UC Santa Cruz
UC Santa Cruz Previously Published Works

Title

The Metapopulation Bridge to Macroevolutionary Speciation Rates: A Conceptual Framework and Empirical Test.

Permalink

Journal

Authors

Januario, Matheus

Pinsky, Malin

Rabosky, Daniel

Publication Date

DOI

PERSPECTIVE OPEN ACCESS

The Metapopulation Bridge to Macroevolutionary Speciation Rates: A Conceptual Framework and Empirical Test

Matheus Januario[1] 🔬 | Malin L. Pinsky[2,3] 🔬 | Daniel L. Rabosky[1]

[1]Museum of Zoology & Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA | [2]Department of Ecology, Evolution, and Natural Resources, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA | [3]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, California, USA

Correspondence: Matheus Januario (januarioml.eco@gmail.com)

ABSTRACT

Whether large-scale variation in lineage diversification rates can be predicted by species properties at the population level is a key unresolved question at the interface between micro- and macroevolution. All else being equal, species with biological attributes that confer metapopulation stability should persist more often at timescales relevant to speciation and so give rise to new (incipient) forms that share these biological traits. Here, we develop a framework for testing the relationship between metapopulation properties related to persistence and phylogenetic speciation rates. We illustrate this conceptual approach by applying it to a long-term dataset on demersal fish communities from the North American continental shelf region. We find that one index of metapopulation persistence has phylogenetic signal, suggesting that traits are connected with range-wide demographic patterns. However, there is no relationship between demographic properties and speciation rate. These findings suggest a decoupling between ecological dynamics at decadal timescales and million-year clade dynamics, raising questions about the extent to which population-level processes observable over ecological timescales can be extrapolated to infer biodiversity dynamics more generally.

1 | Introduction

At the largest scales of time and space, the most striking feature of biological diversity is arguably the extent to which it varies: species richness shows dramatic variability amongst clades, through geological time, and amongst geographic regions. It is increasingly recognised that the causes of these patterns are in part attributable to differences in evolutionary rates of speciation and extinction (Coyne and Orr 2004, Rolland et al. 2023, Ricklefs 2007). Consequently, there has

been widespread interest—amongst ecologists and evolutionary biologists alike—in understanding both the pattern of speciation rate variation and its biological causes. The availability of time-calibrated phylogenetic trees for extant clades, in conjunction with sophisticated comparative and computational tools for estimating evolutionary rates (Jetz et al. 2012, Rabosky et al. 2018, Maliet et al. 2019, Rabosky 2014, Title and Rabosky 2019, Vasconcelos, O'Meara, and Beaulieu 2022), has revealed extensive variability amongst clades in the rate of speciation (Cooney and Thomas 2021, Henao-Diaz and

Pennell 2023, Singhal et al. 2022, Rabosky 2016). This variability cannot readily be explained by any simple association with latitude, biogeographic region or climate (Jetz et al. 2012, Economo et al. 2018, Rabosky et al. 2018, Rabosky 2022, Schluter 2016, Schluter and Pennell 2017) suggesting that biological, clade-specific traits influence the rates at which new species form.

Despite widespread success in documenting the pattern of speciation rate variation amongst extant lineages, it has proven much more difficult to determine which population-level processes—if any—can explain why some lineages are more likely to speciate than others (Henao-Diaz and Pennell 2023; Rabosky 2016). Many traits that should influence speciation or diversification appear to have ambiguous, inconsistent or scale-dependent effects on macroevolutionary dynamics (Anderson and Weir 2022, Anderson et al. 2023, Foisy et al. 2019, Helmstetter et al. 2023, Kraaijeveld, Kraaijeveld-Smit, and Maan 2011). For example, the evolution of reproductive isolation between diverging populations is considered a central component of the speciation process (Mayr 1963; Stankowski and Ravinet 2021). Consequently, species that have traits that promote faster evolution of reproductive isolation should—all else being equal—split into new species at faster rates than species lacking in such traits (Coyne and Orr 2004; Panhuis et al. 2001). This prediction underlies a large comparative literature on trait-dependent diversification, which has focused heavily on traits associated with sexual selection and other processes thought to influence the rate at which reproductive isolation accumulates (Anderson and Weir 2022; Cally et al. 2021; Helmstetter et al. 2023; Kraaijeveld, Kraaijeveld-Smit, and Maan 2011). However, phylogenetic speciation rates appear to be largely uncorrelated with direct (Rabosky and Matute 2013; Matute and Cooper 2021) and indirect (Freeman et al. 2022) measurements of the rate at which reproductive isolation evolves, challenging the hypothesis that the evolution of reproductive isolation itself is a central 'rate limiting control' on the speciation process (Harvey, Singhal, and Rabosky 2019; Rabosky 2016). Furthermore, the rate at which geographically isolated populations diverge genetically is largely uncorrelated with speciation rates (Burbrink et al. 2023; Nitschke et al. 2018; Singhal et al. 2018, 2022; Wacker and Winger 2024), even though such divergence should be correlated with the rate of incipient species formation (Harvey et al. 2017).

In spite of these generally mixed/ambiguous results with respect to traits and speciation (Anderson and Weir 2022; Helmstetter et al. 2023; Kraaijeveld, Kraaijeveld-Smit, and Maan 2011), one of the most important potential controls remains largely unexplored. As pointed out by Mayr (1963, 554) and others (Allmon 1992; Allmon and Sampson 2016; Dynesius and Jansson 2014; Glazier 1987; Harvey, Singhal, and Rabosky 2019; Levin 2000; Schluter 2016), successful speciation does not occur in a demographic vacuum: incipient species or divergent lineages must persist throughout the speciation process (Germain et al. 2021). They must persist for long enough to evolve reproductive isolation itself, and they must persist long enough to evolve sufficiently distinctive traits that we recognise them as distinct species. If reproductive isolation or genetic structure evolves quickly, then these properties may not be the primary

causes of variation in phylogenetic speciation rates. Put another way, it may be easy for incipient species to form through geographic splitting or the development of reproductive isolation, but if these lineages have low probability of long-term persistence, then they will have little impact on realised speciation rates and on biological diversity more generally (Levin 2000; Stanley 1979, 2008). Admittedly, the interaction between speciation and 'extinction of incipient species' raises many difficult conceptual problems and blurs the line between speciation and extinction (Allmon 1992; Allmon and Sampson 2016; Dynesius and Jansson 2014; Glazier 1987; Harvey, Singhal, and Rabosky 2019; Levin 2000; Mayr 1963; Schluter 2016; Stanley 1979, 2008). Regardless, there are clear theoretical reasons to hypothesize that (1) species should vary in properties relating to demographic persistence during the speciation process; (2) that this variation should have a biological basis (e.g., be caused by underlying traits), at least in part and (3) that lineage-specific variation in persistence should be correlated with macroevolutionary speciation rates. Thus, if we want to understand how and why speciation rates vary across the tree of life, we must explore new sources of data into how and why populations (and species) persist differentially across a range of timescales and environmental contexts.

Here, we illustrate a framework for comparing metapopulation-based estimates of persistence in a phylogenetic framework to answer whether within species persistence correlates with speciation rate. Because species traits must be 'heritable' (Lynch 1991) to have consistent (and detectable) effects on phylogenetic diversification patterns, we also test whether metapopulation properties are phylogenetically autocorrelated, such that closely related species are more likely to show similar metapopulation-level persistence. In our empirical analysis, we hypothesize that the dynamical properties of spatial occupancy represent metapopulation-level properties that are important to the persistence of incipient species. In general, the persistence of metapopulations has not been characterised and compared across a large set of vertebrate species, despite the potential importance of such persistence for the success or failure of incipient species (Harvey, Singhal, and Rabosky 2019; Levin 2000; Schluter 2016). We conceptualise persistence as the combination of two qualitatively different components (i.e., two 'persistence axes'): the expected occupancy at equilibrium (which approaches commonness in our spatial context, see below) and the ability to recover a certain occupancy after a disturbance (hereafter, 'resilience'). Disturbances in turn are exogenous discrete events that can push a population closer to extinction and that can potentially happen at many timescales (Capdevila et al. 2020, Harrison 1979, but see Strier 2021). Within this context, the ability to maintain large occupancy (i.e., be common) is an important component of metapopulation-level persistence because geographic ranges frequently are amongst the best factors predicting extinction probability (Harnik, Simpson, and Payne 2012, Harvey, Singhal, and Rabosky 2019, Jablonski 2008, Levin 2000, Smits 2015—but see Rabinowitz 1981). However, the effects of range on persistence at ecological timescales may be best studied from the perspective of dynamical systems properties, as ranges have been observed to change at decadal timescales (e.g., marine ranges in Batt et al. 2017 and Pinsky, Selden, and Kitchel 2020—for general

examples, see Beissinger and Riddell 2021 and MacLean and Beissinger 2017). Resilience captures a complementary feature associated with the temporal dynamism of ranges, as frequently disturbed environments may select for species that are able to quickly recover spatial structure (e.g., a proportion of occupied habitat, or the occupation of specific locations important to the species). Our approach encapsulates important properties related to persistence not only under classic views from metapopulation theory on what persistence is (Hanski 1999; Harrison 1979; Levins 1969) but also properties relevant to extinction at deeper timescales (i.e., occupancy at large spatial scales). Importantly for comparative analyses, our framework is flexible enough to incorporate the many different sources of noise and biases that real datasets face.

2 | Framework Overview

Our general approach involves testing the relationship between speciation rates and metapopulation properties estimated from time-series data on contemporary populations. These estimates are then compared to phylogenetic speciation rates as inferred from molecular phylogenies. We used a simple Markovian model of metapopulation dynamics to estimate rates of colonisation and extirpation (i.e., local, but not global, extinction), and from these rates, we calculated our demographic properties of interest (commonness and resilience).

As an illustrative example, we applied our Markov model to a space–time grid of occupancy probabilities for 189 species of marine demersal (i.e., bottom-associated) fishes from the North American continental shelf, where occupancy probabilities were estimated from long-term survey data in a Multi-State Occupancy Model framework (Batt et al. 2017). We used phylogenetic generalised least squares (PGLS–Felsenstein 1985, Symonds and Blomberg 2014) to test whether these species-level metapopulation traits (commonness, resilience) are correlated with rates of speciation as inferred from a time-calibrated phylogenetic tree of extant fishes.

3 | Metacommunity Context

In our analytical framework, we assume populations occur in spatial cells that are potentially connected to others through colonisation (i.e., all cells and species together represent a metacommunity–Leibold and Chase 2018, but see Grimm, Reise, and Strasser 2003). Thus, the relevant metapopulation quantities are the extirpation (local extinction) rate e and the colonisation rate c. Here the units of these rates are events per cell per year. The model receives as input an empirical time series of probabilities of presence (hereafter, ψ_t) for a given species, in a certain cell, at a certain year t (cell and species notation is omitted for simplicity). In practice, these ψ_t values might be obtained from other types of modelling that account for incomplete detection or other factors, and some of these approaches may allow the estimation of uncertainty in occupancy state simultaneously to the colonisation and extinction rates (for other types of occupancy models, see Bailey et al. 2014, Dorazio and Royle 2005, Kéry 2010, MacKenzie et al. 2003). Occupancy fluctuations caused by individual

movement are assumed to be either irrelevant or averaged out in our estimates. We also implicitly assumed that all patches (i.e., habitat subdivision) within a cell are equally reachable in terms of individual movement.

