Colonisation lags predict sympatric diversity 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, Oxford, The United Kingdom

[2]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, The United Kingdom

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

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

Abstract

Ecological communities are assembled over time, suggesting a crucial role for evolutionary history in determining the diversity of species assemblages. Indeed, differences in the build-up of sympatric diversity across space, time and clades may reflect variation in the historical patterns of speciation and thus time available for colonisation. However, quantifying the role of evolutionary history is challenging and thus its potential effects have received far less attention than the effects of ecological limits to diversity. Here we use a dynamic model of assembly by allopatric speciation, colonisation, and local extinction to test how evolutionary history predicts variation in sympatric diversity across clades of passerine birds. We show that the clade-specific history of speciation strongly impacts the build-up of sympatry over time, and that phylogenetic tree shape statistics are key predictors of differences in sympatric diversity between clades. Within clades the proportion of species that are sympatric is highest in old, species-poor, and phylogenetically balanced clades. These historical factors all promote the average age of extant species, providing more time for species to expand their geographic distributions and colonise communities. We further show that phylogenetic patterns in community structure frequently interpreted as evidence that community assembly is limited by the saturation of ecological niche space can be explained simply by the lag-time to colonisation following speciation. Our results show that the macroevolutionary history of speciation and assembly time is essential in explaining variation in the structure and diversity of ecological communities.

Key words: Sympatry, coexistence, colonisation, ecological limits, niche filling, birds

Introduction

The investigation of patterns in species richness is a central focus of ecology. Much debate has centred around the extent to which these patterns are controlled by current climate or by history (Francis & Currie 2003; Jetz *et al.* 2004; Ricklefs & Latham 1993), where history can mean many things, encompassing the climatic and geological history of regions and the dynamics of evolutionary lineages (Cracraft 1994; Mayfield & Levine 2010; Patterson 1999; Ricklefs 1987; Webb 2000). A recurring theme has been the role of time in limiting the accumulation of species. In some cases, the recency of historical events makes the necessity of time clear. For example, species richness is still rebounding following the eruption of Krakatau in 1883 (Bush & Whittaker 1991) and the retreat of northern hemisphere ice sheets at the end of the Pleistocene (Blois *et al.* 2010). While variation in diversity among biomes and regions has been linked to differences in when these climatic zones or regions arose or were first colonised (Kozak & Wiens 2012; Miller *et al.* 2018; Stephens & Wiens 2003; Wiens 2011), such studies address the accumulation of richness within large geographic areas, rather than the richness of species co-occurring in assemblages.

Because most new species arise in geographic isolation (Coyne & Orr 2004; Phillimore *et al.* 2008), ultimately the accumulation of richness within assemblages depends on the expansion of species geographic ranges and attainment of secondary sympatry (Pigot *et al.* 2018; Price *et al.* 2014; Ricklefs 2004; Ricklefs & Schluter 1993; Tobias *et al.* 2020). Previous studies analysing the transition to sympatry among pairs of sister species have shown that although instances of sympatry among young rapidly diverging lineages exist, most sister species continue to have geographically non-overlapping distributions for millions of years after the initiation of speciation (Price 2010; Weir & Price 2011). By focussing on range overlap between pairs of closely related species, these studies have been able to identify both intrinsic and extrinsic barriers delaying the build-up of sympatry, including limited dispersal, reproductive interference, and competition for ecological resources (Pigot & Tobias 2015; Price 2010; Sexton *et al.* 2009; Weber & Strauss 2016). But this narrow focus has not allowed quantification of how the time available for colonisation potentially limits the richness patterns of entire species assemblages.

If colonisation is a highly stochastic and there is a delay until colonisation process this leads to a key prediction: that current variation in the maximum number of sympatric species across clades should depend on their macroevolutionary history of diversification, and specifically the timing of past speciation events (Mittelbach & Schemske 2015; Pigot & Etienne 2015).

Several studies have explored correlations between patterns of sympatric richness and the phylogenetic properties of clades, but these have often been with the aim of testing for signatures that species richness is regulated by an ecological carrying capacity (Kennedy *et al.* 2018; Larcombe *et al.* 2018). For example, phylogenetic branches representing co-occurring species of Himalayan songbirds appear to exhibit a slowdown in accumulation towards the present as may be expected if species richness is locally capped (Price *et al.* 2014). Weir (2006) found that maximum sympatric richness was higher among those Amazonian bird genera in which the rate of diversification inferred from the reconstructed phylogeny had declined over time (but see (Pie *et al.* 2023)), potentially consistent with the idea that local niche filling has inhibited recent speciation. Other studies have reported strong positive correlations between clade richness and the total number of species in sympatry (Kennedy *et al.* 2017). However, the absence of a process-based model for how sympatric diversity is assembled, has hindered further exploration and interpretation of these phylogenetic patterns. We hypothesise that all these patterns can be explained without invoking ecological limits if colonisation is a stochastic process.

Here we test the role of time in limiting species richness by applying a process-based model to explain differences in sympatric richness across families of oscine and suboscine passerine birds (n = 40 families). These groups are ideal for our analysis because, despite the high overall species diversity (n = 2081 species), comprehensive phylogenetic and geographic data is available for almost all species (BirdLife International and Handbook of the Birds of the World 2020). Furthermore, the taxonomic families comprising these radiations differ widely in both total richness and the maximum number of species found in sympatry. Our analysis consists of the following steps. First, we use geographic range maps to identify the assemblage (~100 km^2 grid cell) with the highest number of sympatric species across the distribution of each family. Second, using the identity of the sympatric species and the divergence times of species within each

family we use a maximum-likelihood approach to estimate the average time from speciation to the colonisation of the assemblage. Third, using simulations we assess the extent to which this stochastic colonisation model can explain observed differences in the maximum proportion of species occurring in sympatry across families and its correlation with key metrics describing the clade's evolutionary history: (*i*) clade richness, (*ii*) phylogenetic imbalance, (*iii*) clade age, (*iv*) the mean age of extant species and (*v*) the temporal constancy in branching times. For this latter property, we examined both how the level of sympatry varies across clades with different temporal patterns of diversification, and the dynamics of sympatric species accumulation within clades. Throughout, we compare explanatory power of the delayed colonisation model to that of a non-historical null model in which there is no time to colonisation following speciation and levels of sympatry depend only on the relative rate of local extinction. Together, our analyses therefore aim to address the role of evolutionary history and time for colonisation in explaining contemporary patterns of species richness.

