Speciation and the Latitudinal Diversity Gradient: Insights from the Global Distribution of Endemic Fish

Patrick J. Hanly,^{1,2,*} Gary G. Mittelbach,^{1,2} and Douglas W. Schemske^{1,3}

1. W. K. Kellogg Biological Station, Michigan State University, Hickory Corners, Michigan 49060; 2. Department of Integrative Biology, Michigan State University, East Lansing, Michigan 48824; 3. Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824

Submitted August 8, 2016; Accepted November 15, 2016; Electronically published April 7, 2017 Online enhancements: appendix. Dryad data: http://dx.doi.org/10.5061/dryad.70sr1.

ABSTRACT: The nearly universal pattern that species richness increases from the poles to the equator (the latitudinal diversity gradient [LDG]) has been of intense interest since its discovery by early naturalhistory explorers. Among the many hypotheses proposed to explain the LDG, latitudinal variation in (1) productivity, (2) time and area available for diversification, and (3) speciation and/or extinction rates have recently received the most attention. Because tropical regions are older and were formerly more widespread, these factors are often intertwined, hampering efforts to distinguish their relative contributions to the LDG. Here we examine the global distribution of endemic lake fishes to determine how lake age, area, and latitude each affect the probability of speciation and the extent of diversification occurring within a lake. We analyzed the distribution of endemic fishes worldwide (1,933 species and subspecies from 47 families in 2,746 lakes) and find that the probability of a lake containing an endemic species and the total number of endemics per lake increase with lake age and area and decrease with latitude. Moreover, the geographic locations of endemics in 34 of 41 families are found at lower latitudes than those of nonendemics. We propose that the greater diversification of fish at low latitudes may be driven in part by ecological opportunities promoted by tropical climates and by the coevolution of species interactions.

Keywords: latitudinal diversity gradient, speciation, endemism, species diversity.

Introduction

The rich diversity of life at tropical latitudes is remarkably consistent across habitats and taxonomic groups (Hillebrand 2004), establishing the latitudinal diversity gradient (LDG) as Earth's dominant biogeographic pattern. Although expla-

* Corresponding author; e-mail: hanlypat@msu.edu.

nations for the LDG date back to the time of Wallace and Darwin, no consensus on the drivers of elevated tropical diversity has yet emerged. Current hypotheses for the LDG focus primarily on temperate/tropical differences in productivity (energy), historical time and area, and rates of speciation or extinction (Gaston 2000; Mittelbach et al. 2007; Brown 2014; Fine 2015; Pigot et al. 2016), but disentangling these and other potential hypotheses for the LDG is challenging. The Earth has a single shared history that can limit the ability to make inferences when potential drivers are inseparable (e.g., the greater age and area of the tropics relative to extratropical regions; Mittelbach et al. 2007).

Latitudinal variation in the rates of both speciation and extinction figures prominently in many hypotheses for the LDG (Mittelbach et al. 2007; Brown 2014), and phylogenetic inference is increasingly employed to estimate these rates for different taxa and apply them to studies of the LDG (Ricklefs 2007; Morlon 2014). It is perhaps surprising then that current phylogenetic analyses have yielded little consensus on either the magnitude or direction of latitudinal differences in rates of speciation or extinction. For example, analyses of bird phylogenies (probably the best studied of all taxonomic groups) have differentially found higher speciation rates at low latitudes (Ricklefs 2006), higher recent speciation at high latitudes (Weir and Schluter 2007), and little difference in speciation rates across latitude (Rabosky et al. 2015; Pulido-Santacruz and Weir 2016). There are many challenges to estimating geographical variation in speciation and extinction rates from phylogenetic data (Morlon 2014; Rabosky and Goldberg 2015), and evolutionary biologists continue to develop new methods to address these challenges (e.g., Rabosky and Huang 2016). In this study, we take a different approach to the question of whether latitude affects speciation by examining the global distribution of endemism in freshwater fish.

Fish represent the bulk of the planet's vertebrate diversity, with nearly 29,000 described species (Froese and Pauly 2016), 41% of which are found in freshwater. Like other

ORCIDs: Hanly, http://orcid.org/0000-0001-9435-9572; Mittelbach, http://orcid.org/0000-0001-6971-9273.

Am. Nat. 2017. Vol. 189, pp. 604-615. © 2017 by The University of Chicago. 0003-0147/2017/18906-57165\$15.00. All rights reserved. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits reuse of the work with attribution. DOI: 10.1086/691535

vertebrates, freshwater fish diversity is greatest in the tropics (Tisseuil et al. 2013), and the diversity of fish increases with area in lakes (Barbour and Brown 1974) and river basins (Oberdorff et al. 1995). Owing to the restricted dispersal of lake fish, a measure of the generation of new species in situ can be obtained by identifying single-lake endemics (defined as species and subspecies confined to a single lake; hereafter simplified to "endemic"), which can persist in their natal lakes sometimes for millions of years (as in the case of deep-water sculpins in Lake Baikal; Sherbakov 1999). The global distribution of endemic fish thus provides a unique record of speciation events, with lakes containing endemic fish found from the equator to latitudes as high as 67.5°N (Lake El'gygytgyn, Siberia). Moreover, lakes with endemic fish range in age from a few thousand to millions of years old and provide natural replicates of lake ages and sizes across latitude (fig. A1; figs. A1-A5 available online). Thus, it is possible to estimate the relative importance of time, biogeography, and environment in the evolution of endemic freshwater fish, providing an opportunity to examine long-standing questions about the relationship between latitude and speciation. Our analysis of in situ speciation at the scale of individual lakes (and river basins) complements studies of endemism and speciation on islands (reviewed in Warren et al. 2015), as well as recent studies on global gradients in vertebrate diversity conducted at the scale of bioregions (Jetz and Fine 2012; Belmaker and Jetz 2015).

