Ecological opportunity alters the timing and shape of adaptive radiation

Edward D. Burress and Milton Tan

1Department of Biological Sciences and Auburn University Museum of Natural History, Auburn University, Auburn, Alabama
2E-mail: edwarddburress@gmail.com
3Division of Infectious Diseases, School of Medicine, Emory University, Atlanta, Georgia

Received June 1, 2017
Accepted August 27, 2017

The uneven distribution of diversity is a conspicuous phenomenon across the tree of life. Ecological opportunity is a prominent catalyst of adaptive radiation and therefore may alter patterns of diversification. We evaluated the distribution of shifts in diversification rates across the cichlid phylogeny and the distribution of major clades across phylogenetic space. We also tested if ecological opportunity influenced these patterns. Colonization-associated ecological opportunity altered the tempo and mode of diversification during the adaptive radiation of cichlid fishes. Clades that arose following colonization events diversified faster than other clades. Speciation rate shifts were nonrandomly distributed across the phylogeny such that they were disproportionately concentrated around nodes that corresponded with colonization events (i.e., of continents, river basins, or lakes). Young clades tend to expand faster than older clades; however, colonization-associated ecological opportunity accentuated this pattern. There was an interaction between clade age and ecological opportunity that explained the trajectory of clades through phylogenetic space over time. Our results indicate that ecological opportunities afforded by continental and ecosystem-scale colonization events explain the dramatic speciation rate heterogeneity and phylogenetic imbalance that arose during the evolutionary history of cichlid fishes.

KEY WORDS: BAMM, colonization, diversification, evolution, macroevolution, speciation.

A central hypothesis about adaptive radiations is that rates of diversification will be fastest early in clade history as lineages rapidly disperse across novel regions of the adaptive landscape and subsequently slow as the niches associated with these regions saturate (Simpson 1953; Schluter 2000). Despite many examples from the fossil record (Foote 1994; Ruta et al. 2006; Hughes et al. 2013), early bursts of phenotypic evolution are rarely observed in comparative phylogenetic data (Harmon et al. 2010). Likewise, signals of diversity-dependent diversification in molecular phylogenies are often inconsistent and suggest that ecological opportunities rather than clade age regulates diversification rates (Rabosky and Lovette 2008; Rabosky 2009a,b; Rabosky et al. 2012; Rabosky and Hurlburt 2015). Regardless, the uneven distribution of diversity among extant clades is a conspicuous phenomenon across the tree of life. This discrepancy may be explained by diversification rate heterogeneity among clades (Kirkpatrick and Slatkin 1993; Chan and Moore 2002; Wiens 2011) or variable clade ages (McPeek and Brown 2007; Shi and Rabosky 2015).

Ecological opportunity has been characterized as relaxed competition in conjunction with ample resource availability (Schluter 2000) and is widely recognized a major catalyst of adaptive radiation (Rabosky & Lovette 2008; Rabosky 2009a; Losos 2010; Yoder et al. 2010; Mahler et al. 2010). There are myriad sources of ecological opportunity that may alter a lineage’s evolutionary trajectory, including the evolution of a key innovation, extinction of a competitor, or the colonization of a novel environment that may provide access to new resources (Losos and de Queiroz 1997; Schluter 2000; Yoder et al. 2010; Wainwright et al. 2012). The occurrence of these events should elicit a burst in diversification rates in response to ecological opportunity and subsequent slowing as those opportunities saturate (Simpson 1953; Schluter 2000; Burress et al. 2016). Lineages proceed through these stages
differently. For example, adaptive landscapes are variable across different organisms and environments and the magnitude of ecological opportunity, such as that afforded by the colonization of a novel ecosystem, will manifest in different ways and to different degrees. For example, among freshwater fishes, there is a large disparity in the ecological opportunities provided by lakes and rivers (Seehausen 2015). Additionally, ecosystem size constrains ecological opportunity and thereby in situ adaptive radiation (Wagner et al. 2014). Furthermore, ecological opportunity may interact with other mechanisms such as sexual selection (Wagner et al. 2012) or hybridization (Meier et al. 2017) that may accentuate the influence of ecological opportunity on adaptive radiation.

Cichlids are the third most species-rich family of fishes and exhibit diversity that is not equally distributed across clades and considerable diversification rate heterogeneity among clades (McMahan et al. 2013). Some clades have diversified rapidly, particularly those of Lakes Malawi and Victoria. For example, the ~700 species endemic to Lake Victoria arose over the last 150,000 years (Meier et al. 2017), whereas other much older lineages have relatively few extant species (McMahan et al. 2013). Cichlids also have a complex biogeographic history. For example, cichlids colonized South America via trans-Atlantic dispersal from Africa (Friedman et al. 2013; Matschiner et al. 2017) and later colonized Middle America also via trans-Atlantic dispersal (Ričan et al. 2013). In addition to these major continental scale colonization events, cichlids have also independently colonized and diversified within myriad lakes throughout East Africa (Wagner et al. 2014) and Middle America (Elmer et al. 2014). In some cases, lakes have been repeatedly colonized after multiple transitions between river and lake environments (Salzburger et al. 2005). These features make cichlid fishes an ideal model to assess the role of colonization-associated ecological opportunity on the tempo and mode of diversification during their adaptive radiation.