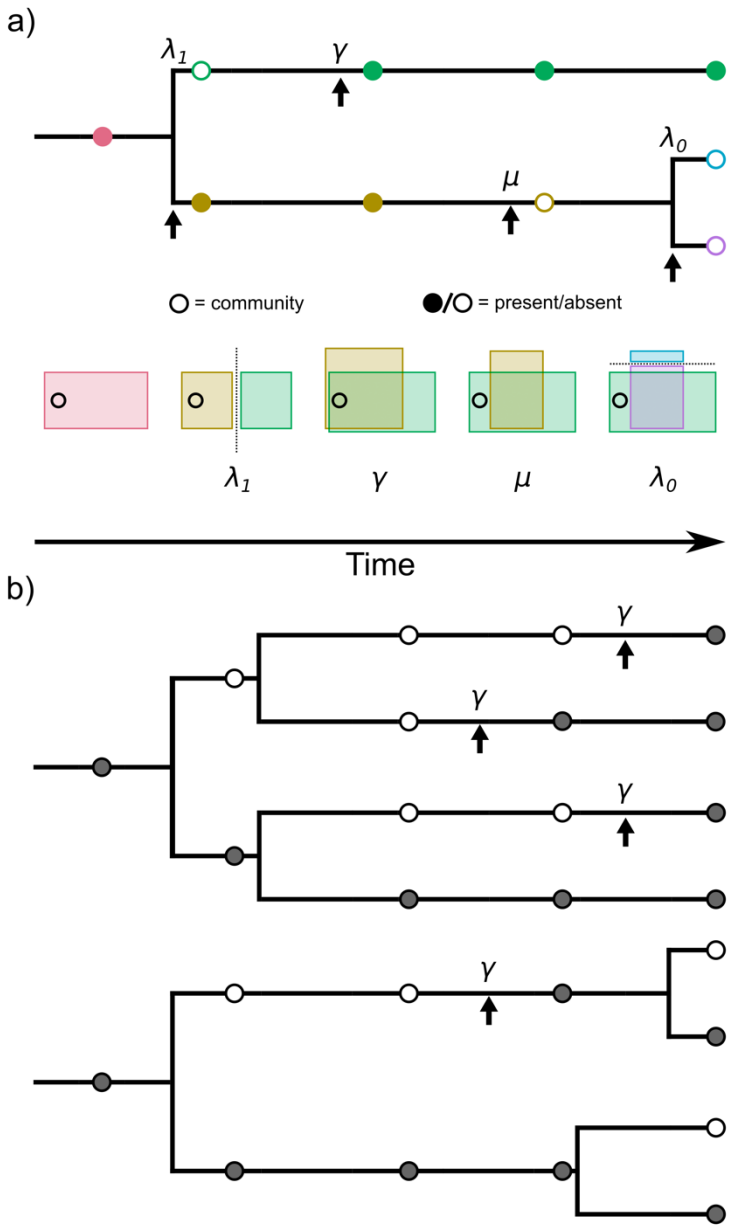

Figure 1 Visualisation of the phylogenetic assembly of communities. a) conceptualisation of the link between allopatric speciation, range expansion and sympatric diversity. Upward-pointing arrows show timing of events. When speciation occurs by the ancestral range (red rectangle) being divided into two (yellow and green), at maximum one species will be present (λ_1) in the assemblage (circle). If the ancestor is not present, neither of the descendants (blue and purple) will be present (λ_0). Lineages may become present by colonising the assemblage (γ) or may become absent by experiencing local extinction or extirpation (μ). As these dynamics play out over time, the colonisation and local extinction of species might be conceptualised as the expansion (green) and contraction of species' ranges (yellow) and the attainment or loss of sympatry between the members of a clade. b) two hypothetical phylogenies, each with four species, with the top phylogeny having older species on average and the bottom younger. Filled circles highlight the presence of a lineage in the assemblage, and arrows indicate colonisation events. For both

phylogenies the ancestor is present in the assemblage, but the dynamics of allopatric speciation and colonisation result in a higher degree of sympatric diversity in the tree with older species.

Materials and Methods

Phylogenetic and geographic range data

We compiled a dataset consisting of oscine and suboscine passerines predominantly endemic to the Americas. These clades were chosen based on their near complete phylogenies and highly detailed spatial data on the geographic distribution of species. The evolutionary relationships between species and divergence times were obtained from two recently published time-calibrated phylogenies containing >95% (oscines) and >98% (suboscines) of described extant species (Barker *et al.* 2015; Harvey *et al.* 2020). Our analysis focused on how the different evolutionary histories of clades impact sympatric diversity. Consequentially, the two main phylogenies are divided into family level clades that had at least five species (n = 25). The use of families provided enough clades to analyse variation in patterns of sympatric richness but also large enough clades to reliably calculate phylogenetic metrics.

We were interested in explaining variation in the maximum level of sympatry clades have attained. To do this we used expert delineated range maps to define assemblages of species occurring in a particular geographic area (BirdLife International and Handbook of the Birds of the World 2020). In contrast to local species inventories which are not sampled systematically, expert range maps provide a comprehensive characterisation of species richness patterns albeit at a limited spatial resolution. The phylogenies and range maps we used are based on different taxonomies and so we aligned these by using the taxonomy provided by the phylogenetic trees and merging or splitting species geographic ranges accordingly. Species that did not have geographical data associated with them (n = 5, equally distributed among clades of > 53 species each) were pruned from the dataset. Range maps were then extracted onto an equal area grid of 96 km^2 resolution, which is the grain size typically used in macroecological analyses to avoid false presences (Jetz & Fine 2012; Pigot *et al.* 2016; Rahbek 2005). For each family-level clade, we identified the grid cell with the highest species richness and recorded the identities of the species present in that cell. Note, that the grid cell with the highest richness could vary across clades. To ensure our results were not dependent on the

152 particular grain size used, especially in regions of steep environmental and species
153 turnover, we repeated our analyses using 24 km^2 grid cells.

