Comparing speciation rates in lakes, rivers, and the sea

Elizabeth Christina Miller[1]

[1]School of Aquatic and Fishery Sciences, University of Washington, Seattle, Washington 98105, U.S.A.

Correspondence to: emille3@uw.edu

ABSTRACT:

The diversity of species inhabiting freshwater relative to marine habitats is striking, given that freshwater habitats make up less than 1% of Earth's water. The most commonly proposed mechanism for this pattern is that freshwater habitats are more fragmented than marine habitats, resulting in increased opportunities for allopatric speciation in freshwater and thus increased speciation rates relative to marine habitats. However, some authors suggest that speciation is faster in sympatry than allopatry, as illustrated by lacustrine radiations such as African cichlids. Sympatric speciation may be common in lakes, while allopatric speciation may be dominant in rivers. If allopatric speciation is inherently slower than sympatric speciation, we might expect differences in rates of speciation between lineages primarily confined to riverine versus lacustrine habitats. These differences within freshwater habitats may be important to consider when comparing diversification rates between freshwater and marine lineages. Here I compared rates of speciation associated with ray-finned fishes in marine, riverine and lacustrine habitats. I also used geographic range data to compare rates of speciation associated with allopatric and sympatric speciation modes in freshwater fishes. I found that lakes had faster rates of speciation than other aquatic habitats. However, most freshwater fish diversity arose through allopatric speciation in rivers, despite their slower speciation rate. Speciation rates were more strongly correlated with habitat (riverine or lacustrine) than present-day range overlap (allopatry or sympatry). Surprisingly, riverine and marine habitats had similar rates of speciation. Models of evolution suggest that lacustrine habitats are evolutionary unstable, and biased transition rates explain the dearth of lacustrine species. Collectively, these results suggest that while allopatric speciation is important for generating freshwater diversity, this mechanism may not be the cause of faster speciation rates between freshwater and marine habitats.

KEY WORDS

Fishes, allopatric speciation, sympatric speciation, speciation rates, lacustrine radiations, diversification

INTRODUCTION:

The great biodiversity of freshwater habitats is striking. Freshwater habitats make up 0.009% of the Earth's water but contain ~40% of ray-finned fish species (Horn 1972). The diversity of freshwater relative to marine fishes is thought to be due to the greater fragmentation of freshwater habitats, creating more opportunities for population subdivision than in the ocean (Horn 1972; May 1994). As such, speciation rates are thought to be higher in freshwater than marine lineages of fishes (Grosberg et al. 2012), although this hypothesis has mixed empirical evidence (Carrete Vega and Wiens 2012; Bloom et al. 2013; Tedesco et al. 2017a, Miller et al. 2018; Betancur-R et al. 2015; Rabosky 2020).

This proposed mechanism for the difference in speciation rates between biomes implies that allopatric speciation has generated much of the world's freshwater diversity (Lynch 1989). This may be the case, as major river systems such as the Amazon are among the most diverse freshwater habitats (Lévêque et al. 2008). However, lakes are also known for their diverse faunas (Lévêque et al. 2008). Ancient lakes such as the African rift lakes and Lake Baikal contain species flocks, or monophyletic groups that underwent rapid speciation within a single lake (Cohen 1995; Cristescu et al. 2010; Wagner et al. 2012; Seehausen 2015). Lacustrine radiations represent some of the best-supported cases of sympatric speciation (Coyne and Orr 2004; Bolnick and Fitzpatrick 2007; Seehausen and Wagner 2014). The relative contribution of allopatric speciation in rivers and sympatric speciation in lakes to freshwater diversity overall is an open question (Seehausen and Wagner 2014).

A related question is the pace of speciation inherent to allopatric versus sympatric speciation modes. Speciation is thought to proceed slowly in allopatry if genetic differences between isolated populations accumulate passively (genetic drift) or through gradual local

adaptation. Reproductive isolation may be achieved faster in sympatry because reinforcement and disruptive selection may actively select against hybridization (Bush 1975; McCune and Lovejoy 1998; Crow et al. 2010). McCune and Lovejoy (1998) showed that sympatric lacustrine sister-species had fewer genetic differences between them than allopatric riverine sister-species, suggesting that reproductive isolation is achieved faster in sympatry. However, their study was limited by the few sister-species pairs with genetic data and statistical methods at the time. Later studies in other fishes found that speciation can also be rapid in allopatry (Hrbek et al. 2002; Near and Bernard 2004). It is unclear what generalities can be drawn from these studies in select groups.

Twenty years after McCune and Lovejoy's (1998) study, the widespread availability of genetic data in public repositories and improved models of diversification allow macroevolutionary hypotheses to be tested in a broader, more rigorous context. A cursory examination of diversification-rate variation among freshwater fish clades (Rabosky et al. 2013; Seehausen and Wagner 2014) suggests that some species-rich clades do not have fast rates (e.g. cyprinids), while some groups with fast rates are not species-rich (e.g. sticklebacks). This variation may be related to contrasting speciation mechanisms among freshwater habitats. Allopatric speciation may be the dominant mode in rivers, because rivers become fragmented or change course through time (Dias et al. 2013). In contrast, rapid sympatric speciation may be primarily found in lakes, because the dimensionality of lakes provides opportunities for disruptive selection and nonrandom mating (Bolnick and Fitzpatrick 2007; Seehausen 2015). At least one case of sympatric speciation within rivers is likely (mormyrid fishes; Sullivan et al. 2002), but in general sympatric speciation may be rare in rivers. If speciation rates are inherently different in allopatry versus sympatry, then we may also expect riverine and lacustrine lineages

to show different rates of speciation. The hypothesis that riverine and lacustrine habitats differ in rates of speciation has been proposed before (Briggs 1966; Lowe-McConnell 1969; Seehausen 2015); but has not been explicitly tested using modern approaches.

The difference in rates between freshwater habitats may be a crucial missing piece to interpreting comparisons between freshwater and marine diversity. Specifically, it is possible that most freshwater diversity arose via allopatry, but these allopatric species accumulated over long periods of time rather than at a fast rate. If so, higher speciation rates in freshwater relative to marine habitats may actually be driven by rapid sympatric speciation in lakes. This would call into question the most commonly proposed mechanism for why speciation rates seem to differ between biomes (opportunities for allopatric speciation).

In this study, I compared rates of speciation between marine and freshwater habitats across ray-finned fishes. However, unlike past studies, I compared speciation rates in riverine and lacustrine habitats separately rather than treating freshwater as a single category. I also used geographic range data for well-sampled groups to compare modes of speciation in rivers and lakes. I found that speciation rates in riverine and marine lineages are actually similar. However, speciation rates in lakes are faster than other aquatic habitats. These results imply that nuance is needed when comparing the origins of marine and freshwater biodiversity.

MATERIAL AND METHODS:

Data acquisition:

I used the largest available time-calibrated phylogeny of actinopterygian fishes for all analyses (Rabosky et al. 2018). I used the maximum likelihood phylogeny including species with

genetic data only. This phylogeny originally contained 11,638 tips. I removed 51 tips that were duplicate or unresolved species.