As Tedesco et al. (2012, pp. 977–978) note, "endemic species have always been fascinating because they should reflect the roles of speciation, extinction and dispersal ultimately responsible for their restricted distribution." The restricted distribution of endemics within particular geographic regions that are often of known history makes it possible to relate both the presence of an endemic species (i.e., evidence of a speciation event) and the number of endemic species (a measure of diversification) to potential drivers of diversity. Recent studies of the distribution and abundance of endemic species in light of factors thought to influence evolutionary rates, including area, age, isolation, and environment, have demonstrated the positive effects of area and isolation on speciation leading to endemicity in multiple taxa (e.g., Anolis lizards on Caribbean islands [Losos and Schluter 2000], Tetragnatha spiders in the Hawaiian archipelago [Gillespie and Baldwin 2010], multiple taxa on islands worldwide [Kisel and Barraclough 2010; reviewed in Warren et al. 2015], flora on islands and mountains worldwide [Steinbauer et al. 2016], angiosperms on islands worldwide [Weigelt et al. 2016]).

In fishes, Tedesco et al. (2012) found that the global richness of riverine endemic species was positively related to drainage basin area and climatic stability, Wagner et al. (2014) demonstrated strong effects of lake area and depth on the number of cichlid species arising via in situ speciation in African lakes, and Doi et al. (2012) hypothesized that lake age and endemism affect food chain length based on isotopic analysis of fish from young and ancient lakes. Although some of the above studies suggest a greater preponderance of endemic species in the tropics compared to the temperate zone (e.g., see fig. 2 in Tedesco et al. 2012), no studies to our knowledge have directly quantified how endemicity varies with latitude, age, and area. Here we develop a data set on the distribution of endemic fish in the world's largest lakes to examine this question.

We compiled data on native fish diversity (endemic and nonendemics), lake age (continuous occupancy), area (and perimeter), latitude, elevation, maximum depth, pH, and productivity (chlorophyll a, total phosphorus, Secchi depth) from 1,949 published sources for 2,746 natural lakes with a surface area \geq 50 km² that are listed in the Global Lakes and Wetlands Database (Lehner and Döll 2004). These data were analyzed to determine how age, area, and latitude together with potential physical, chemical, and biological factors contribute to (1) the probability of at least one endemic occurring in a lake and (2) the total number of endemic fish in a lake (a measure of net diversification). Although we control for the effect of age in our analysis of endemism in lake fishes, we cannot assess the effects of latitude or area on speciation rate per se because the ages of fish species are unknown. Because we cannot estimate how extinctions may have influenced contemporary patterns of endemism, endemic species richness in a lake is best viewed as an estimate of the extent of diversification (speciation minus extinction), and the presence of an endemic species in a lake is evidence of a least one speciation event. To determine whether there is a geographic bias in the occurrence of new species relative to the background distribution of species, we performed a familylevel analysis comparing the latitudinal distribution of endemic fish to that of nonendemic fish.

Methods

Database Construction

A database of endemic fish distributions was assembled for the 2,746 largest natural lakes in the world. These lakes were selected from the Global Lakes and Wetland Database Level 1 (GLWD-1; Lehner and Döll 2004), a compilation of water bodies larger than 50 km² in area that represents an unbiased sampling of lakes worldwide. After removing duplicates and man-made reservoirs that were misattributed as natural lakes, our database has 321 fewer lakes than the GLWD-1. Latitude, longitude, surface area, perimeter, and elevation data for these lakes were obtained from GLWD-1, and we added data on lake age, depth, productivity, pH, temperature, as well as native and endemic fish distributions from a review of nearly 2,000 literature sources (data are deposited in the Dryad Digital Repository: http://dx.doi.org/10.5061 /dryad.70sr1; Hanly et al. 2016).

We performed a Google Scholar query for "[Lake Name]" (including name variations, when applicable) to identify the peer-reviewed literature for each lake. Articles were read until either all target data were collected or no more information could be obtained. In cases where complete data could not be obtained through peer-reviewed literature, an additional Google Search query for "[Lake Name] filetype:pdf" was used to locate gray literature such as technical reports and government documents. Due to the scarcity of substantial literature published on most lakes, no quality screens were performed. As an additional verification measure to ensure the endemic status of each fish, a Google Scholar query for "[Fish Scientific Name]" was carried out to either corroborate or invalidate single-lake endemicity.

We include as endemics those species and subspecies of fish found only in a single lake and its tributaries. We define lake age as the duration that a lake basin has been continuously occupied by water, as estimated from lake sediment cores and from the timing of tectonic activity, glaciation, volcanism, natural damming, and impact events. Nonendemic native fish species names were standardized across lakes using FishBase (Froese and Pauly 2016) to consolidate synonyms since original data sources spanned 8 decades and encompassed many taxonomic revisions. Names of endemic fish were taken from the literature except where misapplication of a species name or taxonomic revision could be determined.

Analyses on Lakes with Age Data

Presence/absence and the numbers of endemic fish taxa were evaluated for the 252 lakes with estimates of age and with complete data for all predictors. The distributions of age, area, and latitude for these lakes are given in the appendix (fig. A2). A generalized linear model (GLM) using the presence or absence of endemic fish as a binomially distributed response, and a logit link function was constructed using the absolute value of lake latitude, hemisphere, logtransformed lake surface area, elevation, and age. A second model was constructed for the log-transformed number of endemic fish species or subspecies in lakes containing at least one endemic fish with the same predictor variables. Lake perimeter was not used in these models as it was strongly correlated with lake area when both were log transformed and standardized for analysis of the probability that a lake contains an endemic (r = 0.90, P < .0001) and for the restricted data set of lakes that had age data and contained endemics (r = 0.96, P < .0001). To facilitate comparison of effect sizes, standardized z-scores were calculated for each variable except hemisphere prior to analysis by centering and scaling each variable based on its mean and standard deviation.