154 *Estimating the dynamics of community assembly*

155 To determine if evolutionary history could explain differences in sympatric richness
156 across clades we applied and modified the Dynamic Assembly Model of Colonisation,
157 Local Extinction and Speciation, 'DAMOCLES' (Pigot & Etienne 2015). DAMOCLES was
158 originally developed as a null model to evaluate patterns of phylogenetic structure for a
159 single community. DAMOCLES models the assembly of a single community via
160 colonisation (γ) and local extinction (μ) along the branches of a phylogeny (Figure 1).
161 These rates are assumed to be constant over time and across lineages, with the
162 composition of the regional pool at time *t* provided by the lineages in the reconstructed
163 phylogeny that are extant at time *t*. Thus, the timing of speciation events through which
164 regional pools evolve are not modelled directly but are instead obtained from the
165 reconstructed phylogeny. We assumed that speciation happens in allopatry and that the
166 grid cell size used to define assemblages (i.e. 24 or 96 km^2) is too small for *in situ*
167 speciation to occur (Coyne & Price 2000; Kisel & Barraclough 2010). This is a reasonable
168 assumption for birds, for which more than 99% of speciation events are estimated to have
169 involved an allopatric phase and where cladogenetic speciation in continuous areas
170 smaller than Madagascar (1600 x 350 km) are extremely rare or debatable (Coyne & Price
171 2000; Phillimore *et al.* 2008). The key implication of allopatric speciation is that for a
172 lineage present in the assemblage, this results in one daughter lineage being present and
173 the other being absent. For a lineage absent from the assemblage, both daughter lineages
174 would also be absent. Thus, speciation does not directly result in an increase in
175 assemblage richness, as this relies on the colonisation of species (Figure 1b).

176 A likelihood approach to estimate γ and μ based on the presence-absence of species
177 in the assemblage and phylogenetic branching times was previously developed by Pigot
178 & Etienne (2015). Here we estimated γ and μ using the presence-absence of species in the
179 grid cell of maximum richness for each clade. Selecting communities from the tail-end of
180 the richness distribution might introduce error into the estimation of these parameters
181 (e.g. (Louca *et al.* 2022)). However, the communities with exceptional richness are of
182 inherent interest, but we also show how the model performs when rates are estimated for
183 an average community in the supplement. To test the effect of delayed colonisation in

184 explaining differences in maximum sympatric richness across clades we modified the
185 likelihood optimisation workflow, so that rather than optimising the likelihood for each
186 family-level phylogeny separately, we estimated a single global rate, varying starting
187 parameter values and optimisers, of γ and μ across all clades for which the cumulative
188 likelihood is maximised. By estimating a single global rate of γ and μ, differences in the
189 expected number of sympatric species across clades arise entirely from differences in
190 their phylogenetic histories and thus time for colonisation, rather than differences in rates
191 of γ and μ. We do however realise that variation in rates among families is present, and
192 explore this by individually estimating the colonisation and extinction for each family, but
193 focus on the global rate models.

194 Using this analytical framework we examined three different assembly scenarios.
195 First, we fitted two different scenarios of delayed colonisation. One where (*i*) both γ and
196 μ were estimated (colonisation and local extinction), and one (*ii*) in which only γ was
197 estimated and μ was set to zero (delayed colonisation only). Second, we fitted a model
198 where γ was fixed at a very high rate (γ = 1000) and μ was estimated. A value γ = 1000
199 equates to an average waiting time from speciation to colonisation of 1000 years.
200 Although arbitrary, this is sufficiently rapid compared the average branch lengths in the
201 phylogenetic trees (2.57 million years), that it captures a scenario in which the lag time to
202 colonisation is absent and which evolutionary history will have no effect on patterns of
203 sympatry. Instead, under this non-historical null model, sympatric richness is simply
204 determined by the ratio between γ and μ, with the latter rate estimated from the data.

205 *Simulating the build-up of sympatric diversity*

206 We compared the fit of each scenario using both AIC values and the ability to reproduce
207 the observed patterns of sympatric richness across clades. Specifically, for each scenario
208 we simulated assembly forward in time along the branches of each individual phylogeny
209 according to the estimated values of γ and μ. Simulations were implemented via a Gillespie
210 algorithm (Gillespie 1977), and modelled from the root of the tree at $t = 0$ until the present.
211 Under this algorithm the transition between state 0 (locally absent) and state 1 (locally
212 present) was modelled via the occurrence of colonisation ($0 \rightarrow 1$) and local extinction (1
213 $\rightarrow 0$) events. The root state of 0 and 1 was randomly assigned with equal probability, and
214 the waiting time till the next event (δ) was drawn from an exponential distribution with a
215 mean equal to the sum of all per-lineage rates. At time $t + \delta$, either a colonisation or local

extinction event occurred to a randomly drawn lineage with a probability equal to the sum of the respective per-lineage colonisation and local extinction rates. This simulation framework was already incorporated in DAMOCLES, but we rewrote the functions in *Rcpp* (Eddelbuettel & Balamuta 2018). The simulation was repeated a thousand times per clade from which we calculated the 95% confidence interval (95% CI) in the expected number of sympatric species.