To begin identifying species occurring in freshwater habitats, I used occurrence records from Tedesco et al. (2017b). In brief, this dataset is based on 1,436 literature sources and contains occurrences of 14,947 actinopterygian species among 3,119 freshwater or estuarine drainage basins globally. I removed occurrences representing invasive species and occurrences marked as uncertain. I matched tip names in the Rabosky et al. (2018) phylogeny to these occurrences, using the validate_names function in rfishbase version 2.99 (Boettiger et al. 2012) to resolve taxonomic incongruences. Of 11,587 species in the phylogeny, 6,717 species were found in 1 or more drainage basin. I coded each freshwater species as either riverine, lacustrine, or occurring in both habitat types using descriptions of its geographic range and habitat in FishBase (Froese and Pauley 2019) and IUCN Red List version 6.2 (2020). Some fishes were endemic to other freshwater habitats including floodplains, swamps, ponds, temporary pools, caves, and springs. Rather than exclude these species or assign new categories, I coded them as riverine. I did this because most of these habitats are created or maintained by fluvial action, and because I expect speciation modes in these habitats to be similar to rivers (i.e. allopatry as the dominant mode). Some of these 6,717 species spend most of their time in estuarine or marine habitats rather than strictly freshwater habitats. I coded these 1,504 species as a fifth category (brackish or euryhaline) rather than assign them to freshwater or marine habitats. I used FishBase, IUCN, and a recent compilation (Corush 2019) to identify the habitat of the remaining 4,870 species that were not recorded in freshwater drainage basins. Of these, 4,139 species were strictly marine, explaining their absence in this database (Tedesco et al. 2017b). I also identified 244 diadromous species, which were excluded from all analyses as these species have a distinct

lifestyle from other freshwater or marine fishes (Corush 2019). Finally, I excluded 311 species without enough information in these sources to confidently assign a habitat category.

Speciation rates among habitats:

I compared speciation rates among habitat types using tip-associated rates (Title and Rabosky 2019) calculated for each species by Rabosky et al. (2018). Speciation rates associated with each tip were downloaded using the R package fishtree version 0.2.0 (Chang et al. 2019). I used three alternative rate types: (1) the DR statistic (Jetz et al. 2012) calculated from the phylogeny of ray-finned fishes with all missing species imputed on the phylogeny; (2) speciation calculated by Bayesian Analysis of Macroevolutionary Mixtures (BAMM; Rabosky 2014) under a time-varying model on the phylogeny with genetic data only; and (3) speciation calculated by BAMM under a time-constant model. Analysis settings used to obtain these rates are detailed in the original study (Rabosky et al. 2018). The DR statistic for any given species is estimated from the number of branches (i.e. splitting events) between the species and the root of the tree and the length of those branches (Jetz et al. 2012). BAMM uses a model of diversification to simulate a posterior distribution of rate shift configurations, in which taxa in the same rate shift cohort have identical tip rates. While -SSE models can also be used to obtain tip rates (Cataeno et al. 2018), I chose to use rates calculated using methods agnostic to traits (DR, BAMM). This is because it is simple to compare rates among several categories at once using these trait-agnostic rates, while -SSE models are presently limited in the number of traits they can model.

I compared tip-associated speciation rates between species in four habitat categories: riverine, both riverine and lacustrine, lacustrine, and marine. Note that tip rates were not available for 142 species (Chang et al. 2019). I excluded euryhaline species from these analyses

because the habitat where speciation occurred may be unclear when species interact with both freshwater and marine habitats. I statistically compared the mean tip rates of the four habitats using phylogenetic ANOVA (Garland et al. 1993) implemented in phytools version 0.6–44 (Revell 2012). I repeated the phylogenetic ANOVA using each of the three different measures of rate (DR, BAMM time-constant, and BAMM time-variable). I also performed a set of analyses comparing only freshwater habitats (three categories: riverine, both riverine and lacustrine, and lacustrine). The phylANOVA function in phytools implements a simulation-based post-hoc test to compare pairwise differences in means, and also uses a Holm-Bonferroni correction for multiple comparisons.

In addition to analyses spanning all actinopterygians, I also compared tip rates among the three freshwater habitats within (1) individual clades and (2) individual biogeographic regions. I did this to examine whether patterns seen across fishes were driven by a single clade or region (e.g. "phylogenetic pseudoreplication"; Maddison and FitzJohn 2014). I identified diverse orders and families whose members were found in all three habitat types. These were the Cypriniformes (n=1,525 species with habitat data), Siluriformes (n=1,005), Cichlidae (n=720), Cyprinodontiformes (n=470), Atheriniformes (n=136), and Cottidae (n=60). I used occurrence records from Tedesco et al. (2017b) to identify species found in six biogeographic regions. These regions were the Neotropics (n=1,665 species), Afrotropics (n=1,109), Indo-Malay (n=1,052), Palearctic (n=794), Nearctic (n=738), and Australasia (n=193).

Mode of speciation:

I tested whether (1) habitats differed in mode of speciation, and (2) whether mode of speciation was related to speciation rates. I first identified freshwater sister-species pairs using

the phylogeny of Rabosky et al. (2018). I only included sister-species pairs from genera that were 80-100% sampled with genetic data. This is because if the true sister is not sampled, the apparent sister-species relationship will appear older, and the relationship may not reflect the conditions during speciation (McCune and Lovejoy 1998).

I obtained fine-scale range information for sister species from the IUCN Red List. Range information for freshwater fishes in IUCN is based on the HydroBASINS delineation of watershed boundaries (Lehner and Grill 2013). Data were downloaded from IUCN on 27 February 2020. Introduced and uncertain occurrences were excluded. I calculated the percent of range overlap for sister-species pairs following Barraclough and Vogler (2000). This was the number of watersheds shared by the two sister species divided by the total number of watersheds where the smaller-ranged sister was found. Each pair was assigned one of three categories: allopatric, partially sympatric, and sympatric. I categorized each sister species pair as allopatric if no drainage basins were shared (percent overlap=0), sympatric if the smaller-ranged species was completely nested in the range of the larger-ranged species (percent overlap=1), and partially sympatric if the percent of overlap was greater than 0 but less than 1. I acknowledge that present-day range overlap may not reflect overlap at time of speciation if species' ranges have shifted (Bolnick and Fitzpatrick 2007). Overlapping sister-species pairs may cautiously be considered as potential cases of sympatric speciation (Coyne and Orr 2004).

In addition to range overlap, I also categorized species pairs by the habitat where speciation presumably took place. If both species in the pair were found in rivers, I coded the pair as riverine. Similarly, if both species were found in lakes, I coded the pair as lacustrine. I coded the pair as "both lacustrine and riverine" if either species was found in both lakes and rivers, or if one species was found in rivers and the other in lakes.