Our objectives were twofold. First, we estimated diversification rates across a family-wide phylogeny of 903 species of cichlid fishes that spans all major lineages. We focused on the relative rates among major clades, rates through time, and associations between rate shifts and colonization-associated ecological opportunity. We hypothesized that rate shifts would be nonrandomly distributed across the phylogeny due to ecological opportunities afforded by colonization events during the evolutionary history of cichlids. Therefore, we predicted that rate shifts would be disproportionally concentrated around nodes that correspond with known colonization events (i.e., colonization of a continent, river basin, or lake). Second, we quantified the relative shapes of different regions of the phylogeny (i.e., the distribution of major clades in phylogenetic space). Lineages explore phylogenetic space as they proceed through the stages of adaptive radiation (Lewitus & Morlon 2016) such that lineages should exhibit differently shaped phylogenies because of different modes of diversification that are apparent depending upon clade age and temporal proximity to changes in ecological opportunity. We hypothesized that ecological opportunity also alters the shape of phylogenies by accentuating phylogenetic characteristics such as phylogenetic expansion, phylogenetic imbalance, and the distribution of branching events through time. Therefore, we predicted that the relationship between phylogenetic shape and clade age would be different between colonization-associated clades and clades that did not arise in association with colonization-associated ecological opportunity.

Methods

PHYLOGENY CONSTRUCTION

To provide a phylogenetic hypothesis for comparative analyses while accommodating phylogenetic uncertainty, we sought to infer a distribution of densely sampled, time-calibrated phylogenetic trees. Synthetic time trees were inferred using several steps, including (1) compiling sequence data from GenBank, (2) bootstrapping maximum-likelihood phylogenetic analysis on these sequence data to infer a distribution of bootstrap trees representing all cichlids, (3) performing time-calibrated phylogenetic analysis to infer a posterior distribution of reference time trees representing a subset of cichlids, and finally, (4) integrating these results to infer divergence times on the bootstrap trees based on the reference time trees.

First, to compile phylogenetic data, we used the PHLAWD pipeline (Smith et al. 2009) to query GenBank for several commonly used genes in cichlid phylogenetics. PHLAWD was designed to rapidly compile phylogenetic datasets for densely sampled species-level phylogenetic analysis. A GenBank database was downloaded and set up using PHLAWD (accessed 2016 Dec 16). For each locus, PHLAWD queries the GenBank database and generates alignments where a single sequence represents each taxon (generally species, but also subspecies if present) (Smith et al. 2009). We queried the mitochondrial genes 12S, 16S, COI, cyt b, ND2, ND4 and the mitochondrial control region; and the nuclear genes 4c4, enc1, RAG1, RAG2, S7 intron 1, SH3PX3, glyt, myh6, plage3, ptr, sreb2, and tblr1. Outgroup taxa were selected from other ovalearians used in a time-calibrated phylogenetic analysis of cichlids (Friedman et al. 2013). Sequences were excluded if they represented unidentified taxa marked with sp., aff., cf., and hybrids by NCBI taxonomy. We also manually included several sequences as user sequences (Table S1). For some species, subspecies were also present in the alignment generated by the PHLAWD pipeline. For these species, we selected the longest sequence from among the subspecies present for each locus to represent the species.
Sequences were concatenated and alignment formats converted using AMAS (Borowiec 2016). PartitionFinder 2.1.1 (Lanfear et al. 2016) was used to determine the best model of evolution and partitioning scheme available for RAxML (Lanfear et al. 2016). A preliminary maximum likelihood tree was inferred, followed by manual inspection to remove incorrectly placed taxa. Finally, a maximum likelihood tree was inferred, and 100 bootstrap replicates were analyzed with branch lengths estimated. A distribution of reference time trees for subsequent congruification were inferred using the xml file made available by Friedman et al. (2013) for their time-calibrated phylogenetic analysis of percomorphs (doi:10.5061/dryad.48f62). Friedman et al. (2013) used 14 fossil calibrations for the ages of 13 outgroup nodes distributed across Percomorpha and the root node using a relaxed-clock analysis in BEAST to calibrate their phylogeny, which included 156 percomorph species, including 91 cichlids. Friedman et al. (2013) provides a relatively young estimate of cichlid age that infers trans-Atlantic dispersal from Africa to South America, which has been corroborated by subsequent analyses (Matschiner et al. 2017). We performed two runs using BEAST 1.8.2 (Drummond et al. 2012). Using burntrees.pl (https://github.com/nylander/Burntrees), burn-in was excluded and 50 evenly spaced time trees were sampled from the remaining posterior distribution of each run, resulting in 100 time trees. We then determined the maximum clade credibility (MCC) tree from these samples using TreeAnnotator.