Multiple Missing Data Imputation on the Full Lake Data Set

To test for the potential effect of predictors for which data were limited (e.g., productivity, native species richness, maximum lake depth), a multiple missing data imputation was performed using chained equations (Buuren and Groothuis-Oudshoorn 2011). Multiple imputation repeatedly generates imputed data sets using all nonmissing data (e.g., the presence and number of endemics for all 2,746 lakes in our case) but draws different plausible values for missing data that reflect the range of uncertainty in those missing data values. Pooled estimates are calculated on the set of multiple imputation results such that significant effects are found only when they are consistent across imputations.

We pooled results from GLMs on 100 imputed data sets, allowing us to obtain parameter estimates using all actual predictor values from the full database of 2,746 lakes. This imputation included lake age, maximum depth, mean pH, maximum surface water temperature, the richness of nonendemic native fish, chlorophyll a, total phosphorus, and maximum Secchi depth. All parameters except pH were log transformed prior to imputation. Native fish species richness was used for only the 288 lakes where sources attempted to catalog all fish species, to prevent including artificially low values for lakes where the full fish community (e.g., nonsport fish) is not well documented. GLMs predicting the probability that a lake contains an endemic and the number of endemics in a lake were evaluated using these additional predictors. We report the fraction of missing information, the ratio of the difference of information in the complete versus the incomplete data sets to the information contained in the complete data set, as well as the total variance due to missing data (λ ; Buuren and Groothuis-Oudshoorn 2011). Although λ was nontrivial and reached values up to 0.5, our use of 100 imputations (an atypically large number of imputations in practice) is approximately double the number needed to produce 95% confidence in confidence interval half widths as well as the estimate of λ (Bodner 2008). Although no similar theoretical generalization on the number of imputations needed to achieve confidence in the estimate of P values is available, variability in the estimate is unlikely to alter the rejection of a null hypothesis at the .05 significance level when $P \leq .01$. This condition is satisfied for all our significant findings except for the effect of lake age (.01 $\leq P \leq$.05), for which we have run a separate, explicit analysis as previously described, and, nevertheless, the variability in these estimates is minimized by the large number of imputations used.

Comparison to Fish Endemism Patterns in Rivers

The distribution of endemics at the regional river-basin scale was evaluated using data from the Fish-SPRICH database (Brosse et al. 2013), which contains 4,193 endemic fish that are restricted to single river basins. GLMs predicting the presence and number of endemic fish were performed using latitude, area, range of elevations within a basin, hemisphere, and native fish richness for 928 of 1,054 river basins without missing data. Age information for river basins is not available. Area, elevational range, and native fish richness were log transformed.

Phylogenetic Considerations and Potential Sampling Bias

A species-level phylogeny for freshwater fish at a global scale does not exist. Therefore, we were unable to employ standard phylogenetic controls in our analyses. Instead, we used other means to examine the potential effects of sampling bias and evolutionary nonindependence on our results. Speciation rates are often estimated per lineage in phylogenetic analyses to control for the number of lineages contributing to the overall speciation rate (Morlon 2014). We examined whether the effect of latitude on the probability that a lake contained an endemic species was the result of an increase in the number of fish lineages per lake at low latitudes relative to high latitudes. We estimated the number of lineages per lake by taking separate sums of the unique number of families and genera in lakes where complete fish species lists were available and tested whether the probability that a lake contains an endemic was influenced by the number of lineages in that lake using GLMs for families and genera.

To determine whether our model output from the data set of 252 lakes with age data was driven by the African Great Lakes, which display extreme fish endemism relative to other lakes, we reevaluated these models for both the presence and the number of endemics after excluding (1) the three largest and most speciose African Great Lakes (Victoria, Tanganyika, and Malawi) and (2) up to eight of the largest African Great Lakes.

To determine whether our results were robust to excluding the remarkable endemism exhibited by the family Cichlidae worldwide, we also reevaluated our original models after removing all endemic cichlids from the lake analysis. To further examine whether our results depend on the distribution of endemics within particular families, the latitudinal centers of distribution of both endemics and nonendemic natives within our lakes database were calculated for each of the 47 fish families containing at least one endemic. The mean latitude of endemic fish occurrence was compared to that of nonendemic native fish for the 41 families with data on both endemic and native distributions. The latitudinal center for each species or subspecies was determined by averaging the latitudes of each of the lakes in which it occurs. Families excluded from this analysis were entirely or predominantly composed of endemics for which no native distribution could be calculated (e.g., the Abyssocottidae of Lake Baikal; Sherbakov 1999).

Finally, as in any geographic comparison of biogeography, our methods require that the division of taxonomic units into species and subspecies is comparable across latitudes such that lineages represent similar subsets of a phylogeny. Moreover, geographic differences in sampling intensity may introduce bias. To evaluate the possible effect of sampling bias, we compiled records of the date of description of the endemic fish in our database and determined whether new endemic fish were being described at different rates in tropical versus extratropical regions.

Results

A total of 1,895 endemic species and 38 subspecies of fish were found, inhabiting 107 lakes worldwide (examples: fig. 1; distribution: fig. 2). For the set of 252 lakes with associated age data, the probability that a lake contains an endemic fish and the total number of endemic fish per lake are significantly associated with increased lake age (GLM, z = 5.225, P < .001), increased area (GLM, z = 2.808, P < .005), and lower latitude (GLM, z = -4.679, P < .0001; fig. 3), and the standardized effect sizes of these variables are similar in magnitude and statistically indistinguishable. Together, these effects are highly predictive of the probability that a lake contains an endemic fish (53.4% of total variance explained; fig. 4). Similarly, lake age (GLM, z = 4.042, P < .001), area (GLM, z = 6.394, P < .0001), and latitude (GLM, z = -3.880, P < .001) are jointly predictive of the number of endemic fish per lake (34.7% of total variance explained; fig. 5), with the probability that a lake contains an endemic increasing with lake age and area and decreasing with latitude. Lakes in the Western Hemisphere had a lower probability of containing an endemic fish (GLM, z =-3.839, P < .001), but for lakes with endemic fish there was no difference in the number of endemic fish between hemispheres (GLM, z = -1.132, P = .26; fig. 3).

