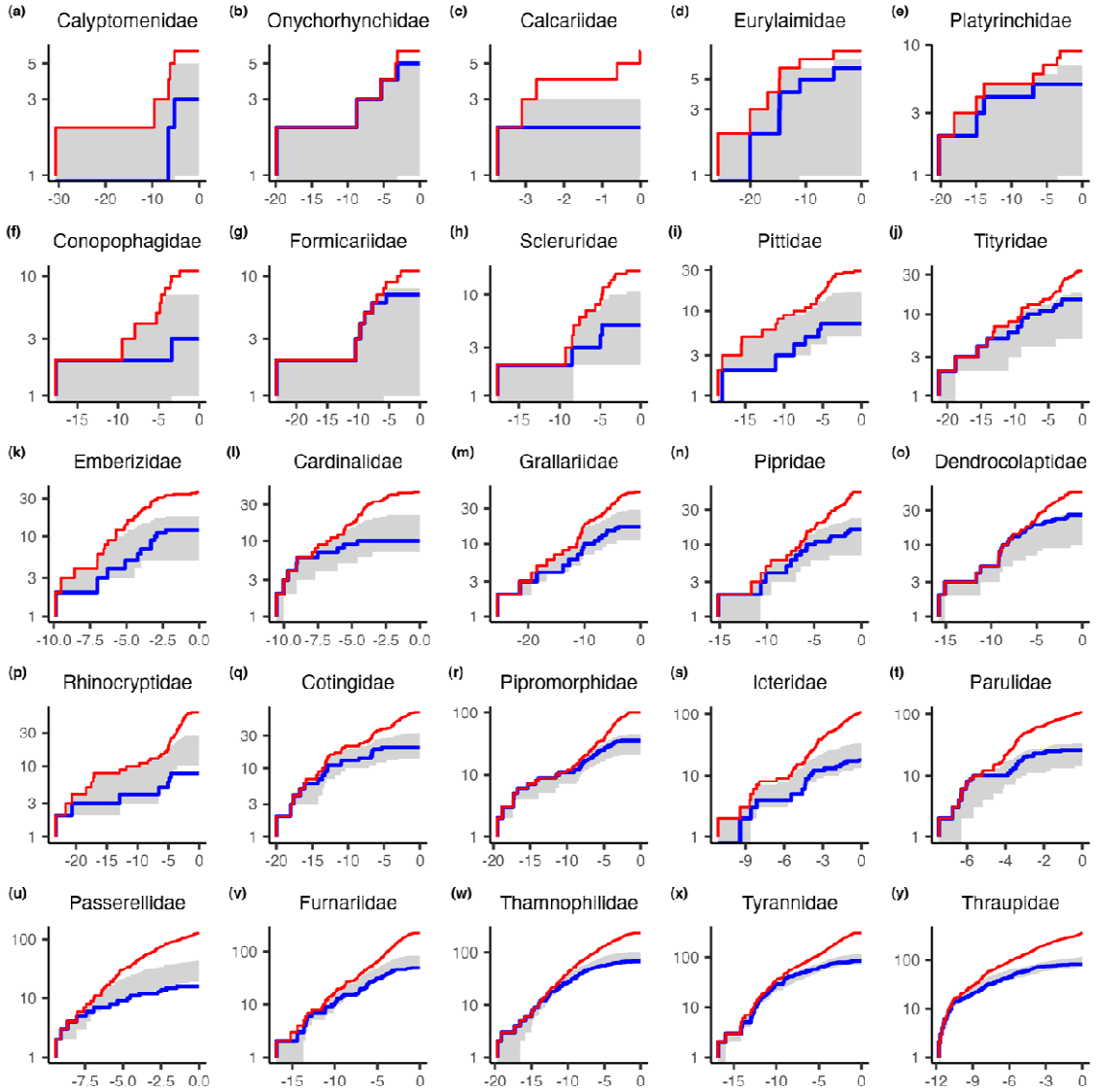


Figure 5 Clade and assemblage lineage-through-time (LTT) plots showing the accumulation of lineages through time within each family (red) and assemblage (blue, i.e., grid cell of maximum richness) compared to model predictions. Shaded areas correspond to the 95% confidence intervals of a rank envelope (Mrkvička *et al.* 2017) for the assemblage LTT expected under the 'historical global rate' model (n = 2500 replicate simulations). Families are ordered from smallest (Calyptomenidae, 6 species) to largest (Thraupidae, 354 species).