4 | Metapopulation Model

We used a two-state Markov model (derived in Section S1) to estimate rates of colonisation and extirpation (i.e., local, but not global, extinction) from a space–time grid of occupancy probabilities. This general approach was initially used to estimate faunal turnover in islands and was subsequently adopted more widely (e.g., Alonso et al. 2015, Gotelli 1991, Hanski 1999, Levins 1969, Pagel 1994, Simberloff 1969). However, the mathematical formulation of our model differs slightly from the standard two-state Markov process because we account for uncertainty in the presence or absence of species within a given cell by using probabilities of occurrence as data (i.e., $\psi_t \in [0,1]$). Thus, our framework integrates over (estimated) uncertainty in species-cell-specific presence. Let 00 and 01 represent the respective events that an initially unoccupied cell (state = 0), after some time interval $\Delta t = t_n - t_{n-1}$, is then found to be unoccupied again or that it changed its state to occupied (state = 1). Conversely, let 10 and 11 denote the corresponding events whereby an initially occupied cell is either unoccupied or occupied during the same time interval. By the law of total probability, we have the relationship $P(00) + P(01) + P(10) + P(11) = 1$ for a given interval Δt. Strictly speaking, each of these four events—and their probabilistic descriptions below—thus encapsulate all transition histories consistent with the observations, while simultaneously allowing for multiple transitions (including reversions) through Δt. We can thus compute the probabilities of each event as (see Section S1.6 for details):

$$
\begin{aligned}
P(01 \mid c, e, \Delta t) &= \frac{c}{c+e} \times (1 - \exp(-(e+c)\Delta t)); \\
P(00 \mid c, e, \Delta t) &= 1 - P(01 \mid c, e, \Delta t); \\
P(10 \mid c, e, \Delta t) &= \frac{e}{c+e} \times (1 - \exp(-(e+c)\Delta t)); \\
P(11 \mid c, e, \Delta t) &= 1 - P(10 \mid c, e, \Delta t)
\end{aligned}
\tag{1}
$$

The likelihood of metapopulational rates generating a given observed sequence X of occupancy probabilities ψ_t at a particular cell is then (Section S1.11):

$$
L(c, e \mid X) = \prod_{n=2}^{N}
\begin{bmatrix}
(1 - \psi_{n-1}) \times (1 - \psi_n) \times P(00 \mid c, e, (t_n - t_{n-1})) + \\
(1 - \psi_{n-1}) \times \psi_n \times P(01 \mid c, e, (t_n - t_{n-1})) + \\
\psi_{n-1} \times (1 - \psi_n) \times P(10 \mid c, e, (t_n - t_{n-1})) + \\
\psi_{n-1} \times \psi_n \times P(11 \mid c, e, (t_n - t_{n-1}))
\end{bmatrix}
\tag{2}
$$

where c and e are the colonisation and extirpation rates as defined earlier, and we omit subscripts associated with cell identity for clarity. The likelihood of the full dataset, for a given species, is the product of the above expression over all cells in the dataset. Note that each term in the above equation is the probability of a given event given the Markov parameters

(e.g., $P(00 | c, e, \Delta t)$) weighted by the probability that the focal event happened, given the data (e.g., $(1 - \psi_{n-1}) \times (1 - \psi_n)$ for the 00 event), and thus accounts for uncertainty in cell state. Parameter estimation was made by optimising the likelihood function (equation 2), jointly considering all same-region cells in which a given species appears (i.e., regardless of latitude, longitude or depth).

5 | Measuring Persistence

After c and e are obtained for each species–region combination, we calculated two metrics that together capture how metapopulation turnover relates to persistence. Our commonness metric $\varepsilon_{raw} = c/(c + e)$ equals the expected occupancy at equilibrium and so describes how common a species is (in a way that is mathematically independent of range size), because it approximates the probability of observing the species occupying a particular patch at any given time (though meta-community sampling can capture just a fraction of range-wide dynamics—see Section S1.8). Due to ε_{raw} values lying in the interval between zero and one, we logit-transformed this variable (and interesting alternatives could involve beta-family phylogenetic methods). We call this transformed metric $\varepsilon = \ln(\varepsilon_{raw}/(1 - \varepsilon_{raw}))$, and it can be used directly in the PGLS analyses employed here. However, to plot and interpret results we re-projected estimates in the original ε_{raw} space, which is more straightforward to interpret.

The complementary persistence metric τ measures 'resilience' and equates the log-transformed turnover rate ($\tau_{raw} = c + e$; so $\tau = \log_2(\tau_{raw})$). τ is proportional to how fast a metapopulation recovers equilibrium occupancy after a perturbation (see Section S1.8), and its units are (logged) events per cell per year. These transformations might make our final persistence components (i.e., τ and ε) nonlinear and so affect coefficient interpretation (Houle et al. 2011, Spake et al. 2023), but their scale reflects proportionality—a desired feature in comparative analyses (Gingerich 2019). The framework employed here then takes advantage of the mathematical independence between the persistence components proposed and aligns τ and ε to construct an interpretable persistence space (Figure 1). The final steps are the analysis of speciation rates with respect to species positions in this two-dimensional space using PGLS and posterior interpretation of its biological relevance in light of the PGLS parameter values and statistical significance, as well as the species patterns in the persistence space (e.g., hypothesized relationship in Figure 1C).

Different combinations of demographic properties (i.e., regions within the persistence space, hereafter also called 'persistence syndromes') serve as proxies for fundamentally different ways in which species-specific metapopulations can behave in the aftermath of environmental disturbance: populations can recover high occupancies quickly (Figure 1A.II), slowly (Figure 1A.I), recover their equilibrium occupancy fast but never reach high overall occupancies (Figure 1A.IV), or take a long time to recover their small average occupancy (Figure 1A.III). Moreover, if these syndromes have phylogenetic signal, all else being equal, clades that tend to recover their occupancy fast (i.e., higher τ) may be more able to persist in environments with frequent perturbations, while clades with overall higher occupancy will tend to persist more in environments whose disturbances tend to affect large areas.

6 | Handling Possible Sources of Noise and Bias

Sampling varies amongst regions in many demographic datasets, including the one we analyse here (Table S1). Because some clades are restricted or preferentially occupy some of these regions, region-specific sampling can bias the residuals associated with shared ancestry amongst species and thus create the illusion of phylogenetic signal in our estimates (see below). To remove variance in species-specific values of ε and τ that may be artificially correlated with region-specific sampling, we used linear mixed models that treat the region as a random intercept and species as a fixed effect. Mixed models are commonly used to estimate parameters in designs with known clustered or nested observations (Harrison et al. 2018). Such models can be robust to violations of their distributional assumptions and can provide unbiased estimates when model parameters are influenced by missing/unknown random effect components (Schielzeth et al. 2020). One can fit independent mixed models for ε and τ and use their fixed effects as final species values that can be used in all downstream comparative analyses. Importantly for the framework, none of the previously-mentioned transformations and parameter estimations is informed by phylogenetic relationships. In the Discussion section, we address the biological relevance of biogeographic regions to persistence and the caveats associated with removing residuals related to regions from species-level persistence estimates.

The uncertainty in occupancy estimates means that different datasets may store different levels of information due to the uncertainty in their values of ψ_t. Inspired by information theory (Shannon 1948), we addressed this possibility in Section S2 through the direct measurement of dataset information, together with simulations of its impact on estimates. We did not fit our Markovian model to species–region datasets that did not have enough information to estimate parameters (according to our information content metric and our simulations). Erroneous estimates can also result from a lack of change in occupancy states (i.e., there is no temporal variation of metapopulation occupancy within the dataset), due to change that is faster than sampling (i.e., analogously to a 'saturated' molecular dataset), or due to unreasonably high or low values of ε_{raw}. We addressed these concerns by filtering out species–region pairs whose τ value is out of the temporal scale of the sampling, or by truncating ε_{raw} in meaningful values (Details in Section S3).