Predictors of sympatric diversity and phylogenetic structure

For each family we compared the maximum proportion of species occurring in sympatry to that expected under our delayed colonisation models and non-historical null model. We used proportional richness because absolute assemblage richness is constrained by the number of species in the clade and thus largely reflects differences in clade richness (Figure 2a, Figure S1). To understand those aspects of phylogenetic history shaping current patterns of sympatry we quantified the following metrics for each family: (*i*) the crown age (Myr), (*ii*) Colless' index of imbalance standardised to the 'proportional to distinguishable arrangements' model in order to be independent of tree size and richness, and where higher values indicate greater imbalance (Bortolussi *et al.* 2006; Mooers & Heard 1997), (*iii*) the relative rate of diversification (speciation – extinction) between the first and second half of the phylogeny according to the ρ metric (Etienne & Rosindell 2012; Pigot *et al.* 2010), and (*iv*) the mean terminal branch length (*mbl*) used as a measure of the average age of extant species. We tested the effect of each of these phylogenetic metrics in predicting differences in proportional sympatric richness across families using a generalised linear model with a binomial error distribution and a logistic link function (Douma & Weedon 2019). We did this using both the empirical data and the expected proportional richness (mean across simulation runs) under our delayed colonisation models to test whether the observed relationships can be explained by a lag time to colonisation. We only considered families with more than five species as calculating phylogenetic properties for smaller trees becomes less informative. For instance, ρ would only be able to obtain a limited set of values as diversification either happens in the first half (ρ = -1) or second half (ρ = 1). This is especially relevant for our final analysis on the phylogenetic structure of the assemblage if only one or two species were to be present in sympatry.

Our analysis of clade phylogenetic structure tests how macroevolutionary history influences current sympatric richness. We also tested whether delayed colonisation influences the appearance of how sympatric diversity has accumulated over time based using the phylogenetic branching times among currently sympatric species. Lineage-through-time (LTT) plots indicate the accumulation of branches in a phylogenetic tree towards the present. Deviations in LTT shape from a constant rate null model of species diversification can provide evidence of non-random changes in diversification rate over time. For example, a non-random slowdown in diversification rate could indicate a process of adaptive radiation, where an initial burst of speciation is followed by a slowdown in speciation rate as ecological niches are filled (McPeek 2008). Similar logic has been applied to the LTT plot constructed from only those lineages that are currently sympatric in an assemblage (Price *et al.* 2014; Weir 2006). The appearance of a slowdown in the assemblage-LTT plot could indicate that rates of colonisation have slowed down over time, for example due to the local filling of niche space. However, our expectation was that such a pattern would also be expected if there is a lag time to colonisation because fewer closely related lineages (and thus branches) will occur in sympatry than expected by chance. To test this, for each clade we produced the lineage-through-time plot calculated at the level of family and at the level of the focal assemblage (i.e. grid cell of maximum sympatric richness) by pruning the family-level phylogeny to only those species that are present. We construct 95% confidence intervals via a rank envelope to deal with statistical non-independence (Mrkvička et al. 2017). Additionally, we calculated ρ for these assemblage-phylogenies, while repeating this for the empirical data, the non-historical model, and the two delayed colonisation scenarios. The aim of this analysis was to test whether phylogenetic properties of assemblages previously interpreted as evidence that ecological limits constrain sympatric diversity can also be explained by our model accounting for macroevolutionary history but in which ecological limits are absent.

Results

Model estimates and predictions

Our final dataset consisted of 7 oscine and 18 suboscine clades that ranged from 6 to 354 species in size. As expected, the maximum number of sympatric species inevitably strongly covaries with clade richness (Figure 2a, Figure S1). Accounting for clade richness and examining the proportion of species in sympatry reveals substantial variation, with

clades ranging from 15.5 to 100% of species in sympatry for respectively the Onychorhynchidae and the Rhinocryptidae.

Estimated rates of colonisation (γ = 0.215 Myr^{-1}) and local extinction (μ = 0.153 Myr^{-1}) across oscine and suboscine birds are similar to the mean rates of species diversification estimated by Jetz *et al.* (2012) across the global bird phylogeny (0.157 species Myr^{-1}). The estimated colonisation rate equates to a mean lag time from speciation to colonisation of 4.65 Myr. The estimated colonisation rate was lower when assuming no extinction (γ = 0.075 Myr^{-1}). The scenario where both γ and μ are estimated ($AIC_{\gamma\mu}$ = 2446.64) provides a substantially better fit (ΔAIC = -52.11) compared to the colonisation-only model (AIC_{γ} = 2498.75), and the non-historical model (AIC_{NH} = 2496.89) . When assemblages were constructed from 24 km^2 grid cells, very similar results were obtained (for the colonisation-extinction model: γ = 0.220 Myr^{-1} and μ = 0.124 Myr^{-1}, $AIC_{\gamma\mu}$ = 2285.07; for the colonisation-only model: γ = 0.060 Myr^{-1}, AIC_{γ} = 2323.93; and the non-historical model: AIC_{NH} = 2310.73).

The model adequacy analysis showed that the delayed colonisation model predicts the empirical proportions of sympatric species for the different clades well. For the colonisation-only scenario and the colonisation + local extinction scenario, the observed proportion of species in sympatry fell within the 95% CI of the expected proportions in 22 and 21 of the 25 clades respectively. Sympatric richness is underestimated for Onychorhynchidae and Dendrocolaptidae, and overestimated for Rhinocryptidae under the colonisation only model (Figure 2c). For the model with local extinction, sympatric richness is also overestimated for Passerellidae. In contrast, a non-historical null model explains levels of sympatry less well, with only 16 out of 25 clades having observed sympatry richness within the 95% CI of the null model. Similar results were obtained at the 24 km^2 grain size (Figure S2), when models were fitted individually to all clades (Figure S4), and when we looked at an average grid cell and not the cells with highest richness (Figure S7).