Our analysis of total endemic species richness includes endemics that may have evolved via cladogenesis (i.e., one species evolves into two or more new species) or by anagenesis (i.e., one species evolves into a single new species). Cladogenesis is most often linked to diversification, as it increases species richness locally (within a lake in our case), whereas anagenesis does not (although anagenetic speciation may increase regional species richness through increases in beta diversity). Coyne and Price (2000), in their analysis of potential sympatric speciation events on islands, proposed that a count of the number of genera with two or more endemic species provides a measure of the number of lineages



Figure 1: Examples of endemic fish species from 11 different fish families. *A*, Adrianichthyidae (*Oryzias nigrimas*), Lake Poso, Indonesia; *B*, Characidae (*Moenkhausia pittieri*), Lake Valencia, Venezuela; *C*, Cichlidae (*Haplochromis nyererei*), Lake Victoria; *D*, Clariidae (*Bathyclarias foveolatus*), Lake Malawi; *E*, Cyprinidae (*Carassius cuvieri*), Lake Biwa, Japan; *F*, Gobiidae (*Benthophilus casachicus*), Caspian Sea; *G*, Mastacembelidae (*Mastacembelus ellipsifer*), Lake Tanganyika; *H*, Melanotaeniidae (*Melanotaenia lacustris*), Lake Kutubu, Papua New Guinea; *I*, Mochokidae (*Synodontis grandiops*), Lake Tanganyika; *J*, Poeciliidae (*Lamprichthys tanganicanus*), Lake Tanganyika; *K*, Salmonidae (*Salmo letnica*), Lake Ohrid, Macedonia. Photos were taken by the following photographers: *A*, Wikimedia Commons user Daiju Azuma (CC-BY-SA 2.5 license); *B*, Wikimedia Commons user Michael Palmer of photorasa.com (CC-BY-SA 4.0 license); *C*, Kevin Bauman of http://www.african-cichlid.com (CC-BY 1.0 license); *D*, Wikispecies user Michael K. Oliver (CC-BY-SA 4.0 license); *E*, Wikimedia Commons user Daiju Azuma (CC-BY-SA 3.0 license); *F*, Yuriy Kvach (Wikimedia Commons user Ykvach; CC-BY-SA 3.0 license); *G*, French Wikipedia user Jtanganyika (CC-BY-SA 3.0 license); *H*, Wikimedia Commons user Thomnight (CC-BY-SA 4.0 license); *I*, Wikimedia Commons user Haps (CC-BY-SA 3.0 license); *J*, Loury Cédric (Wikimedia Commons user Cedricguppy; CC-BY-SA 4.0 license); *K*, Wikimedia Commons user Haps (CC-BY-SA 3.0 license).

that have diversified in situ by cladogenesis. Based on this alternative criterion of speciation, 31 of the 252 lakes with age estimates had evidence of cladogenesis. An analysis of the presence/absence of cladogenesis in these 252 lakes using a GLM demonstrates strong and significant effects of latitude (GLM, estimate = -1.273, P < .001), area (GLM, estimate = 1.562, P < .0001).

Subsequent analyses using a multiple missing data imputation allowed us to examine the effects of native fish species richness, as well as physical, chemical, and productivity variables on endemic fish in all 2,746 lakes. These analyses again revealed strong effects where older lakes, lakes of larger area, and lower-latitude lakes had a greater probability of containing an endemic and having a greater number of endemics. However, no detectable effects of species



Figure 2: Distribution and richness of endemic fish species and subspecies (N = 1,933) in the 2,746 largest natural lakes in the world (surface area $\geq 50 \text{ km}^2$).

richness, productivity, maximum depth, pH, or temperature on endemism were found (table A1; tables A1-A5 available online).

Further analysis of an independently assembled global database of endemic freshwater fish (Fish-SPRICH; Brosse et al. 2013) at the river-basin scale (not individual lakes) corroborates the patterns observed for lake fish. In river basins, endemic fish presence and endemic fish richness decreased significantly with latitude and increased significantly with area, and additional significant, positive effects were observed for both elevational range and native fish richness. The effect of basin age is unknown for this data set: figure A3, table A2.

Much of the global endemicity of lake fish occurs in the African Great Lakes, where it is famously concentrated in a single family, the Cichlidae. Nevertheless, our results were robust to (1) excluding the largest three or eight African Great Lakes from the data set (table A3) and (2) excluding the >1,000 endemic fish in the family Cichlidae, whose mean latitudinal distribution is only 8.1° from the equator (table A4). Moreover, an analysis comparing the geographic centers of distribution of nonendemic fish species to the mean latitudes of endemic fish shows that for 34 of 41 families, endemics are found at comparatively lower latitudes than nonendemics (exact binomial test, P < .0001; fig. 6; table A5). The two families with the greatest number of endemic species, the Cichlidae (n = 745 named species) and Cyprinidae (n = 139), both display this low-latitude bias in the distribution of endemics (table A5).

Tropical lakes have more lineages (as measured by the total number of families or genera) than temperate lakes, but there is no evidence that the number of lineages affects whether a lake contains an endemic (fig. A4). Thus, the greater probably of a speciation event occurring in lowlatitude lakes is not a function of more fish lineages in the tropics. Further, a comparison of temperate and tropical differences in the rate at which endemic species are being described suggests that the greater endemic species richness in tropical regions may actually underestimate latitudinal differences in diversification, since the description of new endemic fish is increasing much faster at tropical than extratropical latitudes (fig. A5).