As before, I used phylogenetic ANOVA (Garland et al. 1993) to test if speciation rates differed in association with range overlap categories (allopatry, partial sympatry, or sympatry), including post-hoc tests. In the phylogeny used for these analyses, one tip represented each sister species pair. The species in the pair that was dropped was arbitrary, since this choice does not affect the resulting phylogeny. In all sister-species pairs, the tip-associated speciation rate for each species in the pair was identical, so this rate represented the speciation rate of the pair. I performed alternative analyses using three measures of speciation: DR, BAMM under a time-varying model, and BAMM under a time-constant model. Note that other options for calculating diversification rate differences associated with traits or geography, such as -SSE models (Cataeno et al. 2018) were not appropriate for these analyses. This is because the trait of interest here is range overlap, which must be a property of two or more species.

I also compared speciation rates among habitat and range overlap categories jointly using phylogenetic ANOVA. Again, I used three alternative measures of speciation rates. Nine categories were possible based on range overlap (allopatry, partial sympatry, or sympatry) and habitat of speciation (riverine, lacustrine, or both rivers and lakes). After limiting my sample to relatively complete portions of the phylogeny, three categories had few or no sister-species pairs: allopatric+lacustrine, partially sympatric+lacustrine, and sympatric+mixed habitats (Fig. 3). These categories were excluded to avoid complications with posthoc tests, leaving six categories for comparison (allopatric+riverine, allopatric+mixed habitats, partially sympatric+riverine, partially sympatric+mixed habitats, sympatric+riverine, and sympatric+lacustrine).

Models of habitat evolution:

To understand why fewer species occur in lakes than rivers, I fit alternative models of habitat evolution across actinopterygians using the fitMk function in phytools. These analyses included five habitat types: marine, brackish/euryhaline, riverine, both riverine and lacustrine, and lacustrine. My goal was to test two potential hypotheses explaining the dearth of lacustrine species. First, I hypothesized that when species transition from marine to freshwater habitats, they first enter estuaries or rivers before entering lakes. To test this hypothesis, I compared the fit of models with direct transitions between marine and lacustrine habitats disabled to models with these transitions enabled. Second, I hypothesized that species leave lakes at a faster rate than they enter lakes. To test this hypothesis, I compared the fit of models with symmetrical transition rates (identical forward-and-back rates) between lakes and other habitats to models with asymmetrical rates. I acknowledge that the lack of marine fossils may cause deep nodes in the fish phylogeny to be erroneously reconstructed as freshwater (Betancur-R et al. 2015). However, this bias should not be a concern here as it primarily affects the deepest nodes among fishes (before 200 million years ago; Betancur-R et al. 2015) while lacustrine radiations are much more recent (e.g. Danley et al. 2012).

In total, I compared the fit of eight alternative models (model details in Table S17). I compared the fit of models using AIC scores. To estimate the amount of time on the phylogeny spent in each habitat, I performed stochastic character mapping of habitat using the "make.simmap" function in phytools. I used the Q-matrix estimated by the best-fitting model to perform 100 simulations of habitat evolution.

RESULTS

Speciation rate differences and habitat

Of 11,032 species used for analyses, 4,000 species were coded as riverine (36.3%), 867 as both lacustrine and riverine (7.9%), 512 as lacustrine (4.6%), 1,514 as brackish or euryhaline (13.7%), and 4,139 as marine (37.5%). Across actinopterygians, lacustrine species had faster speciation rates than riverine, mixed habitat, or marine species (Fig. 1). Phylogenetic ANOVAs were not significant overall when comparing four habitat categories at once; however, pairwise post-hoc tests showed that lacustrine species had significantly different speciation rates than all other habitats (Table S1). When only comparing the three freshwater habitats, phylogenetic ANOVAs were significant overall ($p=0.002$ using DR; Table S2). Pairwise comparisons among the other habitat categories were never significant. Surprisingly, riverine and marine habitats had similar speciation rates (Fig. 1). These results were robust to the measure of speciation rate used (Tables S1–S2). Detailed results for all phylogenetic ANOVAs are found in the Supporting Information.

In general, lacustrine species also tended to have faster speciation rates than other habitats within individual clades and biogeographic regions (Fig. 2). Among clades (Tables S3–S8), there was some uncertainty in whether differences in means were significant depending on each of the three rate types used. In Cypriniformes (the largest group examined) all three phylogenetic ANOVAs were significant, and lakes had significantly faster speciation rates than all other habitat types. In Atheriniformes and Cyprinodontiformes, results were significant using BAMM but not DR, including pairwise differences in rates. In Siluriformes, no phylogenetic ANOVAs were significant overall, yet post-hoc differences between lakes and other habitats were significant using DR. In Cottidae and Cichlidae, no comparisons were significant despite plots showing fast speciation rates in lacustrine species (Fig. 2). There was similar uncertainty in whether differences in mean speciation rates were significant for individual biogeographic

regions (Fig. 2; Tables S9–S14). In the Nearctic region, all phylogenetic ANOVAs were significant, with lakes having faster rates than all other habitats. For the Palearctic and Australasia regions, phylogenetic ANOVAs were significant overall using DR but not BAMM, with lakes again having the fastest rates. For the Indo-Malay region, results were globally significant using BAMM, but pairwise comparisons between lakes and other habitats were only significant using a time-constant model. For the Neotropics, no ANOVAs were globally significant, but lakes had significantly faster rates in pairwise comparisons using all rate types. No comparisons were significant in the Afrotropics, despite apparent differences when the distributions of rates were plotted (Fig. 2). Overall, pairwise comparisons of rates were only ever significant between lacustrine endemics and other habitats, and were never significant among other habitat types in any clade or region (Tables S3–S14).

Habitat and range overlap

In total, 268 freshwater sister-species pairs were appropriate for analyses (i.e. species pairs come from well-sampled portions of the phylogeny). This sample included 166 riverine, 80 mixed-habitat, and 22 lacustrine pairs, and pairs from 35 families with the most sampled families being Cyprinidae (64 pairs), Percidae (58 pairs), and Cichlidae (38 pairs). Allopatry was the most common mode in rivers (Fig. 3). Of 166 riverine sister-species pairs, 83 were allopatric (50%), 61 were partially sympatric (36.7%) and 22 were sympatric (13.3%). Sympatry was the most common speciation mode in lakes. Of 22 lacustrine sister-species pairs, 21 were sympatric, 0 were partially sympatric, and only 1 was allopatric. There were 80 sister-species pairs found in both rivers and lakes, 31 of which were allopatric (38.8%), 44 were partially sympatric (55%), and 5 were completely sympatric (6.3%).

Sympatric sister-species pairs had faster speciation rates than allopatric or partially sympatric pairs (Fig. 3), though this difference was not significant using phylogenetic ANOVA (p=0.053–0.088, Table S15). Sister-species pairs that were both sympatric and lacustrine had faster speciation rates than other habitat-range overlap categories. This difference was significant using speciation rates calculated by BAMM (p=0.028–0.037, Table S16) but not DR (p=0.054). Post-hoc tests did not show any significant differences in pairwise comparisons of rates. There was not a strong difference in speciation rates associated with range overlap within riverine or mixed-habitat pairs (Fig. 3). Note that all 21 pairs that were both sympatric and lacustrine were cichlids.