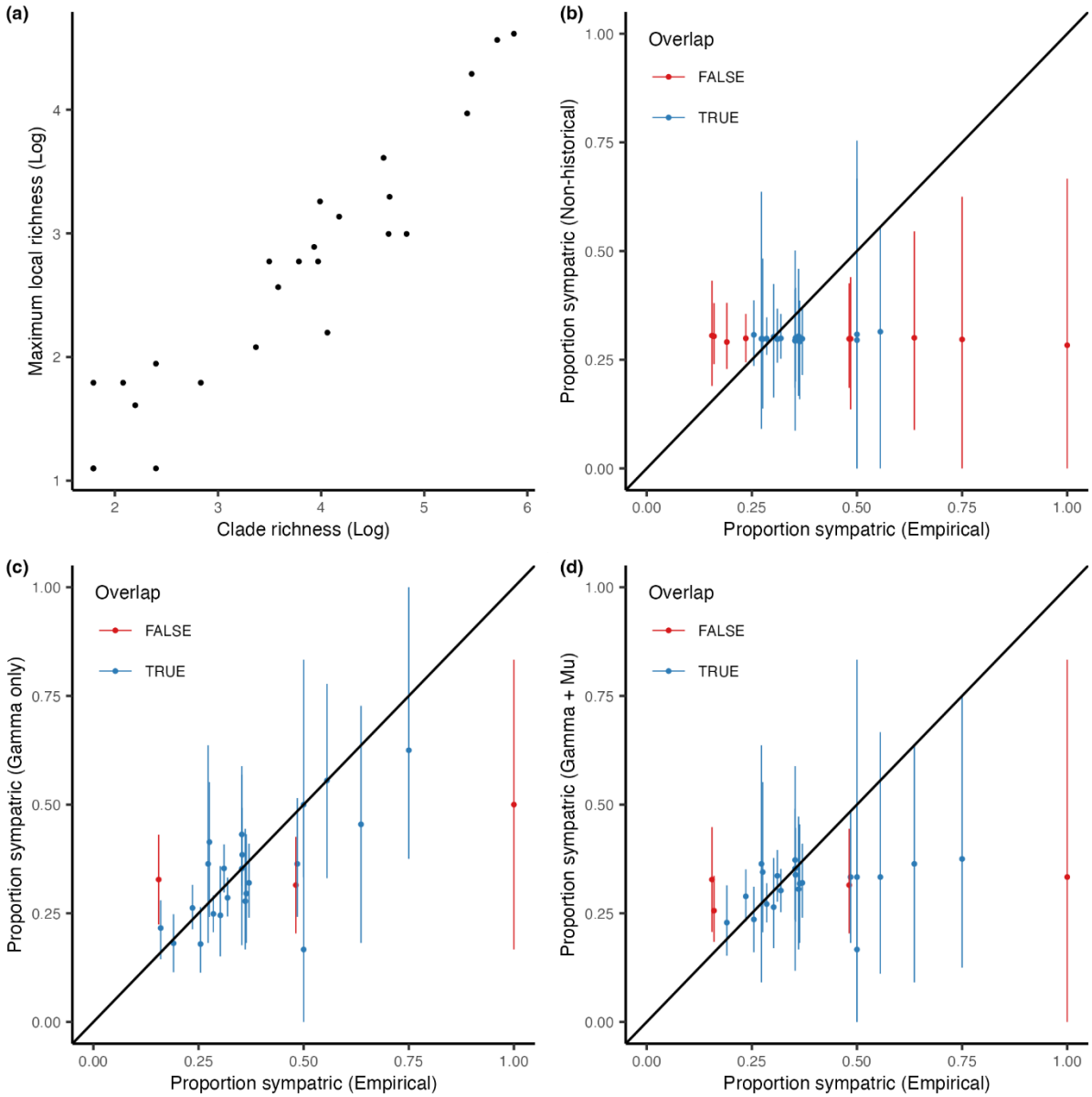

Figure 2 Observed and predicted proportional richness of passerine clades. (a) The empirical relationship between log-transformed sympatric and clade diversity. For respectively (b, c, and d) the predicted proportional sympatric diversity according to the non-historical null model, delayed colonisation, and delayed colonisation and local extinction is plotted against the empirical proportions. The line through the origin shows where the empirical equals the simulated proportion. Bars represent the 95% confidence intervals of the proportional richness of the assemblages recovered under 2500 simulations. Colours indicate if empirical proportions fall within (blue) or outside (red) of the CI's. For the non-historical model 16 out of 25 empirical proportion fall within the 95%CI, whereas 22 out of 25 fall within the CI's of the delayed colonisation model and 21 out of 25 for the delayed colonisation model with extinction.

The predictors of the sympatric diversity across clades

Across avian clades, the proportion of species that are sympatric increases with clade age and mean terminal branch length and declines with clade richness and phylogenetic tree

balance (Figure 3). There was no effect of family-level ρ on the proportion of species in sympatry. These patterns are highly consistent with those predicted by a delayed colonisation model. The only difference is that the delayed colonisation model does not predict any association between levels of sympatry and tree imbalance, while such a pattern is evident in the empirical data. The delayed colonisation + extinction model produces similar trends but fails to predict the decline in the proportion of species in sympatry with clade size, whereas the non-historical model shows no similar nor significant trend at all (not shown) due to little variation in sympatry (Figure 2b). We note that the model in which colonisation and local extinction are estimated separately for each family (Table S1) performs significantly better and predicts all relationships fatefully (Figure S6). The predictions of the models fitted to an average-diversity cell are comparable to the maximum grid cell, but the constrained empirical variation in sympatric diversity make interpretation limited (Figure S8).