Discussion

Lakes, like islands, provide model systems for studying evolution (Warren et al. 2015). Our analyses, along with many others (see "Introduction"), demonstrate a positive effect of age and area on the probability of in situ speciation and the extent of diversification in island-like systems. Here, we provide addition evidence for a strong relationship between latitude and diversification that is independent of age and area effects. Our results show that the probability of a speciation event and the extent of diversification increase with decreasing latitude, as judged by the distribution of endemicity in both lake and river fishes. Previous work on island endemics has used a similar approach to study how in situ speciation is affected by island size, age, and isolation (see Warren et al. 2015 for a recent review), to examine the incidence of sympatric speciation (Coyne and Price 2000) and the spatial scale of speciation on islands (Kisel and Barraclough 2010). However, to our knowledge, no island studies have directly examined how the probability of in situ spe-



Figure 3: Relationship between biogeographic variables and the probability that an endemic species occurs in a lake (*A*) and the number of endemic fish species in a lake (*B*). Boxes represent the direction and magnitude of the standardized effect size of each variable with associated normal-based 95% confidence intervals. Asterisks denote significant predictor variables: one asterisk = P < .01, two asterisks = P < .001, three asterisks = P < .001.

ciation varies with latitude. Recently, Jetz and Fine (2012) and Belmaker and Jetz (2015) examined the influence of historical time, area, and present-day climate on global gradients in terrestrial vertebrate diversity by dividing the Earth into 32 evolutionarily distinct "bioregions." Like our lake analyses, they found strong, positive effects of bioregion age and area on endemic species richness. Unlike our study, Belmaker and Jetz (2015) conclude that diversification rates (estimated phylogenetically) appear to have a relatively minor influence on broad-scale patterns in species richness.

The observation that area is strongly predictive of endemism in lake fishes is consistent with findings in African cichlids (Wagner et al. 2014), *Anolis* lizards (Losos and Schluter 2000), poeciliid fish (Furness et al. 2016), and other taxa (e.g., Kisel and Barraclough 2010), suggesting that there is a minimum area for in situ speciation. We searched the literature for endemism in small lakes (surface areas <50 km², below the minimum size used in our analysis) and found endemics in 74 of these water bodies, including those as small as 0.0028 km² (Tilapia guinasana in Lake Guinas, Namibia; Nxomani et al. 1999). Thus, in situ speciation can occur in small lakes, but despite the abundance of such lakes worldwide (>243,000 lakes between 0.1 and 50 km² in the Global Lakes and Wetlands Database; Lehner and Döll 2004), our finding of only 74 lakes (<50 km²) with endemic fish suggests that these lakes have limited in situ speciation. Area may increase the rate of speciation through larger population sizes, greater habitat heterogeneity (Kisel et al. 2011), and increased environmental stability through time. There are only a few old, deep, very large lakes worldwide, and these also contain the extremes in the number of potential cladogenesis events in multiple genera (using Coyne and Price's [2000] criterion: Lake Malawi 38, Lake Tanganyika 31, Caspian Sea 10, Lake Baikal 9).

Molecular evidence suggests that the pace of divergence in extratropical lakes has been slow, even in those containing relatively large adaptive radiations such as the Caspian Sea and Lake Baikal. For example, the endemic Proto-Caspian gobies of the subfamily Benthophilinae are estimated to have originated 10 million years ago from a common ancestor that diverged into multiple genera 4.29-6.25 million years ago, with the most recent identifiable radiations occurring 1-2 million years ago (Neilson and Stepien 2009). In Lake Baikal, the age of the root of its 33-species cottoid fish radiation is uncertain but is estimated at 1.2-6.5 Ma (Kontula et al. 2003). These deep-rooted divergence events contrast with the rapid divergence of hundreds of species of cichlids in the comparably young Lake Victoria, whose species diverged 15,000-100,000 years ago (Brawand et al. 2014). Another prime example of rapid ecological diversification in a tropical lake is the divergence of Labeobarbus fish in Lake Tana, Ethiopia. This lake has been continuously isolated from the rest of the Nile basin and dried completely within the past 25,000 years (De Graaf et al. 2007). It is hypothesized that a single riverine ancestor colonized Lake Tana and generated a 15-species endemic flock in 10,000-25,000 years (De Graaf et al. 2010), with clear ecological divergence among species, including piscivory, a rare trait in the family Cyprinidae (De Graaf et al. 2008).

Latitudinal Drivers of Speciation beyond Age and Area

Our analysis of endemism in freshwater fish provides fresh insight into how age, area, and latitude are correlated with the probability of speciation and the extent of diversification. Latitude, of course, does not directly affect speciation and diversification but instead is correlated with mechanisms that







Figure 5: Heat map of the estimated number of endemic fish per lake (color scale) in tropical latitudes ($23.43^{\circ}S-23.43^{\circ}N$; *A*) and extratropical latitudes ($>23.43^{\circ}S$ or °N; *B*) as a function of lake age (*X*-axis) and surface area (*Y*-axis) for lakes whose age is known and that contain at least one endemic fish.

may drive these processes. Several hypotheses have linked the LDG to faster diversification at lower latitudes (Mittelbach et al. 2007; Brown 2014; Fine 2015). For example, the evolutionary speed hypothesis (Rohde 1992; Allen and Gillooly 2006) postulates that molecular evolution (nucleotide substitution) is faster at higher temperatures, resulting in higher speciation rates at low latitudes. The biotic interactions hypothesis (Dobzhansky 1950; Schemske 2009; Schemske et al. 2009) proposes that the relatively benign and stable climate of the tropics leads to adaptation governed more by biotic than abiotic factors, resulting in faster speciation because of ongoing coevolution. Stable tropical climates may also result in lower extinction rates in the tropics, contributing to higher rates of diversification (e.g., Pyron 2014; Pulido-Santacruz and Weir 2016).

It has recently been suggested that speciation rates for some taxa are actually higher in temperate regions than in the tropics, because the relatively species-poor temperate zone provides greater opportunities for ecological divergence (e.g., more open niches) and therefore more rapid speciation (reviewed in Schluter 2016). There are well-documented cases of rapid diversification in temperate fishes, including species pairs within multiple families in postglacial lakes in the Northern Hemisphere (Taylor 1999; Schluter 2016), but few of these ecomorphs are formally recognized as species. The case can be made that some temperate-lake ecomorphs deserve species status (e.g., Knudsen et al. 2006; Harrod et al. 2010), but incipient speciation is suspected in many tropical lakes as well (e.g., Nxomani et al. 1999; Wilson et al. 2000; Herder et al. 2008). Our data in fact suggest that the description of tropical fish diversity likely lags that of the temperate zone (fig. A5).