Finally, there may be spatial heterogeneity in metapopulation dynamics across cells within the same biogeographic region and for the same species (e.g., due to source-sink dynamics between different locations or depths), which also can potentially lead to erroneous rate estimates. We addressed this possibility with a mixture model described and discussed in Section S4. If statistical support for the mixture model was found within the species–region dataset, we selected the mixture component with the highest ε to represent the metapopulation parameters for the species, reasoning that population
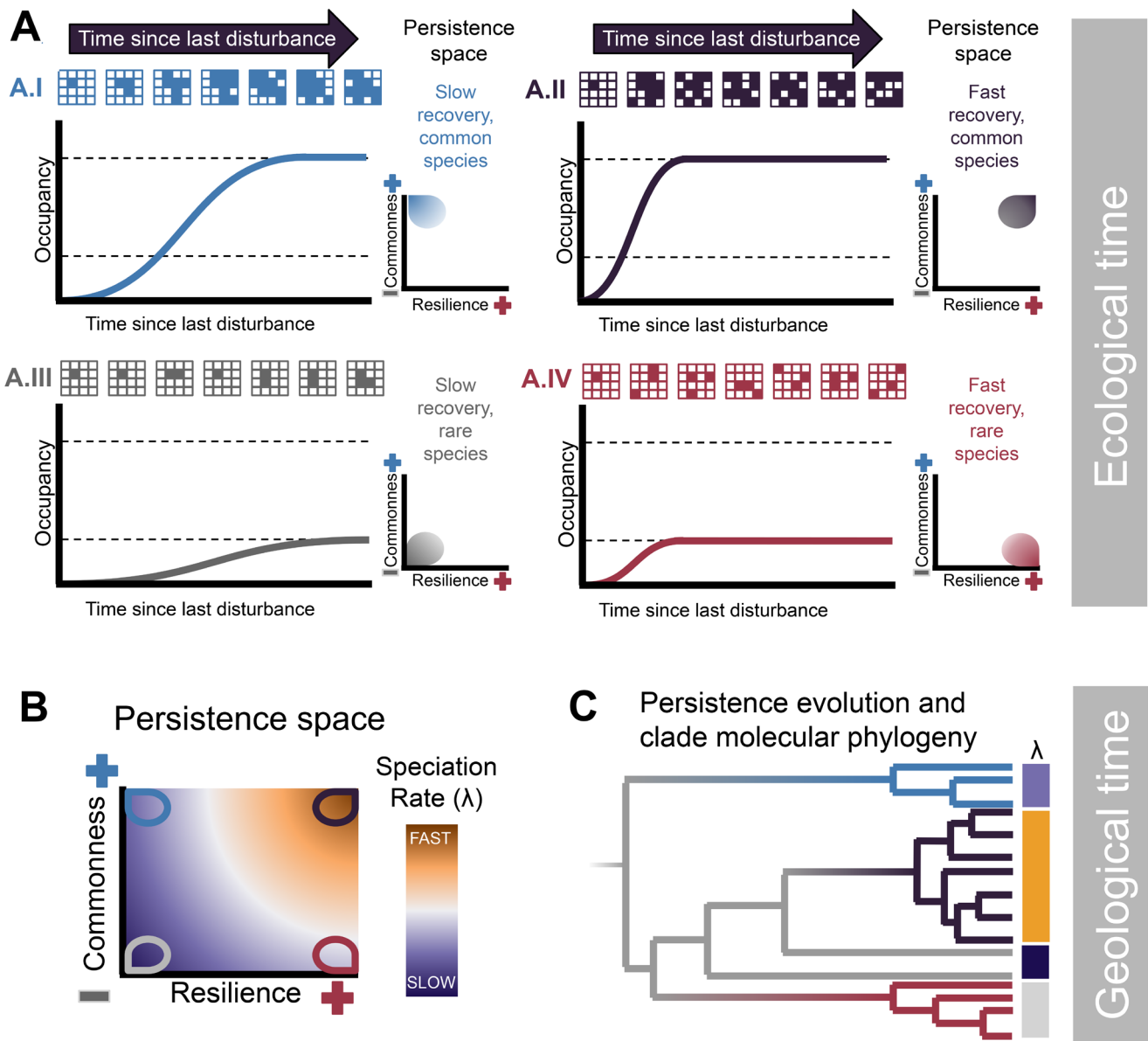
FIGURE 1 | Persistence syndromes and their possible relationship to macroevolutionary speciation rate. (A) Four general types of persistence syndromes (i.e., metapopulation dynamic modes, represented by colours grey, black, blue and red). The persistence space maps those syndromes as a combination of two axes: Resilience (x-axis) and commonness (y-axis). All else being equal, each persistence syndrome would have its characteristic stable structure (e.g., occupancy is illustrated, but abundance could be an alternative), and its average time to recover the stable structure (given by the turnover rate) after a disturbance that decreases the whole metapopulation to a few cells. This is shown both in the temporal series of lattices and in the plots of occupancy through ecological time. Species located in the commonness-prone region of the persistence space (e.g., species A.I, blue) persist mainly by being widespread in the geographic space, as their recovery after perturbations is relatively slow when compared to other species. Species located in the resilience-prone region of the persistence space (e.g., species A.IV, red) would rapidly recover their at-equilibrium spatial occupancy, although would not necessarily be common. Species may be able to maximise the values of both axes (e.g., species A.II, black), meaning such species would be both geographically widespread and rapid in recovering their equilibrium states. Finally, it is possible that some species could have relatively low values of both persistence axes (e.g., species A.III, grey)—and so are, at least theoretically, less prone to persist in the long term, at least when compared to the other species in the space. If a specific persistence syndrome is what confers a higher speciation rate, as shown in (B), then the phylogenetic patterns of persistence at the metapopulation level track the diversification of the clade, as shown in the hypothetical phylogeny of this group ((C), speciation rates shown at tips). One can visualise the phylogenetic signal of persistence within this framework by painting branches that belong to the same persistence syndrome in similar colours. In this plot, if persistence has phylogenetic signal, closely related lineages will have the same branch colours (as shown in (C)).

components with lower ε (= equilibrium occupancy) would be more likely to represent sink populations less relevant for overall species persistence. We address the limitations of this approach in the Discussion section.

Even though this aforementioned list of factors affecting persistence estimates is substantial, it is by no means exhaustive, and further processes that can influence parameter estimations should be considered in future applications of any similar investigation.

7 | Application to Real Ecological Datasets: A Test Using Fish Metacommunities

7.1 | Metacommunity Dataset

We applied the inference framework described above to a space–time occupancy grid for marine fishes from the North American continental shelf. The occupancy estimates modelled by our approach were originally calculated by Batt et al. (2017), and correspond to species-cell-year-specific probabilities of presence (ψ_t) estimated through the aggregation of bottom trawl tows made within the same cell and year. The species set in the dataset include generally demersal (i.e., bottom-associated) ray-finned fishes whose sampling allowed the original occupancy model to converge and whose taxonomy was sufficiently consistent through the sampling years (Batt et al. 2017).

The primary data (i.e., presence-absence observations, not modelled by our framework) was collected through sets of marine bottom trawls made by thousands of scientists (e.g., Maureaud et al. 2021, Maureaud et al. 2024). Species were sampled through the northern Atlantic and Pacific continental shelves, which were divided into cells (i.e., sets of 0.5° latitude-longitude, and 50 m depth bins). Cell size can be arbitrary in our framework, but a large number of cells should, all else being equal, provide more information to estimate parameters. To increase species representativity, the dataset should encompass a large proportion of the range of most evaluated species, which should be reasonably met by the empirical dataset evaluated here because it samples a large proportion of the studied continental shelves (Figure 2). When a certain fish species was captured in at least one tow within a cell and year, its ψ_t is equal to one. When the species was not captured, occupancy modelling could still estimate ψ_t as a probability. This estimate is mostly based on the comparison amongst the tow-specific water temperature measurements between the cells where the species was not observed versus the temperature of tows where the species was captured in the same year and region (see occupancy modelling details on Batt et al. 2017). Consequently, our dataset incorporates both imperfect detection—a key frontier in the study of persistence and range change (Beissinger and Riddell 2021)—as well as the effect of temperature on fish habitat use and activity patterns (Pinsky, Selden, and Kitchel 2020). Even though ψ_t are estimates, we will hereafter refer to them as 'Batt et al. (2017) dataset' because this is the information that can be fit by our Markov model.

Considering all Biogeographic regions (hereafter 'regions') and species, the Batt et al. (2017) dataset spans the years between 1970 and 2014. However, the range of years encompassed varied according to region, and sampling frequency varied through time and especially through regions (Table S1). Batt et al. (2017) dataset encompasses 10 different regions, but we aggregated two because they refer to the US West Coast and differ mostly in sampling frequency—which can be ignored because our approach does not require temporal homogeneity in sampling. Additionally, in the instances where more than one region had ψ_t values for the same cell, depth, and year, we summarised these observations in a single region by making ψ_t equal to one if at least one of the duplicates was a one, or by making ψ_t to be the multiplication of the many ψ_t when they all were smaller than one.

Although the original data (few tows for a large-sized cell) and the occupancy modelling create strong limitations toward inference on other aspects of persistence (e.g., spatial variation in abundance), our dataset captures many desirable features that most other data sources do not. For instance,
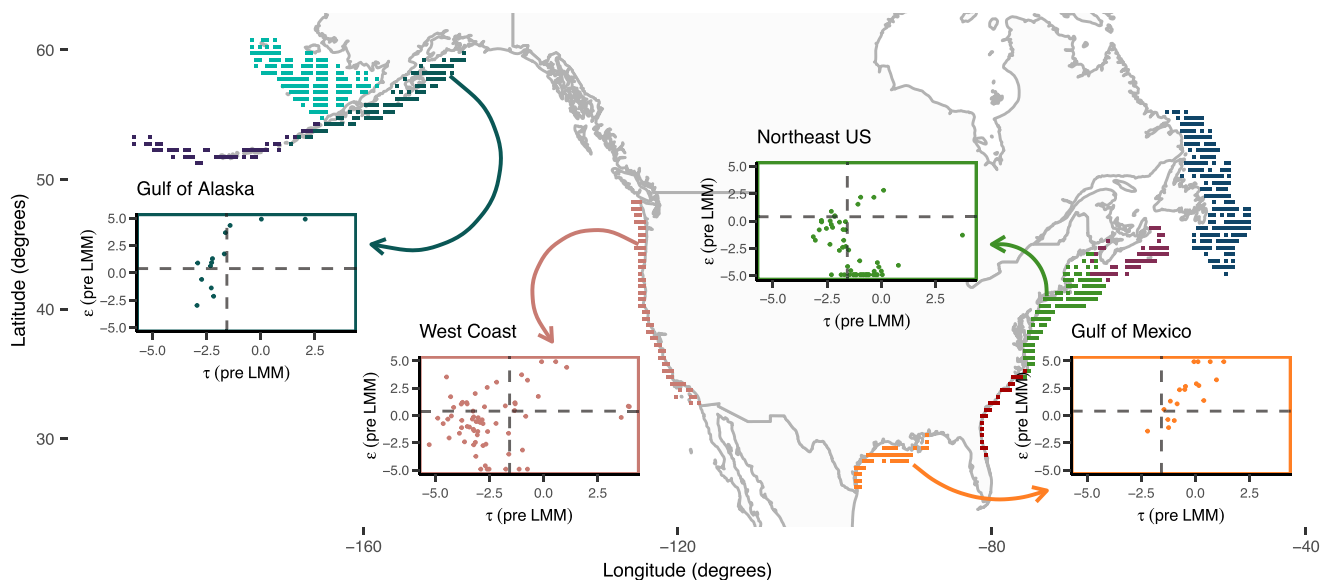
FIGURE 2 | Geographic variation in metapopulation-level persistence of ray-finned fishes. Squares show the spatial positioning of all cells included in this study, coloured by Biogeographic region. Subpanels show examples of region-specific persistence space (i.e., τ and ε estimates for each species–region pair, before the mixed model was employed). Every point within subpanels refers to a species within a region (species–region pair), and different point clouds amongst subpanels show the region-specific trends of persistence within a specific sampled biogeographic region. Dashed grey lines in subplots show the median value on both persistence axes across all species–region pairs (see all regions in Figure S23). τ unit is $\log_2$ metapopulation turnover events per cell per year, and ε unit is log-odds of colonisation.

databases on population-level trends at local scales have been recently assembled (e.g., Capdevila et al. 2020, 2022; Dornelas et al. 2018), but it is hard to reconcile estimates made with different types of data (e.g., sampling methods that require capture with methods that use indirect evidence, or methods based on stage progression matrices with abundance time series). Moreover, dataset standardisation (and thus species comparability) becomes increasingly difficult with phylogenetic distance as appropriate (or even cost-effective) methods for estimating demographic parameters may be context-specific. These difficulties in pairing observations on persistence are particularly challenging in the context of studying speciation rates, as most of the rate variation is associated with distantly related lineages (Singhal et al. 2018, Singhal et al. 2022). Moreover, the spatial coverage of the data is also an important issue as local population dynamics, to the point it is known, seem to be decoupled from global-level measurements of rarity (Daskalova, Myers-Smith, and Godlee 2020), and do not directly incorporate the fact that species can persist by being geographically widespread—meaning that in such cases local abundance plays a minor role in overall persistence. In conclusion, the dataset used here (ψ_t, made by Batt et al. 2017) provides a reasonable compromise (when compared to other data sources) between persistence components that are possible to estimate and that exist at small and large scales.

7.2 | Phylogenetic Signal of Demographic Properties

Because most of the variation in speciation rates occurs amongst distantly related clades, persistence (or any demographic property) can only be a cause of—or correlate with—inferred rates of speciation if it also shows some level of phylogenetic autocorrelation (i.e., phylogenetic signal—Blomberg, Garland, and Ives 2003). Note that, in the absence of phylogenetic signal, the trait would be random with respect to phylogeny and therefore could not be correlated with a phylogenetically structured dependent variable. We are also interested in testing if persistence components evolve independently in the ray-finned fish clade, which we evaluate through a PGLS.