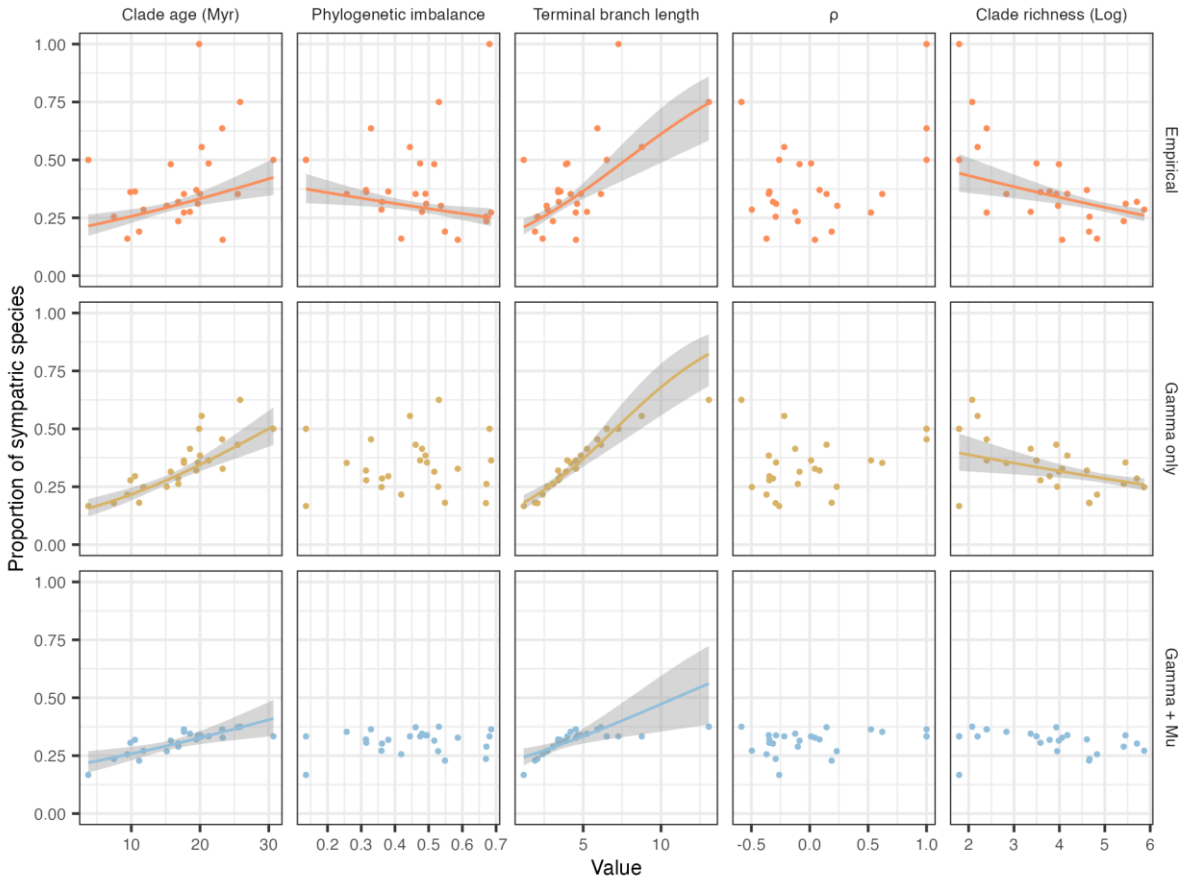

Figure 3 Relationship between evolutionary history and sympatric diversity. Colours and rows denote the empirical relationships for all clades (orange) and relationships simulated under the delayed colonisation scenario without (yellow) and with (blue) local extinction. For the simulation scenarios the mean

proportion across 2500 simulations are plotted. Columns represent the five metrics representing 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. Fitted trends are only shown when significant, and shaded areas represent 95% confidence intervals.

The relationship between clade and assemblage level phylogenetic metrics

We found that assemblage-level phylogenetic trees (i.e. pruned to only consist of those species occurring in the cell of maximum sympatry), show strong temporal slowdowns in branching rate over time (Figure 4, Figure S3). We found that such strong slowdowns are expected under a non-historical null model in 23/25 clades (except Emberizidae and Tyrannidae), suggesting that they can be explained by the pruning of the phylogenetic tree as this method removes more recent branching events (Figure S3). For the delayed colonisation model, slowdowns are weaker than expected for 4 families and stronger than expected for 1 (Calyptomenidae). For the delayed colonisation model with local extinction, the Rhinocryptidae respectively show weaker and the Calyptomenidae stronger slowdowns than expected, with the remaining families no different from model expectations. Thus, we find the slowdown in branching rate observed across assemblage trees is both consistent with that expected under non-historical assembly, or delayed colonisation with local extinction over time, but slowdowns appear too strong under a model with solely delayed colonisation.