Although species-poor environments can provide opportunities for ecological divergence and speciation, the flip



Figure 6: Degree latitude difference in the mean distribution of endemic fish (*Y*-axis) from the latitudinal centers of distribution of the nonendemic species (*X*-axis) for each of 41 families. Across fish families, endemics are distributed at lower latitudes than nonendemics (exact binomial test, P < .0001; below dashed line).

side of the coin is that over time species may become niches for other species. "Every species is potentially a resource on which some other species can in principle specialize or to which another species must adapt" (Vermeij 2005, p. 20). Strong biotic interactions, coupled with relatively benign abiotic conditions, can create ecological opportunities that allow tropical organisms to explore a wider range of niche dimensions than their temperate counterparts, promoting greater species diversity. For example, fish that mainly consume plants or fruits are common in the tropics but rare elsewhere (Horn et al. 2011; Correa et al. 2015). Tropical fishes also display a bewildering array of adaptations rarely seen in temperate fishes, including scale eating, parasite cleaning, and electrical communication (and predation). Some of these unusual traits have evolved independently in multiple families. For example, frugivory occurs in 17 Neotropical fish families (Correa et al. 2015), and scale eating has evolved at least 19 times in tropical lineages (Martin and Wainwright 2013). In addition, adaptive evolution of visual receptors in African cichlids that occupy different light environments (Wagner et al. 2012) and electroreceptors for communication in African catfish (mormyrids; Carlson et al. 2011) are thought to promote the extensive diversification of these groups.

We suggest that the more extensive diversification of fish (and other taxa) at low latitudes may be driven in part by ecological opportunities promoted by tropical climates and by the coevolution of strong species interactions. Endemism in lakes is not unique to fish and has been observed across a wide array of taxa including both benthic and planktonic crustaceans (Marijnissen et al. 2006; Boxshall and Defaye 2008; Väinölä et al. 2008; Von Rintelen et al. 2010; Lorenschat et al. 2014), molluscs (West and Michel 2000; Von Rintelen and Glaubrecht 2005, 2006; Albrecht et al. 2006), and sponges (Meixner et al. 2007). Although freshwater fish follow a classic LDG, it has been suggested that the overall strength of the LDG is weaker in freshwater systems when compared with marine and terrestrial realms (Hillebrand 2004), and some freshwater groups appear to show an inverse LDG (Heino 2001). Comparisons of our findings in fish to other taxonomic groups may improve understanding of why diversification varies with latitude.

In conclusion, a global analysis controlling for the effects of lake age and area reveals a strong and independent effect of latitude on the probably of in situ speciation and on the extent of diversification in freshwater endemic fish. Mechanisms underlying the positive effects of age and area on diversification in island-like systems are well known (Warren et al. 2015); however, understanding why diversification may be enhanced at low latitudes remains a challenge. We suggest that greater diversification in the tropics may be due to biological mechanisms that differ in kind and/or magnitude from those in temperate regions. The LDG may very well reflect a persistent difference in the selective forces across what is not merely a geographic arc but the principal climatic gradient on Earth.

Acknowledgments

We thank F. Condamine, R. Ricklefs, D. Schluter, J. Sobel, two anonymous reviewers, and the editors for comments that greatly improved the manuscript. Support from National Science Foundation grant 1456615 to G.G.M. and D.W.S. is gratefully acknowledged. This is KBS contribution no. 1984.

Literature Cited

- Albrecht, C., S. Trajanovski, K. Kuhn, B. Streit, and T. Wilke. 2006. Rapid evolution of an ancient lake species flock: freshwater limpets (Gastropoda: Ancylidae) in the Balkan Lake Ohrid. Organisms Diversity and Evolution 6:294–307.
- Allen, A. P., and J. F. Gillooly. 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. Ecology Letters 9:947–954.
- Barbour, C. D., and J. H. Brown. 1974. Fish species diversity in lakes. American Naturalist 108:473–489.
- Belmaker, J., and W. Jetz. 2015. Relative roles of ecological and energetic constraints, diversification rates and region history on global species richness gradients. Ecology Letters 18:563–571.
- Bodner, T. E. 2008. What improves with increased missing data imputations? Structural Equation Modeling 15:651–675.
- Boxshall, G. A., and D. Defaye. 2008. Global diversity of copepods (Crustacea: Copepoda) in freshwater. Hydrobiologia 595:195–207.
- Brawand, D., C. E. Wagner, Y. I. Li, M. Malinsky, I. Keller, S. Fan, O. Simakov, et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature 513:375–381.
- Brosse, S., O. Beauchard, S. Blanchet, H. H. Dürr, G. Grenouillet, B. Hugueny, C. Lauzeral, et al. 2013. Fish-SPRICH: a database of freshwater fish species richness throughout the world. Hydrobiologia 700:343–349.
- Brown, J. H. 2014. Why are there so many species in the tropics? Journal of Biogeography 41:8–22.
- Buuren, S., and K. Groothuis-Oudshoorn. 2011. mice: multivariate imputation by chained equations in R. Journal of Statistical Software 45. doi:10.18637/jss.v045.i03.
- Carlson, B. A., S. M. Hasan, M. Hollmann, D. B. Miller, L. J. Harmon, and M. E. Arnegard. 2011. Brain evolution triggers increased diversification of electric fishes. Science 332:583–586.
- Correa, S. B., R. Costa-Pereira, T. Fleming, M. Goulding, and J. T. Anderson. 2015. Neotropical fish-fruit interactions: eco-evolutionary dynamics and conservation. Biological Reviews 90:1263–1278.
- Coyne, J. A., and T. D. Price. 2000. Little evidence for sympatric speciation in island birds. Evolution 54: 2166–2171.
- De Graaf, M., E. Dejen, J. W. Osse, and F. A. Sibbing. 2008. Adaptive radiation of Lake Tana's (Ethiopia) *Labeobarbus* species flock (Pisces, Cyprinidae). Marine and Freshwater Research 59:391–407.
- De Graaf, M., H. J. Megens, J. Samallo, and F. A. Sibbing. 2007. Evolutionary origin of Lake Tana's (Ethiopia) small *Barbus* species: indications of rapid ecological divergence and speciation. Animal Biology 57:39–48.