As discussed in the section 'Handling possible sources of noise and bias', it is possible that any phylogenetic signal in persistence metrics may be artifactual due to factors associated with biogeography, which would manifest for instance in the residuals of our persistence metrics, before the use of the mixed model. Thus, we measured the magnitude of phylogenetic autocorrelation and compared it with other potential sources of bias that could create the impression of phylogenetic signal, as region-specific sampling in the estimates before the use of the mixed model. We compared variance explained by phylogenetic and biogeographical factors by subsetting our full dataset, and taking the persistence estimates of only species that occur in more than one region. We then used deviance tables (i.e., analysis of variance) to measure the percentage of variance (residuals) in ϵ and τ explained by region, taxonomic levels (Family, Genus, Species), as well as other factors related to sampling as information content (h_{rel}—see Section S2) and number of transitions.

Although this is not a formal phylogenetic signal test, it measures the percentage of variance in the empirical estimates that co-vary with different factors. If estimates have a strong phylogenetic signal, they are predicted to have a larger proportion of variance explained by taxonomic levels (especially Genus and Family) than by any other factor.

We formally measured the phylogenetic signal of each persistence axis (with biogeographic effects removed by the mixed model) independently through Pagel's Λ (Lynch 1991; Pagel 1999) and tested for significant proximity to the Brownian motion expectation (implemented in phytools—Revell 2012). Within this context, 'Brownian Motion' is a model where the co-variance in persistence amongst a certain set of species depends on the time since their lineages diverged. Finally, phylogenetic patterns of persistence syndromes were visualised by the construction of a phylo-persistence-space (estimated using the phylomorphospace function in phytools—Revell 2012, see details in Section S3.3 and Sidlauskas 2008).

## 7.3 | Testing the Relationship Between Persistence and Speciation

We extracted species-specific ('tip') speciation rates (Title and Rabosky 2019; Vasconcelos, O'Meara, and Beaulieu 2022) for focal taxa from the R package fishtree (Chang et al. 2019). Tip rates were originally estimated by Rabosky et al. (2018) in a phylogeny with thousands of species through the DR statistic (Jetz et al. 2012, hereafter λ_{DR}), or using models in BAMM (Rabosky 2014, hereafter λ_{BAMM}) where speciation rates where constant or allowed to vary through time.

We correlated our persistence components with tip speciation rates using PGLS, as implemented in phylolm (Tung Ho & Ané, 2014). We also used an automatic stepwise model selection procedure to identify the combination of persistence components (ϵ, τ, additive or multiplicative interaction amongst components, or no effect of persistence on speciation—i.e., a 'null model' with just the intercept) that best explain the variation in tip speciation rates within our dataset (based on AIC).

8 | Persistence Patterns at Ecological Timescales

After filtering out datasets without adequate levels of information content (see Sections S2 and S3), Batt et al. (2017) dataset was reduced to 234 species–region pairs (representing 189 species). Amongst these, 15 (6.4%) of the species–region pairs (representing 13 species—i.e., 6.8%) had evidence for spatial heterogeneity in metapopulation rates and thus had their persistence metrics adjusted accordingly (see Section S4). Video 1 illustrates the original Batt data for four species that represent extremes of the persistence space within the same region, providing a visualisation of the persistence syndromes from the perspective of the Batt et al. (2017) dataset. The cells within each region are shown in Figure 2A, and we show examples of region-specific persistence spaces (i.e., τ and ϵ before the mixed model) in the subpanels of Figure 2 (for all regions, see Figure S24).
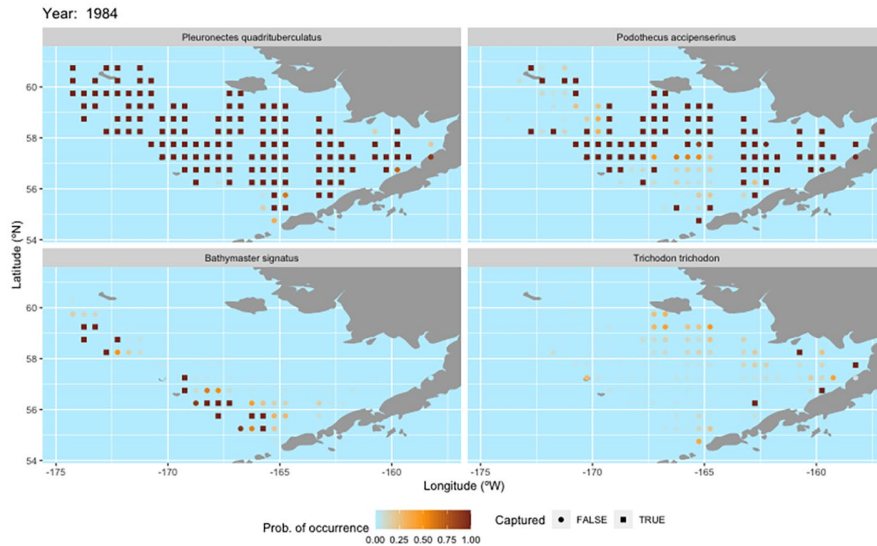
VIDEO 1 | Original data of species at extremes of the persistence space of the East Bering sea. Points show Batt et al. (2017) estimates for every cell with color scale showing ψ_t (probability of occurrence) and squares highlight where $\psi_t = 1$ (i.e., the species was captured in that cell and year). Different frames in the GIF show different sampled years. The species where chosen based on (I) their location (East Bering sea has many cells, all in the same bin depth), (II) the fact these species had the simple (no-mixture) model selected, and (III) their relative positions in the persistence space: *Pleuronectes quadrituberculatus* (Pleuronectidae) occupies the low-τ high-ε region, *Podothecus accipenserinus* (Agonidae) occupies the high-τ high-ε quadrant, *Bathymaster signatus* (Bathymasteridae) occupies the low-τ low-ε region, and *Trichodon trichodon* (Trichodontidae) occupies the high-τ low-ε quadrant. Video content can be viewed at https://onlinelibrary.wiley.com/doi/10.1111/ele.70021

Amongst the 189 species for which we could characterise persistence, 34 (17.9%) occurred in more than one region, and thus allowed the estimation of the proportion of variation in persistence that is attributable to phylogenetic and biogeographical components (Figure 4A). Region explained most variance in persistence components (percentage of residuals explained: 37% for τ; 36.7% for ε), followed by sampling issues (percentage of residuals explained: 20% for τ; 16.9% for ε). However, when one adds the interaction amongst region and taxonomic family, plus all three taxonomic levels, their sum of squares gets closer to the one explained by region (percentage of residuals explained: 31.1% for τ; 34.4% for ε—Figure 4A). Together, all factors analysed explained a large fraction of the variance in our estimates (percentage of residuals explained: 88.7% for τ and 88.1% for ε—Figure 4A).

9 | Persistence Patterns at Geological Timescales

The two persistence components ε and τ were not significantly correlated (slope $p = 0.29$; $R^2 = 0.0004$), which allowed us to treat every persistence axes independently as predictors and as variables to be predicted by traits. When considered in isolation (i.e., in the univariate analysis) only τ has significant phylogenetic signal ($\Lambda_\tau = 0.264$, $p = 0.003$; $\Lambda_\varepsilon = 7.3^{-5}$, $p = 1.0$). Results remain the same if they are evaluated by the pre-mixed model estimates, averaged by species (Table S3). No persistence axis is significant when evaluated in their untransformed scales (results not shown), likely due to these persistence metric geometries not being adequately modelled by Brownian motion (ε_{raw} lies between zero and one, while τ_{raw} is unevenly dispersed between 0.04 and 24 metapopulation turnover events per cell per year).

Some fish families seem to occupy specific regions of the persistence space (Figure 3)—for instance, Scianidae (drums, $n = 8$, Figures 3 and 4C) seem to majorly occupy the low-ε, intermediate-τ quadrant, while both Zoarcidae eelpouts ($n = 10$, Figures 3 and 4D) and Gadidae cods ($n = 9$, Figures 3 and 4E) show sub-clade specific patterns in the persistence space. However, those patterns are not universal as clades with both few (paralichthyid flounders; $n = 7$, Figures 3 and 4F), and many species in our dataset (e.g., Sebastidae rockfishes ($n = 22$, Figures 3 and 4G); Pleuronectidae flounders ($n = 24$, Figures 3 and 4H)) seemed to occupy a very large proportion of the persistence space. The phylogenetic distribution of persistence syndromes in all other species is shown in Figure 3.

10 | Persistence and Speciation Rate

Even though our species pool encompasses many phylogenetically-independent changes in speciation rate (Figure 5A), no significant correlation was found between persistence components (τ and ε) and phylogenetic speciation rates (Table 1, Figure 5B,C). The stepwise procedure selected null models in all estimates of speciation rate (null model intercept estimates ± standard errors: $\lambda_{DR} = -3.55 \pm 0.90$; $\lambda_{BAMM\ (var.\ rate)} = -3.71 \pm 0.54$; $\lambda_{BAMM\ (const.\ rate)} = -3.48 \pm 0.52$, all in $\log_2$ speciation events per lineage per million year). Moreover, we found that both fast and slow-speciating lineages occupy all quadrants of the persistence space (Figure 5D,E). Results on speciation remained qualitatively unchanged if the persistence space was constructed with untransformed variables τ_{raw} and ε_{raw} (results not shown).
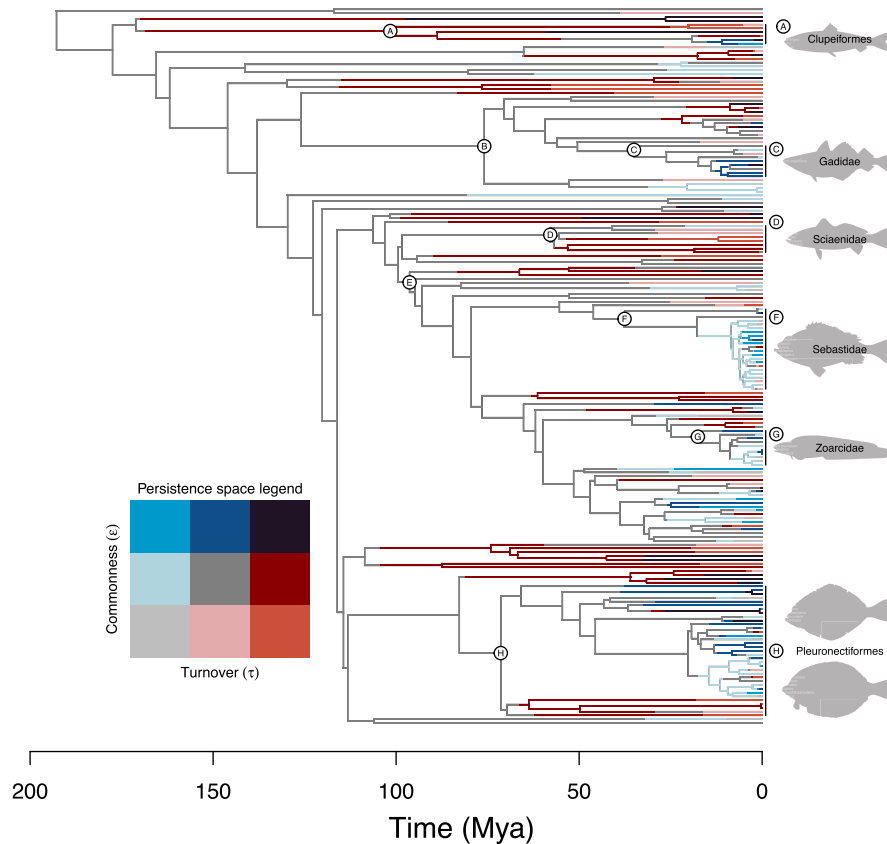
FIGURE 3 | Phylogenetic signal in metapopulation-level persistence of ray-finned fishes. Branch colours and tips are coloured based on their position in the persistence space (details in Section S3.3). Clades with similarly coloured branches emphasise the phylogenetic signal of persistence syndromes (like hypothesized in Figure 1C), while state changes on branches (e.g., multiple colours on single branches) reflect a linear interpolation between node states, projected into the 9-colour persistence space. Consequently, larger or faster colour transitions emphasise branches or clades with more macroevolutionary lability in persistence syndromes. Many of the older branches of the phylogeny are grey, likely because the diffusive nature of the Brownian motion model tends to assign ancestral traits to central parameter values. Labelled nodes and sets of tips grouped by vertical black bars emphasise the species within our dataset that belong to some fish clades with relatively more representation in our dataset (legend: A = Clupeiformes, B = Gadiformes, C = Gadidae, D = Sciaenidae, E = Scorpaeniformes, F = Sebastidae, G = Zoarcidae, H = Pleuronectiformes). All silhouette credits: Nathan Hermann (CC0 1.0).