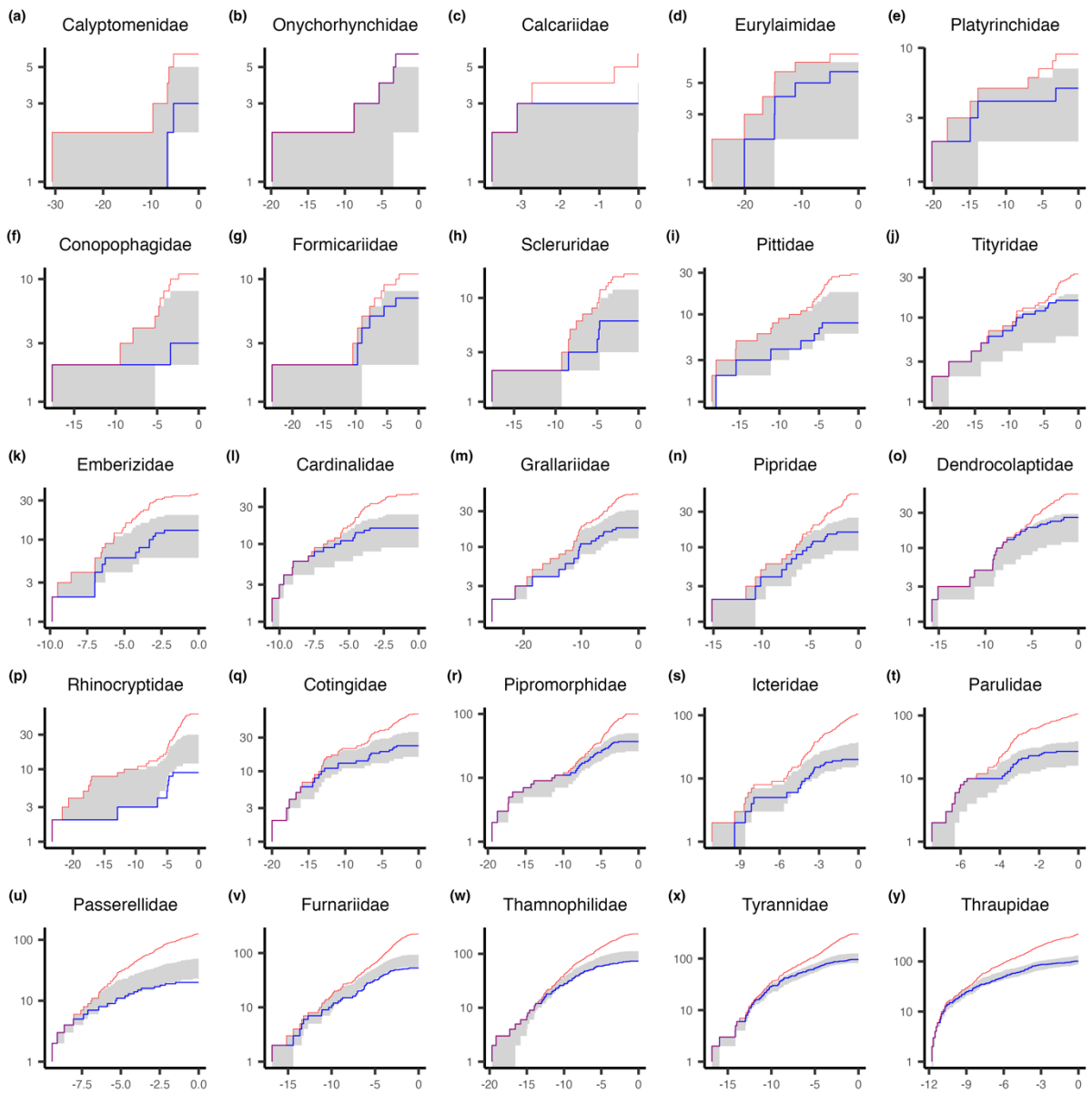

Figure 4 Clade and assemblage lineage-through-time (LTT) plots showing the accumulation of lineages through time within each family (red) and assemblage (blue, i.e. grid cell of maximum richness). Shaded areas correspond to the 95% confidence intervals of a rank envelope (Mrkvička *et al.* 2017) for the assemblage LTT under the delayed colonisation model with local extinction (n = 2500). Families are ordered from smallest (Calyptomenidae, 6 species) to largest (Thraupidae, 354 species).

Discussion

The evolutionary history of speciation, extinction and colonisation has been argued to be of fundamental importance to biological assemblages (Ricklefs 1987). However, a mechanistic perspective of how assemblages are taking shape via allopatric speciation and secondary sympatry has not been incorporated into the study of sympatric diversity patterns (Mittelbach & Schemske 2015). Here, by expanding upon a macroevolutionary

369 model of community assembly where the dynamics of diversification, colonisation, and
370 local extinction each play a role, we show that evolutionary history and the time available
371 for colonisation are of fundamental importance in explaining variation in levels of
372 maximum sympatry across clades. We show that a model incorporating only phylogenetic
373 history and assuming as little as a single rate of colonisation across a set of avian families
374 with radically different ecologies, can explain substantial variation in the maximum
375 proportion of species that are sympatric within any clade. This historical model
376 outperforms a null model in which evolutionary history is ignored. The proportion of
377 sympatric diversity and the phylogenetic properties that explain it are largely consistent
378 between simulations of our model and the empirical data. Both analyses show that the
379 predictors of proportional sympatric diversity are consistent with variables that correlate
380 with the time available for species to colonise the assemblage following speciation. Lastly,
381 we show that assemblage-phylogenetic patterns previously interpreted as evidence that
382 sympatric diversity is regulated by ecological limits (e.g. slowdowns in diversification rate
383 within an assemblage phylogeny), can equally be explained by a delay to colonise the
384 assemblage post-speciation, or even by the stochastic removal of species.

385 Earlier studies seeking to explain variation in regional diversity have focused on
386 the role of time for speciation, where it is argued that the earlier colonisation of a region
387 provides more time for speciation events to occur (Hutter *et al.* 2013; Stephens & Wiens
388 2003). However, these studies did not capture evolutionary history beyond the initial
389 colonisation of a region. This means that for a given clade age or time of initial
390 colonisation, differences in the historical dynamics of diversification and when speciation
391 events happen, does not have any effect on the local diversity of the assemblage. However,
392 as we show here, the timing of past speciation events is vital for determining sympatric
393 species diversity. The older the species are, the more time they have had to colonise the
394 assemblage after speciation in allopatry. Thus, by not accounting for the finer-grained
395 historical patterns of individual speciation events, the role of time in explaining variation
396 in sympatric diversity may have been underestimated.

397 Macroecological studies have variously used phylogenetic properties,
398 diversification rates, and diversity measures to infer ecological processes such as
399 competition from biogeographic patterns. For example, Weir (2006) showed that clades
400 of Amazonian birds showing stronger slowdowns in diversification rate over time

401 measured by the γ-statistic (Pybus & Harvey 2000) had higher absolute levels of
402 sympatric diversity, and interpreted this as evidence that niche space within assemblages
403 was saturated, inhibiting further range expansions and speciation (Weir 2006). We find
404 no relationship between slowdowns and proportional sympatric diversity in either the
405 empirical or the simulated data (Figure 3). We argue that this distinction is caused by the
406 inherent positive relationship between absolute and assemblage richness (Figure S1), and
407 the dependence of the γ-statistic on tree size (Pigot *et al.* 2010). Our results further
408 support this by finding negative relationships between the γ-statistic and assemblage
409 richness for empirical and simulated data without invoking ecological limits (Figure S9).
410 This does not rule out the possibility that ecological limits have caused a slowdown in
411 diversification rate (Rabosky & Hurlbert 2015), but shows that relationships between
412 phylogenetic metrics and sympatric diversity do not provide evidence of ecological limits
413 if not properly evaluated with null models and unbiased metrics.