Model of habitat evolution

The model of habitat evolution with the best support did not allow direct transitions between marine and lacustrine habitats, with all other transition rates differing (Table S17). Transition rates estimated by this model suggest that lineages enter freshwater habitats from the ocean by first becoming euryhaline (Fig. 4). Direct transitions from rivers into lakes were rare (0.001). Transitions from riverine to mixed habitats (both rivers and lakes) were more common (0.071), as were transitions from mixed habitats to lakes (0.011). The transition rate from lakes to mixed habitats was higher than the reverse (0.041), as was the rate from mixed habitats back to rivers (0.359). Direct transitions from lakes back to rivers were also relatively common (0.021). Altogether, these patterns suggest that lineages leave lakes at a faster rate than they enter lakes.

Overall, only 1.9% of the phylogeny was reconstructed as endemic to lacustrine habitats (mean across 100 simulations; Table S17). In comparison, 32.3% was reconstructed in riverine

habitats, 6.5% in mixed habitats, 17.0% in euryhaline/brackish habitats, and 42.2% in marine habitats.

DISCUSSION

In this study, I compared rates of speciation across ray-finned fishes in riverine, lacustrine, and marine habitats. I also compared modes of speciation associated with habitats, as well as fit alternative models of habitat evolution. I found three main results. First, rates of speciation in lakes are faster than riverine or marine habitats, while rates are similar between riverine and marine habitats. Second, sympatric speciation is mostly associated with lakes and allopatric speciation is mostly associated with rivers, but speciation rates may be driven by habitat more than range overlap *per se*. Third, fishes have spent comparatively little time in lakes over their phylogenetic history, at least in part due to biases in transition rates among habitats. I interpret these results and discuss their significance below.

Past studies have found faster rates of speciation in freshwater and marine lineages of fishes (Bloom et al. 2013; Tedesco et al. 2017a, Miller et al. 2018). This is expected if finer habitat fragmentation leads to more opportunities for speciation in freshwater than marine habitats (Horn 1972; May 1994; Grosberg et al. 2012). However, the results shown here support a counter-intuitive, more nuanced explanation for how diversity accumulated in each biome. Consistent with expectations, I found that most freshwater diversity did evolve in rivers, likely through allopatric speciation (Figs. 3, 4). However, contrary to expectations, speciation rates in rivers are modest on average and not generally faster than in the ocean (Fig. 1). Speciation within lakes, possibly via sympatric speciation, occurs much faster than speciation in rivers but accounts for far less freshwater diversity overall. Collectively, these factors suggest that the vast

majority of diversity of freshwater fishes is the result of steady accumulation of allopatric species over long periods of time, rather than a substantially higher overall speciation rate in freshwater compared to marine waters. This points to a combination of stability (e.g. protection from extinction) and time-for-speciation for explaining why freshwater habitats have high diversity relative to the area of water available, which has also been suggested by past studies (Carrete Vega and Wiens 2012, Seehausen and Wagner 2014, Betancur-R et al. 2015; see also Miller and Wiens 2017).

How do we reconcile the observation that speciation rates do not differ between riverine and marine habitats with evidence to the contrary from other studies? For example, Manel et al. (2020) found that freshwater fishes have higher intraspecific genetic diversity than marine fishes. However, genetic diversity was only weakly related to local species richness. Furthermore, population subdivision may not have a straightforward relationship to speciation rates (Harvey et al. 2019). The faster rates of speciation in freshwater habitats inferred by some past studies (Bloom et al. 2013, Miller et al. 2018), may be a consequence of including both lacustrine and riverine habitats as a single “freshwater” category, and not reflective of a general difference between marine and freshwater biomes. Indeed, Rabosky (2020) recently found that apparently faster average freshwater rates of speciation were driven by cichlids and ostariophysans alone, and were not broadly characteristic of marine-freshwater transitions in fishes. Finally, it is worth noting that some marine lineages are known for fast speciation rates (Rüber and Zardoya 2005; Miller and Wiens 2017; Rabosky et al. 2018), despite the apparent lack of physical barriers in the ocean. The mechanisms for speciation in the marine environment may be as varied as those found among freshwater habitats (Bowen et al. 2013), and thus a general expectation for low speciation rates in the ocean may be overly simplistic.

It is difficult to disentangle differences in speciation rates inherent to allopatry/sympatry from differences inherent to rivers/lakes. McCune and Lovejoy (1998) suggested that the reason for fast speciation rates in lakes was that sympatry was the dominant speciation mode. However, all the sympatric sister-species pairs in their comparisons were lacustrine. The comparisons in this study also included sympatric sister-species pairs that are riverine (Fig. 3). From these comparisons, it seems that range overlap is not associated with faster speciation in rivers (Fig. 3). That is, faster speciation in sympatry may just be a lacustrine phenomenon. It is also likely that some sister-species in rivers did arise through allopatry but have since come into secondary sympatry. Thus, the rate of speciation in sympatric riverine pairs may still reflect that typical of allopatric speciation.

I only sampled 1 allopatric lacustrine species-pair (where sister species were found in separate lakes), so I cannot say whether allopatric lacustrine fishes diversify slower than sympatric lacustrine fishes. One concern is that apparent sympatric speciation in lakes may actually be “micro-allopatric” speciation (Seehausen and van Alphen 1999). To distinguish between sympatric and allopatric speciation within a single lake, finer-scale ecological and spatial data for sister species are needed. Still, past studies (McCune and Lovejoy 1998; Seehausen and van Alphen 1999) suggested that allopatric speciation cannot wholly account for the diversity of African cichlids in Lakes Malawi and Victoria, because under the average allopatric speciation rate it would take more time to generate the hundreds of endemic cichlid species than these lakes have existed. This may be true for other lacustrine species flocks as well, especially since lakes tend to be geologically young (see below). The trend for sympatry to be more common in lakes than rivers suggests at least a correlation between mode and rate of speciation, even if allopatric speciation is harder to detect within lakes.

If rates of speciation are so fast in lakes, then why do lakes contain a small proportion of freshwater diversity? Models of habitat evolution (Fig. 4; Tables S17, S18) suggest at least two reasons. First, transitions into lakes are restricted. Direct transitions from the ocean into lakes were not allowed in the best-fitting model (Table S17); this is intuitive, as lakes are fed by rivers and do not interface directly with the ocean. Inferred transition rates suggest a likely path to entering lakes is by first entering rivers, then both rivers and lakes, then finally becoming endemic to lakes (Fig. 4). Since there are several evolutionary steps needed to become endemic to lakes, lacustrine radiations are expected to be nested in the phylogeny relative to lineages in other habitats (Fig. 4). Since these radiations may have less time-for-speciation (Stephens and Wiens 2003) it is not surprising that they contribute fewer species than rivers overall. Second, transition rates out of lakes are higher than those into lakes (Fig. 4). This suggests that being endemic to lakes is an evolutionarily unstable strategy. These results can be framed in light of the geological age of rivers and lakes. On a local scale, conditions in rivers change frequently and seasonally, while conditions in lakes are stable. However, lakes are younger than rivers over deep timescales, because lakes dry out and disappear entirely (McCune 1987; Lowe-McConnell 1969; Cristescu et al. 2010). Rivers also go through dramatic changes (Shepard et al. 2010) but may still exist in different forms, allowing the long-term persistence of lineages. Indeed, the fossil record documents past lacustrine radiations that went extinct and were replaced as lakes dried out and refilled (McCune 1987). Lacustrine radiations may avoid extinction by entering rivers (Seehausen and Wagner 2014), explaining the high transition rates out of lakes.