614 The American Naturalist

— 2010. Preliminary insight into the age and origin of the Labeobarbus fish species flock from Lake Tana (Ethiopia) using the mtDNA cytochrome b gene. Molecular Phylogenetics and Evolution 54:336–343.

- Dobzhansky, T. 1950. Evolution in the tropics. American Scientist 38:209-221.
- Doi, H., M. J. Vander Zanden, and H. Hillebrand. 2012. Shorter food chain length in ancient lakes: evidence from a global synthesis. PLoS ONE 7:e37856.
- Fine, P. V. A. 2015. Ecological and evolutionary drivers of geographic variation in species diversity. Annual Review of Ecology, Evolution, and Systematics 46:369–392.

Froese, R., and D. Pauly. 2016. FishBase. http://www.fishbase.org/.

- Furness, A. I., D. N. Reznick, and J. C. Avise. 2016. Ecological, evolutionary and human-mediated determinants of poeciliid species richness on Caribbean islands. Journal of Biogeography 43:1349–1359. Gaston, K. G. 2000. Global patterns in biodiversity. Nature 405:220–
- 227.
- Gillespie, R. G., and B. G. Baldwin. 2010. Island biogeography of remote archipelagos: interplay between ecological and evolutionary processes. Pages 358–387 *in* J. B. Losos and R. E. Ricklefs, eds. The theory of island biogeography revisited. Princeton University Press, Princeton, NJ.
- Hanly, P. J., G. G. Mittelbach, and D. W. Schemske. 2016. Data from: Speciation and the latitudinal diversity gradient: insights from the global distribution of endemic fish. American Naturalist, Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.70sr1.
- Harrod, C., J. Mallela, and K. K. Kahilainen. 2010. Phenotypeenvironment correlations in a putative whitefish adaptive radiation. Journal of Animal Ecology 79:1057–1068.
- Heino, J. 2001. Regional gradient analysis of freshwater biota: do similar biogeographic patterns exist among multiple taxonomic groups? Journal of Biogeography 28:69–76.
- Herder, F., J. Pfaender, and U. K. Schliewen. 2008. Adaptive sympatric speciation of polychromatic "roundfin" sailfin silverside fish in Lake Matano (Sulawesi). Evolution 62:2178–2195.
- Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. American Naturalist 163:192–211.
- Horn, M. H., S. B. Correa, P. Parolin, B. J. A. Pollux, J. T. Anderson, C. Lucas, P. Widmann, A. Tjiu, M. Galetti, and M. Goulding. 2011. Seed dispersal by fishes in tropical and temperate fresh waters: the growing evidence. Acta Oecologica 37:561–577.
- Jetz, W., and P. V. A. Fine. 2012. Global gradients in vertebrate diversity predicted by historical area-productivity dynamics and contemporary environment. PLoS Biology 10:e1001292.
- Kisel, Y., and T. G. Barraclough. 2010. Speciation has a spatial scale that depends on levels of gene flow. American Naturalist 175:316– 334.
- Kisel, Y., L. McInnes, N. H. Toomey, and C. D. L. Orme. 2011. How diversification rates and diversity limits combine to create large-scale species-area relationships. Philosophical Transactions of the Royal Society B 366:2514–2525.
- Knudsen, R., A. Klemetsen, P. A. Amundsen, and B. Hermansen. 2006. Incipient speciation through niche expansion: an example from the Arctic charr in a subarctic lake. Proceedings of the Royal Society B 273:2291–2298.
- Kontula, T., S. V. Kirilchik, and R. Väinölä. 2003. Endemic diversification of the monophyletic cottoid fish species flock in Lake Baikal explored with mtDNA sequencing. Molecular Phylogenetics and Evolution 27:143–155.

- Lehner, B., and P. Döll. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. Journal of Hydrology 296:1–22.
- Lorenschat, J., L. Pérez, A. Correa-Metrio, M. Brenner, U. von Bramann, and A. Schwalb. 2014. Diversity and spatial distribution of extant freshwater ostracodes (Crustacea) in ancient Lake Ohrid (Macedonia/Albania). Diversity 6:524–550.
- Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary speciesarea relationship. Nature 408:847–850.
- Marijnissen, S. A., E. Michel, S. R. Daniels, D. Erpenbeck, S. B. Menken, and F. R. Schram. 2006. Molecular evidence for recent divergence of Lake Tanganyika endemic crabs (Decapoda: Platythelphusidae). Molecular Phylogenetics and Evolution 40:628–634.
- Martin, C. H., and P. C. Wainwright. 2013. On the measurement of ecological novelty: scale-eating pupfish are separated by 168 my from other scale-eating fishes. PLoS ONE 8:e71164.
- Meixner, M. J., C. Lüter, C. Eckert, V. Itskovich, D. Janussen, T. von Rintelen, A. V. Bohne, J. M. Meixner, and W. R. Hess. 2007. Phylogenetic analysis of freshwater sponges provide evidence for endemism and radiation in ancient lakes. Molecular Phylogenetics and Evolution 45:875–886.
- Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, et al. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters 10:315–331.
- Morlon, H. 2014. Phylogenetic approaches for studying diversification. Ecology Letters 17:508–525.
- Neilson, M. E., and C. A. Stepien. 2009. Escape from the Ponto-Caspian: evolution and biogeography of an endemic goby species flock (Benthophilinae: Gobiidae: Teleostei). Molecular Phylogenetics and Evolution 52:84–102.
- Nxomani, C., A. J. Ribbink, and R. Kirby. 1999. DNA profiling of *Tilapia guinasana*, a species endemic to a single sinkhole, to determine the genetic divergence between color forms. Electrophoresis 20:1781–1785.
- Oberdorff, T., J. F. Guégan, and B. Hugueny. 1995. Global scale patterns of fish species richness in rivers. Ecography 18:345–352.
- Pigot, A. L., J. A. Tobias, and W. Jetz. 2016. Energetic constraints on species coexistence in birds. PLoS Biology 14:e1002407.
- Pulido-Santacruz, P., and J. T. Weir. 2016. Extinction as a driver of avian latitudinal diversity gradients. Evolution 70:860–872.
- Pyron, R. A. 2014. Temperate extinction in squamate reptiles and the roots of latitudinal diversity gradients. Global Ecology and Biogeography 23:1126–1134.
- Rabosky, D. L., and E. E. Goldberg. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. Systematic Biology 64:340–355.
- Rabosky, D. L., and H. Huang. 2016. A robust semi-parametric test for detecting trait-dependent diversification. Systematic Biology 65:181–193.
- Rabosky, D. L., P. O. Title, and H. Huang. 2015. Minimal effects of latitude on present-day speciation rates in New World birds. Proceedings of the Royal Society B 282:20142889.
- Ricklefs, R. E. 2006. Global variation in the diversification rate of passerine birds. Ecology 87:2468–2478.
- ———. 2007. Estimating diversification rates from phylogenetic information. Trends in Ecology and Evolution 22:601–610.
- Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527.
- Schemske, D. W. 2009. Biotic interactions and speciation in the tropics. Pages 219–239 in R. K. Butlin, J. R. Bridle, and D. Schluter, eds.