11 | Discussion

In this perspective, we describe a framework for measuring the properties of species along two axes that contribute to variation in metapopulation persistence and occupancy at large spatial scales. Our conceptual focus is the relationship between metapopulation dynamics and phylogenetic speciation rates, but our approach could easily be applied to other types of comparative data (e.g., traits). We estimated variation in two components that lead to markedly different turnover and occupancy dynamics across a species range (Figure 1): an index of lineage-specific commonness (ε, occupancy at equilibrium), and the turnover rate (τ, a measure of recoverability or resilience). Although the dataset used here was generated by extensive bottom trawl sampling and occupancy modelling (Batt et al. 2017), any other method that estimates ψ_t (e.g., abundance data, mark-recapture, distribution modelling) can also be used in our framework. We explored several potential sources of bias using simulations (see Section S2), but other types of uncertainty related to the underlying data warrant further exploration. In particular, the starting point for our study was the probabilistic occupancy dataset estimated as part of an earlier study (Batt et al. 2017), and thus we did not

address uncertainty in the detection probabilities. This would have entailed re-analysis at the level of primary capture data (e.g., raw trawl data), but can be done with available occupancy models (e.g., Bailey et al. 2014, Dorazio and Royle 2005, Kéry 2010, MacKenzie et al. 2003). Our empirical analyses demonstrated that metrics related to metapopulation-level resilience (i.e., turnover rate) co-vary according to phylogenetic relationships (i.e., have phylogenetic signal).

Ecological processes associated with demographic persistence are increasingly considered to be a key frontier for understanding the mechanisms of speciation (Allmon 1992, Allmon and Sampson 2016, Dynesius and Jansson 2014, Harvey, Singhal, and Rabosky 2019, Levin 2000, Rabosky 2016, Rolland et al. 2023, Stankowski et al. 2024). However, despite extensive and phylogenetically independent variation in both the rate of speciation and demographic properties in our dataset, we found that position in persistence space (i.e., metapopulation properties related to persistence) is uncorrelated with phylogenetic speciation rates.

Our approach measures species-specific persistence syndromes, but it does not directly characterise the fate of intra-specific units
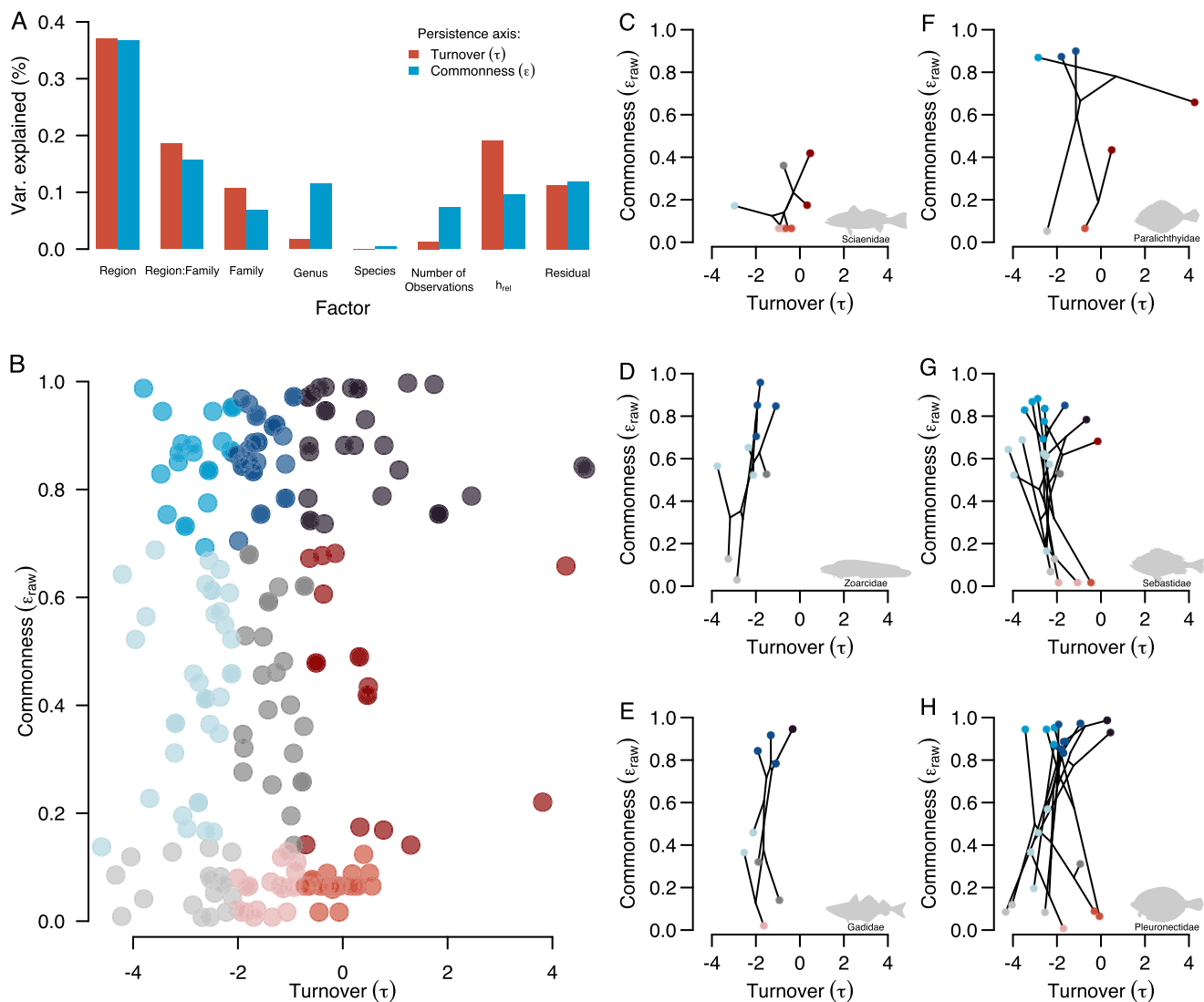
FIGURE 4 | Variation in metapopulation-level persistence. (A) shows the proportion of variance explained by regional and taxonomic factors in the Markov model estimates that only include multi-region species (34 species, 79 region-species pairs), as well as factors related to biogeographic regions, phylogenetic components (here represented by taxonomic levels) and sampling issues (see Section S2 for explanations on h_{rel}), as well as residual variation. (B) shows the empirical persistence space for all species of fishes, with the y-axis showing ε_{raw}. (C–H) show different fish families as filled points (colour-coded as in B), their estimated 'phylo-persistence-space', while the rest of the species are not shown. (B–H) Points representing species are colour-coded by quadrant occupied by that species in the full space. Note black lines connecting edges in (C–H) show within-family phylogeny topology. In this phylomorphospace projection, phylogenetic signal is shown by fewer crosses amongst distant branches, together with slow diffusion in this space as the phylogeny bifurcates (note the difference between (C–E) and (F–H)). τ unit is $\log_2$ metapopulation turnover events per cell per year, and ε_{raw} unit is the proportion of turnover due to colonisation. All silhouette credits: Nathan Hermann (CC0 1.0).

(e.g., incipient species) that give rise to new species-level biodiversity. It is possible that demographic properties we observe at ecological timescales (i.e., decades to centuries), are not representative of processes important at deeper timescales or larger spatial scales (Levin 2000; Harrison 2001; Stanley 2008), and thus metapopulation-level persistence may be unrelated to the factors that mediate the persistence of incipient species. In addition, direct connections between disturbances observed at ecological timescales and environmental change that is relevant to diversification remains poorly known. These issues of scale raise questions about the adequacy of τ and ε as persistence components relevant to speciation. Moreover, other potentially 'rate limiting' steps of the speciation process (e.g., the evolution of reproductive isolation or the ecological differences amongst species) may be unrelated, or even

negatively related, to metapopulation persistence. However, demographic traits—especially those that may correlate with the 'metapopulation dynamics of speciation' (Schluter 2016) are virtually unexplored with respect to phylogenetic speciation rate, and our approach represents a first attempt to link demographic processes occurring over ecological timescales to macroevolutionary patterns. Moreover, regardless of whether τ and ε are good proxies for persistence at scales relevant to speciation, the empirical literature collectively shows a widespread failure of any general type of trait in predicting phylogenetic speciation rates (Anderson et al. 2023; Helmstetter et al. 2023).