414 Another pattern where our model results alter previous explanations regards
415 temporal slowdowns in the rate of branching among species present in an assemblage.
416 The appearance of such a slowdown is not simply a result of the well-known artefact of
417 incomplete sampling of a phylogeny (Cusimano & Renner 2010), and has thus been
418 interpreted as evidence that rates of colonisation slow down over time as niches are filled
419 (McPeek 2008; Price *et al.* 2014). However, our results show that the appearance of such
420 a slowdown arises even when rates of colonisation are constant over time. This is because
421 allopatric speciation results in an overdispersed phylogenetic structure of assemblages
422 (Pigot & Etienne 2015), with fewer closely related lineages co-occurring than expected
423 from random sampling. While it was recognised a decade ago that the appearance of
424 slowdowns in diversification rates could arise because speciation takes time (i.e.
425 protracted speciation (Etienne & Rosindell 2012)), here we show that the appearance of
426 slowdowns can arise because colonisation takes time. Conditions that would need to be
427 met for delayed colonisation to have no impact on assemblage phylogenetic metrics is that
428 speciation primarily happens *in situ* so that species do not have to colonise the area from
429 outside, or that *in situ* cladogenetic speciation is absent from the region against which the
430 assemblage is being compared. The process of delayed colonisation should therefore be
431 considered in many macroecological studies that focus on local assemblages that are small

but regional assemblages that are large compared to the spatial scale of cladogenetic speciation.

Our estimates of the time to sympatry for species in our clades is similar with earlier findings for sister pair analyses. For instance, Pigot & Tobias (2013) found that it takes on average 14.52 Myr for sister species of Furnariidae to reach secondary sympatry (Pigot & Tobias 2013). This is remarkably similar to the 13.36 (96 km^2) and 16.59 Myr (24 km^2) for species to attain sympatry estimated in our analysis (note that Furnariidae are included as one of the families in our analysis). This suggests not only a degree of conservatism across clades in the factors that limit range expansions and assembly, but also that the rate at which species transition to sympatry is consistent between sister species and more distantly related species. However, it is important to note that our model does not identify the factors that delay the attainment of sympatry and cause a lag between speciation and colonisation of the assemblage. Processes such as dispersal limitation, the build-up of reproductive isolation, and interspecific competition have all been shown to be of importance to secondary sympatry between sister species (Pigot *et al.* 2018; Weir & Price 2011), and may explain the lag-time to colonisation overall.

Previously, evidence that time matters for the build-up of diversity has been used as evidence arguing against ecological limits to diversity (Stephens & Wiens 2003; Wiens 2011). This evidence is controversial and has been scrutinised extensively, on the basis of both methodological and conceptual issues (Rabosky 2012; Rabosky & Benson 2021; Stadler *et al.* 2014). Our study supports a key role for time in the build-up of sympatric diversity and shows that this may provide an alternative, and potentially more parsimonious, explanation for assemblage level phylogenetic patterns previously interpreted as evidence of ecological limits. However, while we show patterns consistent with constant rates of colonisation and local extinction, a key next step will be to develop a diversity-dependent colonisation model to evaluate if the fit to the empirical data improves when assuming ecological limits. Furthermore, while our analysis focusses on explaining variation in the proportion of species that are sympatric, our model is silent to the causes of variation in clade level richness and the processes regulating diversification rates over time. In other words, variation in evolutionary history between clades that determine sympatric diversity, such as clade richness, the length of phylogenetic branches or the degree of phylogenetic imbalance, could be caused by differences in ecological

limits or the strength of species interactions (Reijenga *et al.* 2021). Therefore, our results do not rule out ecological limits regulate diversity at the level of clades and assemblages.

Our results show that a simple model based only on the macroevolutionary history of clades enables robust predictions about the variation in diversity and phylogenetic structure of assemblages. We focussed our analysis of predicting maximum sympatric diversity for each clade. However, different assemblage configurations can occur by chance, alternative stable states can be facilitated by which species arrives first (i.e. priority effects) (Fukami 2015; Reijenga *et al.* 2021) and levels of sympatry often vary substantially across different parts of a clades distributions. Understanding the causes of these differences in assemblage diversity within clades will be an important next step in determining how sympatric diversity is constrained by ecology and evolutionary history. Furthermore, our stochastic model lacks the ability to make precise predictions about assemblage composition and in particular which species are likely to be present. For example, while our model predicts that the probability of both sister lineages co-occurring locally should increase with the time since they diverged, we cannot predict which lineage will be present at any given time as both are governed by identical colonisation rates. An important extension of our model would therefore be to incorporate key environmental and ecological factors, such as differences in dispersal ability and niche divergence (Reijenga *et al.* 2023), that are expected to influence the likelihood of range expansions and which combinations of species are able to coexist. While the simple models we have employed here provide a powerful approach for isolating the effect of time for colonisation on sympatric diversity, a complete understanding of the role of evolutionary history in community assembly is likely to require more complex models that account for the interactions between species and the feedback between local diversity and the dynamics of species diversification.

Acknowledgements

We thank James Rosindell and Luis Valente for comments and discussion.

Funding

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

References

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.

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.

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

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. & 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.

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.

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.

Larcombe, M.J., Jordan, G.J., Bryant, D. & Higgins, S.I. (2018). The dimensionality of niche space allows bounded and unbounded processes to jointly influence diversification. *Nat. Commun.*, 9, 4258.

Louca, S., Henao-Diaz, L.F. & Pennell, M. (2022). The scaling of diversification rates with age is likely explained by sampling bias. *Evolution*, 76, 1625–1637.

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.

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. (2013). Species interactions constrain geographic range expansion over evolutionary time. *Ecol. Lett.*, 16, 330–338.

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.

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. & 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. & 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.

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. (2004). A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.*, 7, 1–15.

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.

Ricklefs, R.E. & Schluter, D. (Eds.). (1993). *Species Diversity in Ecological communities: historical and geographic perspectives*. *Species Divers. Ecol. Communities*. University of Chicago Press Chicago.

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.

Tobias, J.A., Ottenburghs, J. & Pigot, A.L. (2020). Avian Diversity: Speciation, Macroevolution, and Ecological Function. *Annu. Rev. Ecol. Evol. Syst.*, 51, 533–560.

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.