Speciation and patterns of diversity. Cambridge University Press, Cambridge.

- Schemske, D. W., G. G. Mittelbach, H. V. Cornell, J. M. Sobel, and K. Roy. 2009. Is there a latitudinal gradient in the importance of biotic interactions? Annual Review of Ecology, Evolution, and Systematics 40:245–269.
- Schluter, D. 2016. Speciation, ecological opportunity, and latitude. American Naturalist 187:1–18.
- Sherbakov, D. Y. 1999. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. Trends in Ecology and Evolution 14: 92–95.
- Steinbauer, M. J., R. Field, J. A. Grytnes, P. Trigas, C. Ah-Peng, F. Attorre, H. J. B. Birks, et al. 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. Global Ecology and Biogeography 25:1097–1107.
- Taylor, E. B. 1999. Species pairs of north temperate freshwater fishes: evolution, taxonomy, and conservation. Reviews in Fish Biology and Fisheries 9:299–324.
- Tedesco, P. A., F. Leprieur, B. Hugueny, S. Brosse, H. H. Dürr, O. Beauchard, F. Busson, and T. Oberdorff. 2012. Patterns and processes of global riverine fish endemism. Global Ecology and Biogeography 21:977–987.
- Tisseuil, C., J. F. Cornu, O. Beauchard, S. Brosse, W. Darwall, R. Holland, B. Hugueny, P. A. Tedesco, and T. Oberdorff. 2013. Global diversity patterns and cross-taxa convergence in freshwater systems. Journal of Animal Ecology 82:365–376.
- Väinölä, R., J. D. S. Witt, M. Grabowski, J. H. Bradbury, K. Jazdzewski, and B. Sket. 2008. Global diversity of amphipods (Amphipoda: Crustacea) in freshwater. Hydrobiologia 595:241–255.
- Vermeij, G. J. 2005. From phenomenology to first principles: toward a theory of diversity. Proceedings of the California Academy of Sciences 56:12–23.
- Von Rintelen, K., M. Glaubrecht, C. D. Schubart, A. Wessel, and T. Von Rintelen. 2010. Adaptive radiation and ecological diversification of Sulawesi's ancient lake shrimps. Evolution 64:3287– 3299.

- Von Rintelen, T., and M. Glaubrecht. 2005. Anatomy of an adaptive radiation: a unique reproductive strategy in the endemic freshwater gastropod *Tylomelania* (Cerithioidea: Pachychilidae) on Sulawesi, Indonesia and its biogeographical implications. Biological Journal of the Linnean Society 85:513–542.
- 2006. Rapid evolution of sessility in an endemic species flock of the freshwater bivalve *Corbicula* from ancient lakes on Sulawesi, Indonesia. Biology Letters 2:73–77.
- Wagner, C. E., L. J. Harmon, and O. Seehausen. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. Nature 487:366–369.
- 2014. Cichlid species-area relationships are shaped by adaptive radiations that scale with area. Ecology Letters 17:583–592.
- Warren, B. H., D. Simberloff, R. E. Ricklefs, R. Aguilée, F. L. Condamine, D. Gravel, H. Morlon, et al. 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecology Letters 18:200–217.
- Weigelt, P., M. J. Steinbauer, J. S. Cabral, and H. Kreft. 2016. Late Quaternary climate change shapes island biodiversity. Nature 532:99– 102.
- Weir, J. T., and D. Schluter. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315: 1574–1576.
- West, K., and E. Michel. 2000. The dynamics of endemic diversification: molecular phylogeny suggests an explosive origin of the thiarid gastropods of Lake Tanganyika. Advances in Ecological Research 31:331–354.
- Wilson, A. B., K. Noack-Kunnmann, and A. Meyer. 2000. Incipient speciation in sympatric Nicaraguan crater lake cichlid fishes: sexual selection versus ecological diversification. Proceedings of the Royal Society B 267:2133–2141.

Associate Editor: Susan Harrison Editor: Alice A. Winn



"During the past summer the author had no opportunity of fishing in the Raritan River, at or about New Brunswick, at which point the specimen was taken; but among a number of small collections from that river, no specimen of this cyprinoid occurred. From other streams, generally not in the basin of the Raritan, isolated specimens have occurred, and the distribution seems to be without reference to salt water, although the type, and two other specimens, were taken from streams having direct access to the sea." Figured: *Hybognathus*. From "Further Notes on New Jersey Fishes" by Charles C. Abbott (*The American Naturalist*, 1871, 12:717–720).