It is somewhat unexpected that rivers have lower speciation rates than lakes, since the fragmentation of rivers has been shown to promote population subdivision and endemism (Dias et al. 2013). However, earlier literature posited that lakes should have faster speciation rates for

ecological and structural reasons. Briggs (1966) and Lowe-McConnell (1969) thought that rates of speciation should be faster in peripheral habitats (giving lakes as an example), even if they are evolutionarily unstable systems. Speciation in rivers may be limited if there are fewer available niches compared to lakes, while lineages entering lakes may experience ecological release and undergo niche evolution (Seehausen 2015). Since rivers are locally unstable, riverine fishes are more likely to be opportunistic feeders. In turn, the stable conditions of lakes allow for trophic specialization. In addition, since lakes tend to be deeper than rivers, they offer three dimensions to partition habitats: onshore-offshore, shallow-deep, and along the shoreline (Seehausen and Wagner 2014; Seehausen 2015). This may promote species coexistence and allow for faster speciation rates.

The most famous examples of lacustrine radiation are cichlids in the African rift lakes (Wagner et al. 2012; Seehausen 2015). However, rapid speciation in lakes is not limited to cichlids nor African lakes (Fig. 2). Each clade and biogeographic region examined showed a tendency for faster speciation in lakes compared to rivers. In several cases, this tendency was not statistically significant using phylogenetic ANOVA (Tables S3–S16). In spite of the inconsistency of statistical significance, faster speciation in lakes may indeed be a widespread phenomenon. First, the degrees of freedom in statistical tests will be reduced in phylogenetic tests if the trait of interest is phylogenetically clustered and unreplicated (Felsenstein 1985; Maddison and Fitzjohn 2014). Both conditions are expected for lacustrine radiations because transitions into lakes are rare (Fig. 4) and by definition species flocks are closely related to each other. Second, among six clades and six biogeographic regions there were no cases where pairwise comparisons among other habitat types were significant; all significant pairwise comparisons in this study were between lacustrine endemics and the other habitats (Supporting

Information). Third, comparisons were significant when all actinopterygian fishes were included in the same analysis (Tables S1–S2). This suggests that the addition of all clades allowed for phylogenetic replication and increased power.

In conclusion, lacustrine fishes have faster speciation rates than other aquatic habitats. However, most overall freshwater diversity arose in rivers. Surprisingly, riverine and marine fishes had similar rates of diversification, despite past literature predicting opposite effects on speciation rates in these habitats (Horn 1972; May 1994; Grosberg et al. 2012). While allopatric speciation is indeed important for generating freshwater diversity, this mechanism may not be the cause of faster speciation rates between freshwater and marine habitats. Instead, these results point to alternative explanations for the high diversity of freshwater habitats relative to available water volume, including long-term diversification and lineage persistence.

ACKNOWLEDGEMENTS

I thank John Wiens and Luke Tornabene for helpful discussion related to this manuscript. I was supported by an NSF Postdoctoral Fellowship (DBI-1906574).

REFERENCES

Barraclough, T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. Am. Nat. 155:419–434.

Betancur-R, R., G. Orti, and R. A. Pyron. 2015. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. Ecol. Lett. 18:441–450.

Bloom, D. D., J. T. Weir, K. R. Piller, and N. R. Lovejoy. 2013. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of new world silversides (Atherinopsidae). Evolution 67:2040–2057.

Boettiger, C., D. T. Lang, and P. C. Wainwright. 2012. rfishbase: exploring manipulating and visualizing FishBase data from R. J. Fish. Biol. 81, 2030–2039.

Bolnick, D. I., and B. M. Fitzpatrick. 2007. Sympatric speciation: models and empirical evidence. Annu. Rev. Ecol. Evol. S. 38:459–487.

Bowen, B. W., L. A. Rocha, R. J. Toonen, S. A. Karl, et al. 2013. The origins of tropical marine biodiversity. Trends Ecol. Evol. 28:359–366.

Briggs, J. C. 1966. Zoogeography and evolution. Evolution 20:282–289.

Bush, G. L. 1975. Modes of animal speciation. Annu. Rev. Ecol. Syst. 6:339–64

Caetano, D. S., B. C. O'Meara, and J. M. Beaulieu. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. Evolution 72:2308–2324.

Carrete Vega, G., and J. J. Wiens. 2012. Why are there so few fish in the sea? Proc. R. Soc. Lond. B. 279:2323–2329.

Chang, J., D. J. 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 Ecol. Evol. 10:1118–1124.

Cohen, A. S. 1995. Paleoecological approaches to the conservation biology of benthos in ancient lakes: a case study from Lake Tanganyika. J. N. Am. Benthol. Soc. 14:654–668.

Corush, J. B. 2019. Evolutionary patterns of diadromy in fishes: more than a transitional state between marine and freshwater. BMC Evol. Biol. 19:168.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer, Sunderland, MA.

Cristescu, M. E., S. J. Adamowicz, J. J. Vaillant, and D. G. Haffner. 2010. Ancient lakes revisited: from the ecology to the genetics of speciation. Mol. Ecol. 19:4837–4851.

Crow, K. D., H. Munehara, and G. Bernardi. 2010. Sympatric speciation in a genus of marine reef fishes. Mol. Ecol. 19:2089–2105.

Danley, P. D., M. Husemann, B. Ding, L. M. DiPietro, E. J. Beverly, and D. J. Peppe. 2012. The impact of the geologic history and paleoclimate on the diversification of East African cichlids. Int. J. Evol. Biol. 2012:574851.

Dias, M. S., J. Cornu, T. Oberdorff, C. A. Lasso, and P.A. Tedesco. 2013. Natural fragmentation in river networks as a driver of speciation for freshwater fishes. Ecography 36:683–689.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Froese, R., and D. Pauly. 2019. FishBase Version 12/2019. Available at http://www.fishbase.org/.

Garland Jr., T., A. W. Dickerman, C. M. Janis, and J. A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. Syst. Biol. 42:265–292.

Grosberg, R. K., G. J. Vermeij, and P. C. Wainwright. 2012. Biodiversity in water and on land. Curr. Biol. 22:R900–R903.

Harvey, M. G., S. Singhal, and D. L. Rabosky. 2019. Beyond reproductive isolation: demographic controls on the speciation process. Annu. Rev. Ecol. Evol. S. 50:75–95.

Horn, M. H. 1972. The amount of space available for marine and freshwater fishes. Fish. Bull. U.S. 70:1295–1297.

Hrbek, T., F. Küçük, T. Frickey, K. N. Stölting, R.H. Wildekamp, and A. Meyer. 2002. Molecular phylogeny and historical biogeography of the *Aphanius* (Pisces, Cyprinodontiformes) species complex of central Anatolia, Turkey. Mol. Phylogenet. Evol. 25:125–137.