Even though persistence syndromes do not predict speciation rates in our dataset, they do correlate with several other
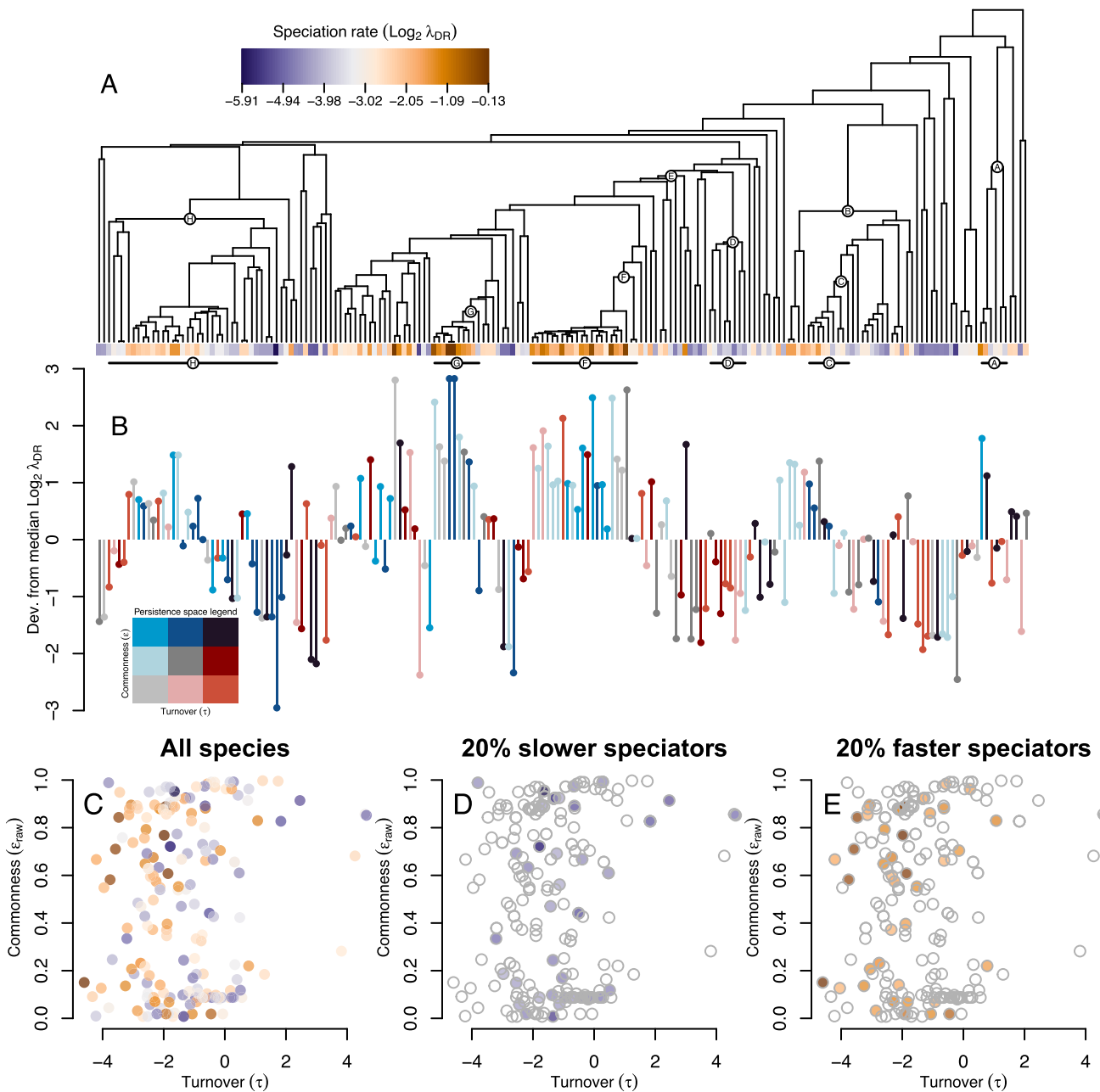
FIGURE 5 | Relationship between phylogenetic speciation rate and ray-finned fish persistence syndromes. (A) Phylogenetic distribution of speciation rates across the ray-finned fish phylogeny pruned to the 189 species in our empirical dataset. Tip colours denote log-transformed tip speciation rates (λ_{DR}). Note that focal set of taxa spans many evolutionarily-independent shifts in speciation rate and in persistence syndromes (for the later, see also Figure 3). Labelled nodes and sets of tips grouped by vertical black bars emphasise the species within our dataset that belong to some fish clades with relatively more representation in our dataset (legend: A = Clupeiformes, B = Gadiformes, C = Gadidae, D = Sciaenidae, E = Scorpaeniformes, F = Sebastidae, G = Zoarcidae, H = Pleuronectiformes—same clades as Figure 3). (B) Tip speciation rates are expressed as residual deviations from the dataset median; colours denote the corresponding persistence syndrome of each tip. Note that persistence syndromes appear randomly scattered across both fast- and slow-speciating lineages. (C) Persistence space (i.e., same as Figure 2B), with species (points) colour-coded according to their values of λ_{DR}. (D) persistence space, but only the 20% slower speciating lineages are colour-coded by their phylogenetic speciation rate, and other lineages are shown as grey points. (E) same as panel D, but in it only the 20% faster speciating lineages are colour-coded. If speciation rate was strongly correlated with persistence, same-coloured points would appear only in specific quadrants in panels C, D or E (see also Figure 1B). Speciation rate unit is $\log_2$ events per lineage per million years. τ unit is $\log_2$ metapopulation turnover events per cell per year, and ε_{raw} unit is the proportion of turnover due to colonisation. Figure S25 shows scatterplots between each individual persistence component and all speciation rate estimates (λ_{DR} and λ_{BAMM}).

important factors that structure fish biodiversity. Biogeographic regions, for instance, differ in solar energy, productivity, bottom topography, sea current dynamics and geological history, amongst other factors that could potentially favour species (or clades) with different persistence syndromes (Allmon 1992; Levin 2000). Indeed, we find that a large percentage of variation

TABLE 1 | Correlation amongst persistence components and different estimates of speciation rate ($\log_2$ transformed). In all cases (speciation estimates), the model selected by the stepwise procedure (= 'best model', $\Delta L = 0.0$) has only the intercept, and thus has no effect on either persistence component or their interaction. Moreover, even the full models (4 parameters total, all shown in this table) provide only slight improvements in prediction (R^2) or likelihood (ΔL), when compared to the selected models. Table S4 shows all parameter estimates. Speciation rate units are $\log_2$ events per lineage per million years.

Speciation rate estimate	ΔL	R^2	τ			ε			$\tau{:}\varepsilon$		
			Est.	S.E.	p	Est.	S.E.	p	Est.	S.E.	p
λ_{BAMM} (var. rate)	1.110	0.011	-0.01	0.01	0.16	0.01	0.01	0.52	-3.9^{-3}	-8.8^{-3}	0.65
λ_{BAMM} (const. rate)	0.977	0.010	-0.01	0.01	0.19	0.01	0.01	0.42	-5.2^{-3}	-8.4^{-3}	0.53
λ_{DR}	0.915	0.009	$-2.0e^{-4}$	0.02	0.99	-0.03	0.02	0.22	-0.03	0.01	0.33

Abbreviations: const. rate = time-continuous diversification rates in BAMM; Est. = estimated value; S.E. = standard error; var. rate = time-variable diversification rates in BAMM; ΔL = Log-likelihood difference from the selected ('best') model; $\tau{:}\varepsilon$ = multiplicative interaction parameter.

in persistence syndromes is explained by biogeographic region (Figure 4A). However, because sampling in our dataset covaries with biogeographic regions (Table S1) we cannot separate the effects of these two components. Moreover, the mixed model employed here to remove sampling effects might have also removed biologically relevant effects (see Beissinger and Riddell 2021 for examples) that may be associated with biogeographic regions, with cascading effects in posterior analyses (phylogenetic signal, PGLS).

One advantage of measuring persistence at the metapopulation level is that doing so allows relatively direct observation of extirpation events. These are not the same as true extinctions of incipient species, whose persistence estimation may require large fossil datasets of relatively higher temporal resolution. While truly exceptional fossil records exist (e.g., Fenton et al. 2021; Graham 1994), it is unclear whether they allow both the identification of incipient species as well as their persistence. Furthermore, robust comparative tests would require a sufficiently detailed record to estimate amongst-clade variance in persistence. It is also unclear whether other indirect methods, including estimates of genetic variation through time (e.g., Carnaval et al. 2009), distribution modelling (Tessarolo et al. 2021) and recent extinctions (e.g., Quaternary extinctions—Lorenzen et al. 2011) provide meaningful proxies for the persistence of incipient species.

The persistence components evaluated in our approach at ecological timescales open up post hoc ecological questions such as 'Do higher-turnover fish species recover their occupancy faster after disturbances like heat waves?' (see also Fredston et al. 2023). These questions are outside the scope of our investigation here, but such estimates could be used to explain events of the past or to predict faunal responses to specific threats (though such responses can be complex—Fredston et al. 2023). At geological timescales, spatial occupancy may be more relevant to macroevolutionary dynamics than local abundance (Harnik 2011, Harnik, Simpson, and Payne 2012, see also Daskalova, Myers-Smith, and Godlee 2020), and may correlate with geographic range dynamics (Batt et al. 2017, Beissinger and Riddell 2021, Pinsky, Selden, and Kitchel 2020)—one of the best predictors of extinction rates (Harnik 2011, Harnik, Simpson, and Payne 2012, Harvey, Singhal, and Rabosky 2019, Jablonski 2008, Levin 2000, Smits 2015).

The species evaluated here occupy most of the theoretically possible regions of ε_{raw}, and most of the τ values fall within our observable timescale (i.e., $\tau \in [2^{-7}, 2^4]$), indicating relevant interspecific variation in metapopulation dynamics. However, when both persistence axes are considered together, species seem to not occupy the persistence space associated with high τ low ε_{raw} (Figure 4B). This is the quadrant that technically is the least 'stable' of the persistence space (high turnover closer to zero occupancy), which might imply that our estimates are biologically meaningful. However, this 'hole' in the persistence space may also reflect species that are rarely captured by bottom trawls or that are associated with information-poor data (e.g., low h_{rel}), and were thus removed by our dataset filtering. Moreover, interpretations based on holes in spaces constructed with comparative data should be seen with scepticism due to the fact that many different processes can generate them (Polly 2023). An additional limitation of our approach is that our data may be biased against infrequently observed species: by virtue of their 'rarity' (Rabinowitz 1981), they may be insufficiently sampled (or not sampled at all) to allow robust estimates of metapopulation parameters—which limits correlations between metrics of extinction risk and our persistence components. This also illustrates a major limitation of any research on persistence: rare species may provide critical data points, and yet it will be precisely these taxa that are most poorly represented in comparative datasets. Finally, metapopulation dynamics depend not only on biological attributes but also on the variation in patch suitability (Hanski 1999; Harrison 1979; Levins 1969), and in metapopulation size—which our approach does not measure. Thus it is possible that at least a fraction of our persistence syndromes (τ and ε) reflect not an intrinsic property of a species, but the way its population is particularly affected by natural, temporal variation in patch suitability (Levin 2000). This is arguably a component of the way species persist, but interpretation is much less straightforward because it involves the interaction between lineage-specific and environment-specific factors.

By using a mixture model for parameter estimation, we attempted to accommodate some amongst-cell heterogeneity in colonisation-extirpation dynamics, including metapopulations that conform to simple source-sink dichotomies. Surprisingly, only 6.8% of species were best fit by the mixture model, but many uncertainties remain. It is also unclear if choosing the mixture with the highest ε is the best protocol to represent species (see Section S4). These results also raise questions related to

spatial completeness and the estimation of metapopulation-level persistence. These issues highlight the difficulties of characterising a species' complex trait by only a pair of values and the necessity of explicitly incorporating intraspecific variation and spatial breadth in future analyses (see also Section S4).