IUCN 2020. The IUCN Red List of Threatened Species. Version 6.2. Available at https://www.iucnredlist.org.

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.

Lehner, B., and G. Grill. 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process 27:2171–2186

Lévêque, C., T. Oberdorff, D. Paugy, M. L. J. Stiassny, and P. A. Tedesco. 2008. Global diversity of fish (Pisces) in freshwater. Hydrobiologia 595:545–567.

Lowe-McConnell, R. H. 1969. Speciation in tropical freshwater fishes. Biol. J. Linn. Soc. 1:51–75.

Lynch, J. D. 1989. The gauge of speciation: on the frequencies of modes of speciation. Pp. 527–553 in D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer Associates, Sunderland, MA.

Maddison, W. P., and R. G. FitzJohn. 2014. The unsolved challenge to phylogenetic correlation tests for categorical characters. Syst. Biol. 64:127–136.

Manel, S., P.-E. Guerin, D. Mouillot, S. Blanchet, L. Velez, C. Albouy, and L. Pellissier. 2020. Global determinants of freshwater and marine fish genetic diversity. Nat. Commun. 11:692.

May, R. M. 1994. Biological diversity: differences between land and sea. Phil. Trans. R. Soc. Lond. B 343:105–111.

McCune, A. R. 1987. Lakes as laboratories of evolution endemic fishes and environmental cyclicity. Palaios 2:446–54

McCune, A. R., and N. R. Lovejoy. 1998. The relative rate of sympatric and allopatric speciation in fishes. Pp. 172–185 in D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford Univ. Pres, New York, NY.

Miller, E. C., and J. J. Wiens. 2017. Extinction and time help drive the marine-terrestrial biodiversity gradient: is the ocean a deathtrap? Ecol. Lett. 20:911–921.

Miller, E. C., K. T. Hayashi, D. Song, and J. J. Wiens. 2018. Explaining the ocean's richest biodiversity hotspot and global patterns of fish diversity. Proc. Roy. Soc. B. 285:20181314.

Near, T. J., and M. F. Benard. 2004. Rapid allopatric speciation in logperch darters (Percidae: *Percina*). Evolution 58:2798–808.

Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS ONE 9:e89543.

Rabosky, D. L. 2020. Speciation rate and the diversity of fishes in freshwaters and the oceans. J. Biogeogr. in press (doi: 10.1111/jbi.13839).

Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4:1958.

Rabosky, D. L., J. Chang, P. O. Title, P. F. Cowman, L. Sallan, M. Friedman, K. Kaschner, C. Garilao, T. J. Near, M. Coll, et al. 2018. An inverse latitudinal gradient in speciation rate for marine fishes. Nature 559:392– 395.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

Rüber, L., and R. Zardoya. 2005. Rapid cladogenesis in marine fishes revisited. Evolution 59:1119–1127.

Seehausen, O., and J. J. M. van Alphen. 1999. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? Ecol. Lett., 2, 262–271.

Seehausen, O. 2015. Process and pattern in cichlid radiations- inferences for understanding unusually high rates of evolutionary diversification. New Phytol. 207:304–312.

Seehausen, O., and Wagner, C. E. (2014). Speciation in freshwater fishes. Annu. Rev. Ecol. Evol. S. 45:621–651.

Shephard, G. E., R. D. Müller, L. Liu, and M. Gurnis. 2010. Miocene drainage reversal of the Amazon River driven by plate-mantle interaction. Nat. Geosci. 3, 870–875.

Stephens, P.R., and J. J. Wiens. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. Am. Nat. 161:112–128.

Sullivan, J. P., S. Lavoué, and C. D. Hopkins. 2002. Discovery and phylogenetic analysis of a riverine species flock of African electric fishes (Mormyridae: Teleostei). Evolution 56:597–616.

Tedesco, P. A., E. Paradis, C. Lévêque, and B. Hugueny. 2017a. Explaining global-scale diversification patterns in actinopterygian fishes. J. Biogeogr. 44:773–783.

Tedesco, P. A., O. Beauchard, R. Bigorne, S. Blanchet, L. Buisson, L. Conti, J. Cornu, M. S. Dias, G. Grenouillet, B. Hugueny, et al. 2017b. A global database on freshwater fish species occurrence in drainage basins. Sci. Data 4:170141.

Title, P. O., and D. L. Rabosky. 2019. Tip rates, phylogenies, and diversification: What are we estimating, and how good are the estimates? Methods Ecol. Evol. 10:821–834

Wagner, C. E., L. J. Harmon, and O. Seehausen. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. Nature 487:366–369.

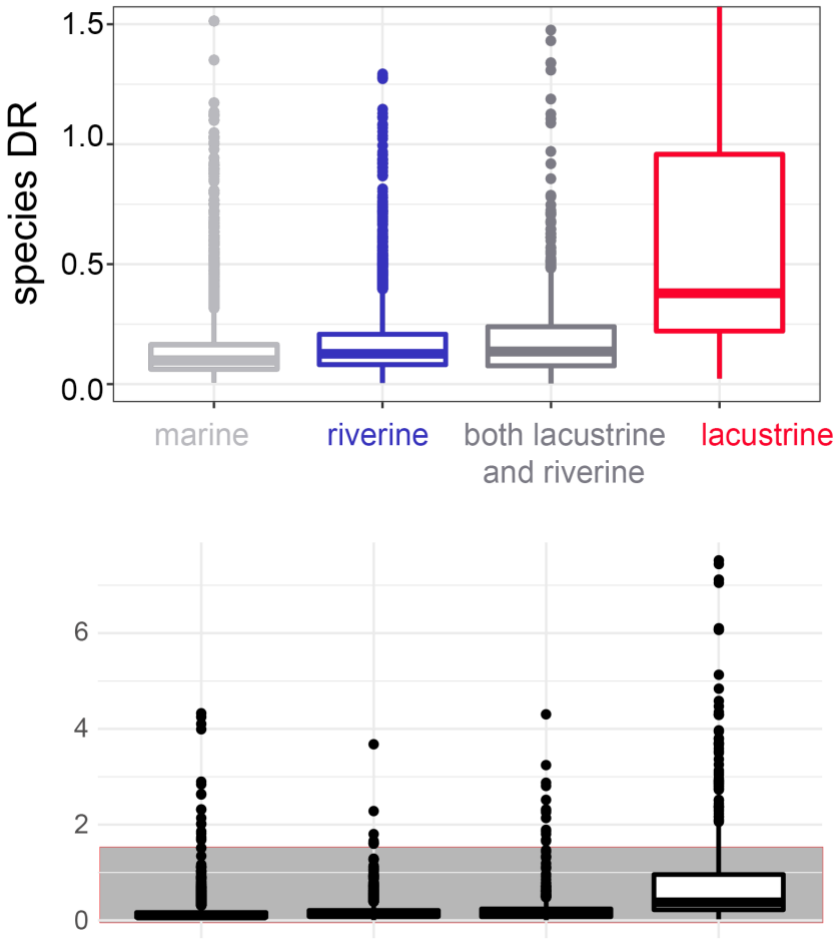

Figure 1. Tip-associated speciation rates across ray-finned fishes (n=9,376 species) in relation to habitat. Top panel shows close-up of medians (grey inset of bottom panel); bottom panel shows all points including outliers. Results were similar using speciation rates calculated from other methods (BAMM time-constant or time-variable; Tables S1, S2).

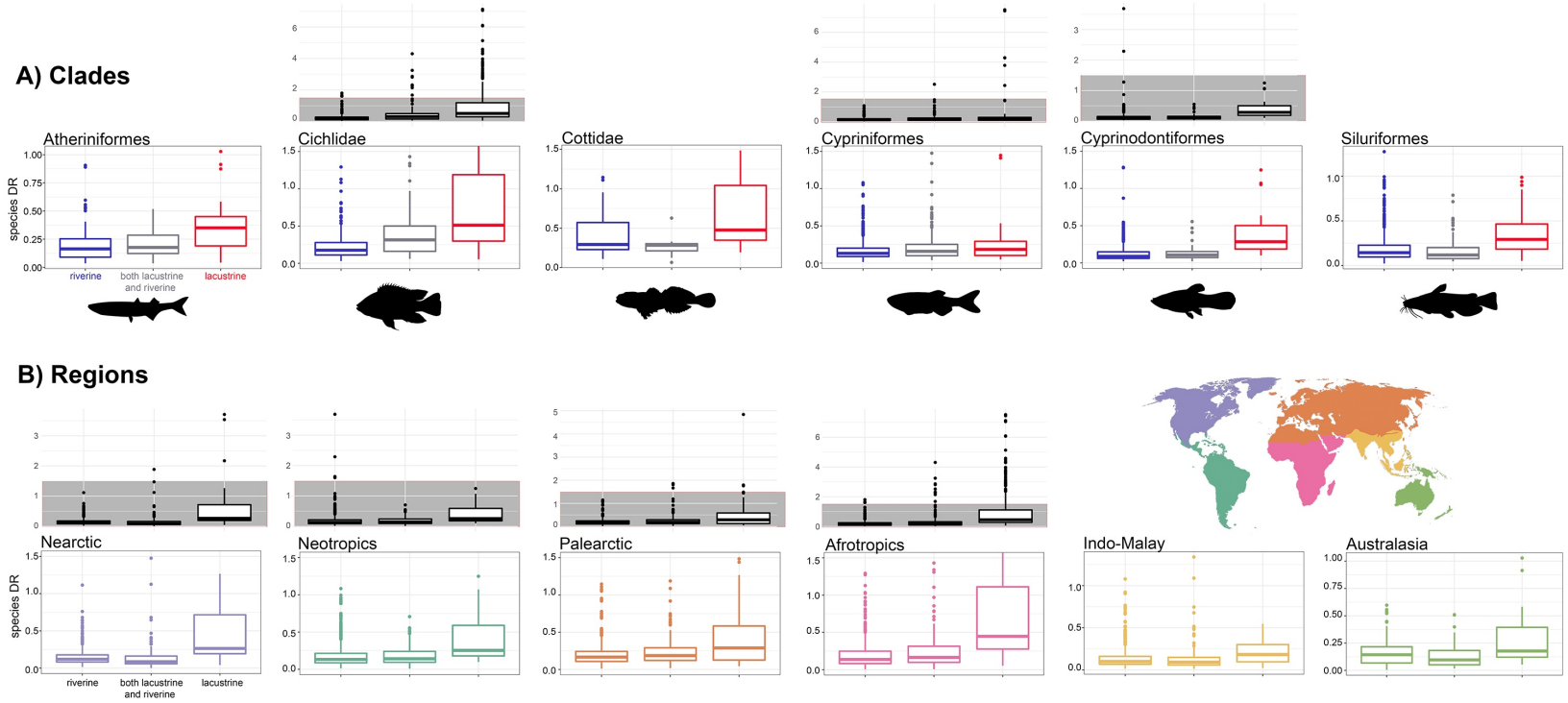
A) Clades
Atheriniformes
Cichlidae
Cottidae
Cypriniformes
Cyprinodontiformes
Siluriformes
species DR
riverine
both lacustrine and riverine
lacustrine
B) Regions
Nearctic
Neotropics
Palearctic
Afrotropics
Indo-Malay
Australasia
species DR
riverine
both lacustrine and riverine
lacustrine

Figure 2. Tip-associated speciation rates related to freshwater habitat among individual clades (A) and biogeographic regions (B). In all panels, left boxplot=riverine, middle boxplot=both habitats, and right boxplot=lacustrine. If needed, top panel shows all points including outliers, and bottom panel shows close-up of medians (grey inset of top panel). Fish icons from PhyloPic (http://phylopic.org/) with credit to Anja Schach, Lily Hughes, Tyler McCraney, and Milton Tan. Higher taxonomy follows Rabosky et al. (2018). Regionalization follows Tedesco et al. (2017). Full results of phylogenetic ANOVAs are given in the Supporting Information (Tables S3–S16).