Our conclusions are conditional on our estimates of tip-level speciation rates and on the phylogeny assumed for comparative analysis, which implies limitations common to most (if not all) studies of phylogenetic speciation rates. For instance, taxonomic oversplitting/lumping will bias speciation rates (Freeman and Pennell 2021). Differential taxonomic sampling across clades in the phylogeny used to estimate speciation can also bias results (Chang et al. 2020). Similarly, amongst-clade variability in rates of molecular evolution can bias branch length estimates (Bromham and Cardillo 2019) and thus speciation rates. Our estimates also do not account for phylogenetic uncertainty (i.e., variation in topology or branch lengths amongst equally plausible phylogenies), which might be consequential to both speciation rate and the PGLS framework. Nonetheless, the 'tip rates' used here are relatively robust to other factors that bias speciation rate estimates through deep time (Louca and Pennell 2020, Maliet et al. 2019, Title and Rabosky 2019, Vasconcelos, O'Meara, and Beaulieu 2022). Similarly to all other studies of trait-dependent diversification, our study assumes at least some degree of consistency in the relationship between speciation and metapopulation persistence (our focal trait) through time. However, our focus on recent speciation events (i.e., the use of tip rates) does not assume homogeneity across the entire history of the clade, because observed relationships will generally reflect variation in speciation rate that has arisen during the past 20 million years or less (Rabosky et al. 2018).

The demography of speciation represents a key frontier for understanding large-scale patterns of biological diversity. Researchers have long recognised that the 'success' of incipient species is an important component of the speciation process (Allmon 1992; Allmon and Sampson 2016; Glazier 1987; Mayr 1963; Stanley 1979, 2008), but only recently we have accumulated the appropriate comparative tools and data to characterise the phylogenetic structure and macroevolutionary implications of persistence and other demographic traits. Here, we provide a roadmap for estimating these dynamic features of contemporary populations and for assessing their significance for biodiversity patterns more generally. Using an example dataset on marine fishes, we demonstrate that persistence syndromes are estimable and appear to covary amongst closely related species. However, ecological-time persistence estimates cannot predict phylogenetic speciation rates. We still have much to learn about how best to extract meaningful and comparable persistence estimates amongst species, and the framework we present here is a tangible step towards this important goal in biodiversity theory.

Author Contributions

Data Availability Statement

Data and scripts are archived in DRYAD (link: https://doi.org/10.5061/dryad.51c59zwgv). Supporting Informations are archived in Zenodo (link: https://doi.org/10.5281/zenodo.13730262).

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ele.70021.

References

Allmon, W. D. 1992. "A Causal Analysis of Stages in Allopatric Speciation." In *Oxford Surveys in Evolutionary Biology*, edited by D. Futuyma and J. Antonovics, 219–258. New York, USA: Oxford University Press.

Allmon, W. D., and S. D. Sampson. 2016. "The Stages of Speciation: A Stepwise Framework for Analysis of Speciation in the Fossil Record." In *Species and Speciation in the Fossil Record*, edited by M. M. Yacobucci and W. D. Allmon, 121–167. Chicago and London: University of Chicago Press.

Alonso, D., A. Pinyol-Gallemí, T. Alcoverro, and R. Arthur. 2015. "Fish Community Reassembly After a Coral Mass Mortality: Higher Trophic Groups Are Subject to Increased Rates of Extinction." *Ecology Letters* 18: 451–461.

Anderson, B., J. Pannell, S. Billiard, et al. 2023. "Opposing Effects of Plant Traits on Diversification." *iScience* 26: 106362.

Anderson, S. A. S., and J. T. Weir. 2022. "The Role of Divergent Ecological Adaptation During Allopatric Speciation in Vertebrates." *Science* 378: 1214–1218.

Bailey, L. L., D. I. MacKenzie, and J. D. Nichols. 2014. "Advances and Applications of Occupancy Models." *Methods in Ecology and Evolution* 5: 1269–1279.

Batt, R. D., J. W. Morley, R. L. Selden, M. W. Tingley, and M. L. Pinsky. 2017. "Gradual Changes in Range Size Accompany Long-Term Trends in Species Richness." *Ecology Letters* 20: 1148–1157.

Beissinger, S. R., and E. A. Riddell. 2021. "Why Are Species' Traits Weak Predictors of Range Shifts?" *Annual Review of Ecology, Evolution, and Systematics* 52: 47–66.

BlackRMills. 2021. "MetBrewer: Color Palettes Inspired by Works at the Metropolitan Museum of Art."

Blomberg, S. P., J. R. T. Garland, and A. R. Ives. 2003. "Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile." *Evolution* 57: 717–745.

Bromham, L., and M. Cardillo. 2019. "Was the Diversification of Mammals due to Luck?" In *Origins of Biodiversity: An Introduction to Macroevolution and Macroecology*. Oxford, England: Oxford University Press.

Burbrink, F. T., S. Ruane, N. Rabibisoa, A. P. Raselimanana, C. J. Raxworthy, and A. Kuhn. 2023. "Speciation Rates Are Unrelated to the Formation of Population Structure in Malagasy Gemsnakes." *Ecology and Evolution* 13: e10344.

Cally, J. G., D. Stuart-Fox, L. Holman, J. Dale, and I. Medina. 2021. "Male-Biased Sexual Selection, but Not Sexual Dichromatism, Predicts Speciation in Birds." *Evolution* 75: 931–944.

Capdevila, P., I. Stott, M. Beger, and R. Salguero-Gómez. 2020. "Towards a Comparative Framework of Demographic Resilience." *Trends in Ecology & Evolution* 35: 776–786.

Capdevila, P., I. Stott, J. Cant, et al. 2022. "Life History Mediates the Trade-Offs Among Different Components of Demographic Resilience." *Ecology Letters* 25: 1566–1579.

Carnaval, A. C., M. J. Hickerson, C. F. B. Haddad, M. T. Rodrigues, and C. Moritz. 2009. "Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot." *Science* 323: 785–789.

Chang, J., D. L. Rabosky, S. A. Smith, and M. E. Alfaro. 2019. "An *r* Package and Online Resource for Macroevolutionary Studies Using the Ray-Finned Fish Tree of Life." *Methods in Ecology and Evolution* 10: 1118–1124.

Chang, J., D. L. Rabosky, and M. Alfaro. 2020. "Estimating Diversification Rates on Incompletely Sampled Phylogenies: Theoretical Concerns and Practical Solutions." *Systematic Biology* 69: 602–611.

Coyne, J. A., and H. A. Orr. 2004. "Speciation and Macroevolution." In *Speciation*, 411–445. Sunderland, Massachusetts: Sinauer associates.

Cooney, C. R., and G. H. Thomas. 2021. "Heterogeneous Relationships Between Rates of Speciation and Body Size Evolution Across Vertebrate Clades." *Nature Ecology & Evolution* 5: 101–110.

Daskalova, G. N., I. H. Myers-Smith, and J. L. Godlee. 2020. "Rare and Common Vertebrates Span a Wide Spectrum of Population Trends." *Nature Communications* 11: 4394.

Dorazio, R. M., and J. A. Royle. 2005. "Estimating Size and Composition of Biological Communities by Modeling the Occurrence of Species." *Journal of the American Statistical Association* 100: 389–398.

Dornelas, M., L. H. Antão, F. Moyes, et al. 2018. "BioTIME: A Database of Biodiversity Time Series for the Anthropocene." *Global Ecology and Biogeography* 27: 760–786.

Dynesius, M., and R. Jansson. 2014. "Persistence of Within-Species Lineages: A Neglected Control of Speciation Rates." *Evolution* 68: 923–934.

Economo, E. P., N. Narula, N. R. Friedman, M. D. Weiser, and B. Guénard. 2018. "Macroecology and Macroevolution of the Latitudinal Diversity Gradient in Ants." *Nature Communications* 9: 1778.

Felsenstein, J. 1985. "Phylogenies and the Comparative Method." *American Naturalist* 125: 1–15.

Fenton, I. S., A. Woodhouse, T. Aze, et al. 2021. "Triton, a New Species-Level Database of Cenozoic Planktonic Foraminiferal Occurrences." *Scientific Data* 8: 160.

Foisy, M. R., L. P. Albert, D. W. W. Hughes, and M. G. Weber. 2019. "Do Latex and Resin Canals Spur Plant Diversification? Re-Examining a Classic Example of Escape and Radiate Coevolution." *Journal of Ecology* 107: 1606–1619.

Fredston, A. L., W. W. L. Cheung, T. L. Frölicher, et al. 2023. "Marine Heatwaves are not a Dominant Driver of Change in Demersal Fishes." *Nature* 621: 324–329.

Freeman, B. G., and M. W. Pennell. 2021. "The Latitudinal Taxonomy Gradient." *Trends in Ecology & Evolution* 36: 778–786.

Freeman, B. G., J. Rolland, G. A. Montgomery, and D. Schluter. 2022. "Faster Evolution of a Premating Reproductive Barrier Is Not Associated With Faster Speciation Rates in New World Passerine Birds." *Proceedings of the Royal Society B: Biological Sciences* 289: 20211514.

Germain, R. M., S. P. Hart, M. M. Turcotte, et al. 2021. "On the Origin of Coexisting Species." *Trends in Ecology & Evolution* 36: 284–293.

Gingerich, P. D. 2019. "Introduction." In *Rates of Evolution: A Quantitative Synthesis*, 1–25. Cambridge, United Kingdom: Cambridge University Press.

Glazier, D. S. 1987. "Toward a Predictive Theory of Speciation: The Ecology of Isolate Selection." *Journal of Theoretical Biology* 126: 323–333.

Gotelli, N. J. 1991. "Metapopulation Models: The Rescue Effect, the Propagule Rain, and the Core-Satellite Hypothesis." *American Naturalist* 138: 768–776.

Graham, R. W. 1994. "FAUNMAP: A Database Documenting Late Quaternary Distributions of Mammal Species in the United States." Illinois State Museum, 25.

Grimm, V., K. Reise, and M. Strasser. 2003. "Marine metapopulations: A Useful Concept?" *Helgoland Marine Research* 56: 222–228.

Hanski, I. 1999. *Metapopulation Ecology. Oxford Series in Ecology and Evolution*. Oxford, England: Oxford University Press.

Harnik, P. G. 2011. "Direct and Indirect Effects of Biological Factors on Extinction Risk in Fossil Bivalves." *Proceedings of the National Academy of Sciences* 108: 13594–13599.

Harnik, P. G., C. Simpson, and J. L. Payne. 2012. "Long-Term Differences in Extinction Risk Among the Seven Forms of Rarity." *Proceedings of the Royal Society B: Biological Sciences* 279: 4969–4976.

Harrison, G. W. 1979. "Stability Under Environmental Stress: Resistance, Resilience, Persistence, and Variability." *American Naturalist* 113: 659–669.

Harrison, S. 2001. "Do Taxa Persist as Metapopulations in Evolutionary Time?" In *Biodiversity Dynamics. Turnover of Populations, Taxa, and Communities*, 19–30. New York, USA: Columbia University Press.

Harrison, X. A., L. Donaldson, M. E. Correa-Cano, et al. 2018. "A Brief Introduction to Mixed Effects Modelling and Multi-Model Inference in Ecology." *PeerJ* 6: e4794.

Harvey, M. G., G. F. Seeholzer, B. T. Smith, D. L. Rabosky, A. M. Cuervo, and R. T. Brumfield. 2017. "Positive Association Between Population Genetic Differentiation and Speciation Rates in New World Birds." *Proceedings of the National Academy of Sciences* 114: 6328–6333.

Harvey, M. G., S. Singhal, and D. L. Rabosky. 2019. "Beyond Reproductive Isolation: Demographic Controls on the Speciation Process." *Annual Review of Ecology, Evolution, and Systematics* 50: 75–95.

Helmstetter, A. J., R. Zenil-Ferguson, H. Sauquet, et al. 2023. "Trait-Dependent Diversification in Angiosperms: Patterns, Models and Data." *Ecology Letters* 26: 640–657.

Henao-Diaz, L. F., and M. Pennell. 2023. "The Major Features of Macroevolution." *Systematic Biology* 72: 1188–1198.

Houle, D., C. Pélabon, G. P. Wagner, and T. F. Hansen. 2011. "Measurement and Meaning in Biology." *Quarterly Review of Biology* 86: 3–34.

Jablonski, D. 2008. "Species Selection: Theory and Data." *Annual Review of Ecology, Evolution, and Systematics* 39: 501–524.

Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, and A. O. Mooers. 2012. "The Global Diversity of Birds in Space and Time." *Nature* 491: 444–448.

Kéry, M. 2010. "Nonstandard GLMMs 1: Site-Occupancy Species Distribution Model." In *Introduction to WinBUGS for Ecologists*, edited by M. Kéry, 237–252. Boston, MA: Academic Press.

Kraaijeveld, K., F. J. L. Kraaijeveld-Smit, and M. E. Maan. 2011. "Sexual Selection and Speciation: The Comparative Evidence Revisited." *Biological Reviews* 86: 367–377.

Leibold, M. A., and J. M. Chase. 2018. *Metacommunity Ecology. Monographs in Population Biology.* Princeton, Massachusetts: Princeton University Press.

Levin, D. 2000. "The Origin, Expansion, and Demise of Plant Species." In *Oxford Series in Ecology and Evolution.* Oxford, England: Oxford University Press.

Levins, R. 1969. "Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control." *Bulletin of the Entomological Society of America* 15: 237–240.

Lorenzen, E. D., D. Nogués-Bravo, L. Orlando, et al. 2011. "Species-Specific Responses of Late Quaternary Megafauna to Climate and Humans." *Nature* 479: 359–364.

Louca, S., and M. W. Pennell. 2020. "Extant Timetrees Are Consistent With a Myriad of Diversification Histories." *Nature* 580: 502–505.

Lynch, M. 1991. "Methods for the Analysis of Comparative Data in Evolutionary Biology." *Evolution* 45: 1065–1080.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. "Estimating Site Occupancy, Colonization, and Local Extinction When a Species is Detected Imperfectly." *Ecology* 84: 2200–2207.

MacLean, S. A., and S. R. Beissinger. 2017. "Species' Traits as Predictors of Range Shifts Under Contemporary Climate Change: A Review and Meta-Analysis." *Global Change Biology* 23: 4094–4105.

Maliet, O., F. Hartig, and H. Morlon. 2019. "A Model With Many Small Shifts for Estimating Species-Specific Diversification Rates." *Nature Ecology & Evolution* 3: 1086–1092.

Maureaud, A., R. Frelat, L. Pécuchet, et al. 2021. "Are We Ready to Track Climate-Driven Shifts in Marine Species Across International Boundaries? – A Global Survey of Scientific Bottom Trawl Data." *Global Change Biology* 27: 220–236.

Matute, D. R., and B. S. Cooper. 2021. "Comparative Studies on Speciation: 30 Years Since Coyne and Orr." *Evolution* 75: 764–778.

Maureaud, A. A., J. Palacios-Abrantes, Z. Kitchel, et al. 2024. "FISHGLOB_data: An Integrated Dataset of Fish Biodiversity Sampled With Scientific Bottom-Trawl Surveys." *Scientific Data* 11: 24.

Mayr, E. 1963. "The Genetics of Speciation." In *Animal Species and Evolution*, 516–556. Cambridge, USA: Harvard University Press.

Nitschke, C. R., M. Hourston, V. Udyawer, and K. L. Sanders. 2018. "Rates of Population Differentiation and Speciation Are Decoupled in Sea Snakes." *Biology Letters* 14: 20180563.

Pagel, M. 1994. "Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 255: 37–45.

Pagel, M. 1999. "Inferring the Historical Patterns of Biological Evolution." *Nature* 401: 877–884.

Panhuis, T. M., R. Butlin, M. Zuk, and T. Tregenza. 2001. "Sexual Selection and Speciation." *Trends in Ecology & Evolution* 16: 364–371.

Pinsky, M. L., R. L. Selden, and Z. J. Kitchel. 2020. "Climate-Driven Shifts in Marine Species Ranges: Scaling From Organisms to Communities." *Annual Review of Marine Science* 12: 153–179.

Polly, P. D. 2023. "Extinction and Morphospace Occupation: A Critical Review." *Cambridge Prisms: Extinction* 1: e17.

Rabinowitz, D. 1981. "Seven Forms of Rarity." In *The Biological Aspects of Rare Plant Conservation*, edited by H. Synge, 205–217. Chichester, UK: Wiley.

Rabosky, D. L. 2014. "Automatic Detection of Key Innovations, Rate Shifts, and Diversity Dependence on Phylogenetic Trees." *PLoS One* 9: e89543.

Rabosky, D. L. 2016. "Reproductive Isolation and the Causes of Speciation Rate Variation in Nature." *Biological Journal of the Linnean Society* 118: 13–25.

Rabosky, D. L. 2022. "Evolutionary Time and Species Diversity in Aquatic Ecosystems Worldwide." *Biological Reviews* 97: 2090–2105.

Rabosky, D. L., J. Chang, P. O. Title, et al. 2018. "An Inverse Latitudinal Gradient in Speciation Rate for Marine Fishes." *Nature* 559: 392–395.

Rabosky, D. L., and D. R. Matute. 2013. "Macroevolutionary Speciation Rates Are Decoupled From the Evolution of Intrinsic Reproductive Isolation in Drosophila and Birds." *Proceedings of the National Academy of Sciences* 110: 15354–15359.

Revell, L. J. 2012. "Phytools: An R Package for Phylogenetic Comparative Biology (And Other Things)." *Methods in Ecology and Evolution* 3: 217–223.

Ricklefs, R. E. 2007. "Estimating Diversification Rates From Phylogenetic Information." *Trends in Ecology & Evolution* 22: 601–610.

Rolland, J., L. F. Henao-Diaz, M. Doebeli, et al. 2023. "Conceptual and Empirical Bridges Between Micro- and Macroevolution." *Nature Ecology & Evolution* 7: 1181–1193.

Schielzeth, H., N. J. Dingemanse, S. Nakagawa, et al. 2020. "Robustness of Linear Mixed-Effects Models to Violations of Distributional Assumptions." *Methods in Ecology and Evolution* 11: 1141–1152.

Schluter, D. 2016. "Speciation, Ecological Opportunity, and Latitude." *American Naturalist* 187: 1–18.

Schluter, D., and M. W. Pennell. 2017. "Speciation Gradients and the Distribution of Biodiversity." *Nature* 546: 48–55.

Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: 379–423.

Sidlauskas, B. 2008. "Continuous and Arrested Morphological Diversification in Sister Clades of Characiform Fishes: A Phylomorphospace Approach." *Evolution* 62: 3135–3156.

Simberloff, D. S. 1969. "Experimental Zoogeography of Islands: A Model for Insular Colonization." *Ecology* 50: 296–314.

Singhal, S., G. R. Colli, M. R. Grundler, G. C. Costa, I. Prates, and D. L. Rabosky. 2022. "No Link Between Population Isolation and Speciation Rate in Squamate Reptiles." *Proceedings of the National Academy of Sciences* 119: e2113388119.

Singhal, S., H. Huang, M. R. Grundler, et al. 2018. "Does Population Structure Predict the Rate of Speciation? A Comparative Test Across Australia's Most Diverse Vertebrate Radiation." *American Naturalist* 192: 432–447.

Smits, P. D. 2015. "Expected Time-Invariant Effects of Biological Traits on Mammal Species Duration." *Proceedings of the National Academy of Sciences of the United States of America* 112: 13015–13020.

Spake, R., D. E. Bowler, C. T. Callaghan, et al. 2023. "Understanding 'It Depends' in Ecology: A Guide to Hypothesising, Visualising and Interpreting Statistical Interactions." *Biological Reviews* 98: 983–1002.

Stankowski, S., A. D. Cutter, I. Satokangas, et al. 2024. "Toward the Integration of Speciation Research." *Evolutionary Journal of the Linnean Society* 3: kzae001.

Stankowski, S., and M. Ravinet. 2021. "Defining the Speciation Continuum." *Evolution* 75: 1256–1273.

Stanley, S. M. 1979. "The Nature of Quantum Speciation." In *Macroevolution: Pattern and Process*, 143–178. San Francisco, USA: W. H. Freeman and Company.

Stanley, S. M. 2008. "Predation Defeats Competition on the Seafloor." *Paleobiology* 34: 1–21.

Strier, K. B. 2021. "The Limits of Resilience." *Primates* 62: 861–868.

Symonds, M. R. E., and S. P. Blomberg. 2014. "A Primer on Phylogenetic Generalised Least Squares." In *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*, edited by L. Z. Garamszegi, 105–130. Berlin, Heidelberg: Springer.

Tessarolo, G., R. J. Ladle, J. M. Lobo, T. F. Rangel, and J. Hortal. 2021. "Using Maps of Biogeographical Ignorance to Reveal the Uncertainty in Distributional Data Hidden in Species Distribution Models." *Ecography* 44: 1743–1755.

Title, P. O., and D. L. Rabosky. 2019. "Tip Rates, Phylogenies and Diversification: What Are We Estimating, and How Good Are the Estimates?" *Methods in Ecology and Evolution* 10: 821–834.

Tung Ho, L., and C. Ané. 2014. "A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models." *Systematic Biology* 63: 397–408.

Vasconcelos, T., B. C. O'Meara, and J. M. Beaulieu. 2022. "A Flexible Method for Estimating Tip Diversification Rates Across a Range of Speciation and Extinction Scenarios." *Evolution* 76: 1420–1433.

Wacker, K. S., and B. M. Winger. 2024. "An Elevational Phylogeographic Diversity Gradient in Neotropical Birds Is Decoupled From Speciation Rates." *American Naturalist* 203: 362–381.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.