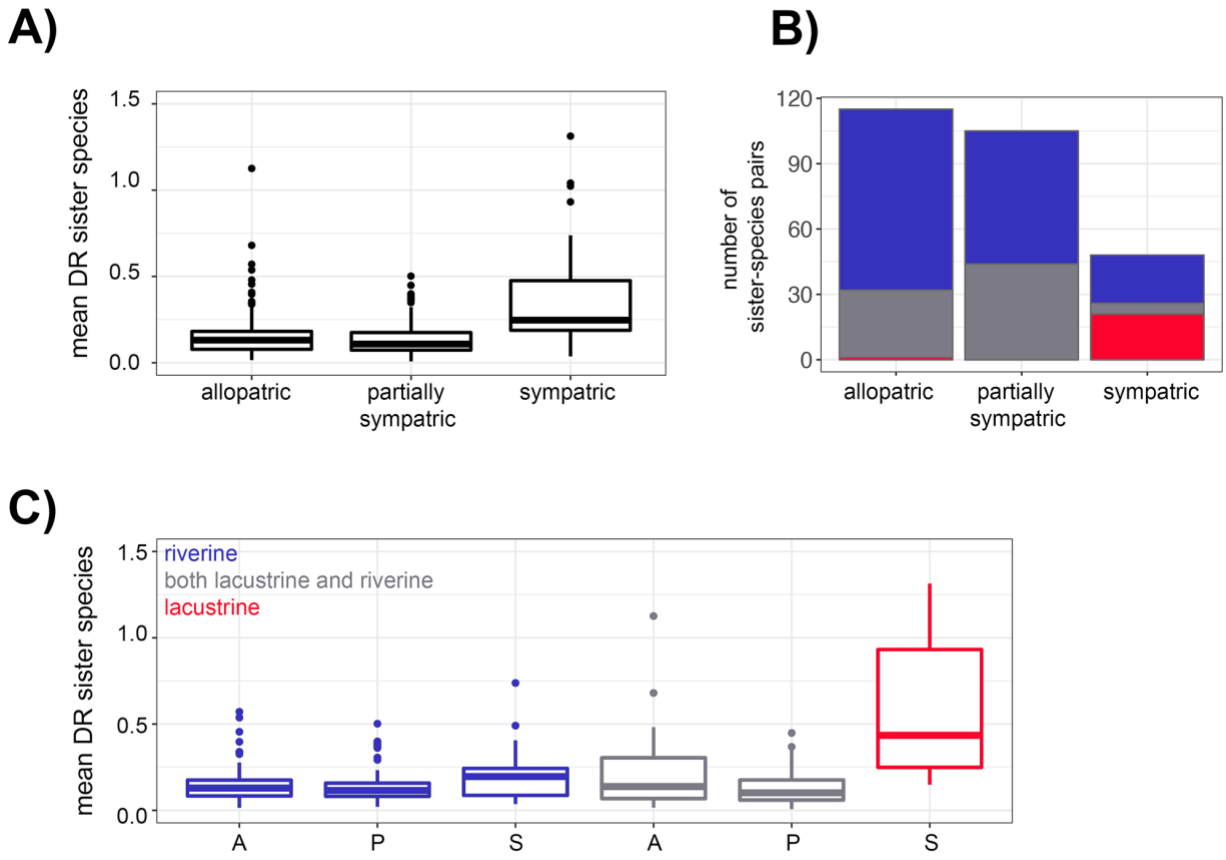

Figure 3. (A) Speciation rates of freshwater sister-species pairs (n=268 pairs) associated with present-day range overlap. Tip rates were always identical between the two sister species. (B) Counts of sister-species pairs by range overlap and habitat. Here, blue=riverine, grey=both riverine and lacustrine, and red=lacustrine. (C) Speciation rates of sister-species pairs associated with both habitat and range overlap (A=allopatric, P=partially sympatric, S=sympatric). Two outlier sympatric lacustrine pairs with DR=2.8 and 6.1 are not shown in panels A or C for clarity. Results were similar using speciation rates calculated from other methods (BAMM time-constant or time-variable; Tables S15, S16).

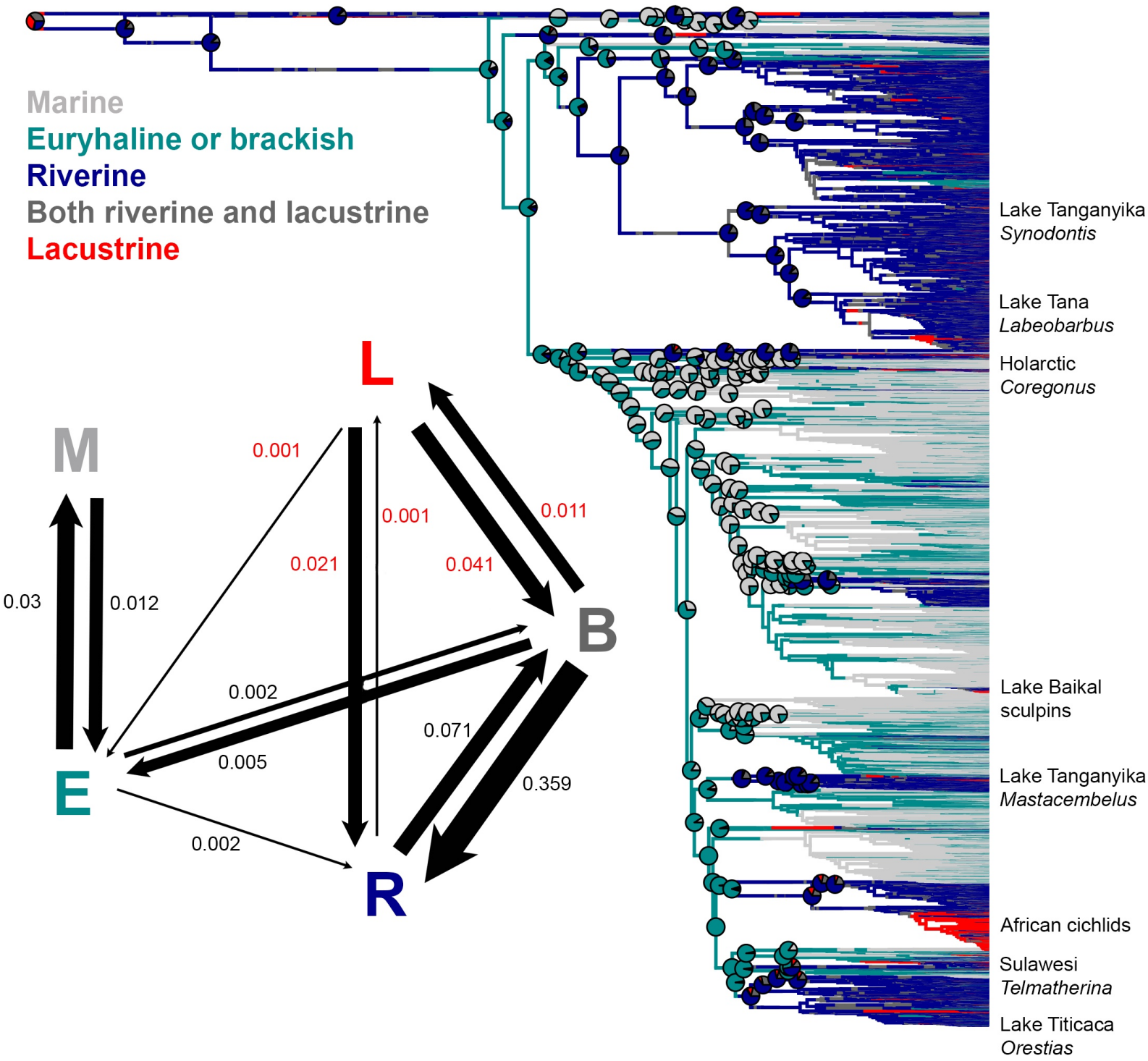
Marine
Euryhaline or brackish
Riverine
Both riverine and lacustrine
Lacustrine
M
L
B
E
R
0.03
0.012
0.001
0.001
0.021
0.011
0.041
0.002
0.005
0.071
0.359
0.002
Lake Tanganyika
Synodontis
Lake Tana
Labeobarbus
Holarctic
Coregonus
Lake Baikal
sculpins
Lake Tanganyika
Mastacembelus
African cichlids
Sulawesi
Telmatherina
Lake Titicaca
Orestias

Figure 4. Ancestral state reconstruction of habitat on the phylogeny of ray-finned fishes (Rabosky et al. 2018). Colors on branches represent the results of a single stochastic map. Pies represent the probabilities of each ancestral habitat from 100 stochastic maps. Pies are only shown on deep nodes for clarity. Transition rates were calculated from the best-fit evolutionary model (Tables S17, S18). In this model, direct transitions between marine and lacustrine habitats were disallowed, and all other transition rates were variable. Arrows are scaled proportionally to the rate magnitude. Select lacustrine radiations are enumerated on the phylogeny.