We timed the maximum-likelihood tree using the MCC tree as a reference by the congruification method (Eastman et al. 2013) implemented in geiger version 2 (Pennell et al. 2014). The congruification software automatically finds matching nodes between a reference time tree to a target tree based on shared taxa. The node ages of shared nodes from the time tree are then used as secondary calibrations to calibrate the target tree, and divergence times were estimated using treePL (Smith and O’Meara 2012). We used the congruify and write.treePL functions to perform the congruification and generate the treePL input files for analysis. The treePL analysis was primed to determine optimal settings, and the tree was then time-calibrated with the thorough setting. This was also repeated for the 100 bootstrap trees and 100 sampled time trees to allow for comparative analyses accommodating uncertainty in tree topology and divergence time estimates. To maximize the number of shared nodes between each bootstrap tree and time tree pair, trees were paired to minimize the Robinson–Foulds distance between them, computed using phangorn (Schliep 2011). This was done iteratively without replacement; hence, for each bootstrap tree, the nearest time tree was selected, and this time tree was excluded from the set paired to the remaining bootstrap replicates. As with the ML and MCC tree congruification, each treePL analysis was primed to determine optimal settings, and divergence times were estimated with the thorough setting.

**DELIMITATION OF FOCAL CLADES**

We delimited 27 nonoverlapping clades across the cichlid phylogeny that exhibit variable species diversities and clade ages (Table 1). Ten of the 27 focal clades are defined by nodes that correspond to known colonization events that were represented in our phylogeny (i.e., of continents, river basins, or lakes). The focal nodes (i.e., colonization nodes hereafter) for these clades were defined as the node where all descending tips occur in the same region, except for South America and Lake Tanganyika, which included subsequent colonization events into other regions (i.e., Middle America and the Malagarasi River, respectively). The remaining 17 clades are not associated with colonization events. To evaluate the role of colonization-associated ecological opportunity, we used these two sets of clades—colonization-associated clades that arose after a colonization event and noncolonization-associated clades that did not arise following such events. We defined all clades that met these criteria: (1) clades must be nonoverlapping and (2) clades must consist of at least eight taxa to permit estimation of diversification rate and phylogeny shape. We first defined colonization-associated clades based on the literature and our phylogeny, then defined all possible noncolonization-associated clades given the two aforementioned criteria. A third criteria was used for the noncolonization-associated clades, which were (3) defined as large as possible to maximize the number of taxa represented in our focal clades. Our global scale has some drawbacks because of low resolution in some parts of the phylogeny. We exclude some small radiations that arose following colonization events due to poor representation in our phylogeny that prevented robust estimation of diversification rates and phylogeny shape; for example, some of the small African lake radiations such as the Alcolapia flock of Lake Natron (Ford et al. 2016; Kavembe et al. 2016) and the Coptodon flock of Lake Bermin (Martin et al. 2015). Additionally, some radiations are simply too small (i.e., species pairs) to robustly assess speciation rates and phylogeny shape (e.g., Amphilophus of the Nicaraguan crater lakes; Elmer et al. 2014). We also excluded one potential clade of Crenicichla that includes several such small radiations and therefore could not be unambiguously coded as a noncolonization-associated clade. Similarly, some taxa had long branches that were not allied with a clade and therefore were not encompassed in a focal clade. The species composition of each clade is delimited in Figure S1.

**RATES OF SPECIATION**

To quantify time-varying diversification rates, we implemented BAMM version 2.5.0 (Rabosky 2014; Rabosky et al. 2014; Shi and Rabosky 2015). Priors were computed using setBAMMPriors in BAMMtools. Clade-specific sampling probabilities were specified for major clades of cichlids based on the proportion of species in the tree of the total number of species described according to...
To assess the influence of colonization-associated ecological opportunity on the distribution of rate shifts across the cichlid phylogeny, we compared the marginal odds ratios of shifts at 11 nodes that correspond with colonization events and nodes that do not correspond to colonization events. Colonization nodes were defined to capture major biogeographic shifts during the evolutionary history of cichlids. These 11 nodes include the crown of each of the 10 aforementioned colonization-associated clades (Table 1) and the node corresponding with the initial colonization of Lake Tanganyika. A shift to a new diversification rate regime may not be coincident with a colonization event, but rather occur after some degree of lag following the colonization event (Burress et al. 2016); therefore, we also included the two subsequent descendant nodes for each colonization node, as well as each of their descendant nodes (up to four nodes, if they did not end in tips). This scheme can capture both immediate and delayed shifts to new rate regimes following each colonization event.

**RECONSTRUCTION OF PHYLOGENETIC SPACE**

We compared the shapes of the phylogenies of 27 focal clades (Table 1) using the RPANDA package, which allows for model-free assessment of macroevolutionary patterns by summarizing information contained in trees (Morlon et al. 2016). In this pro-
procedure, spectral density profiles are calculated using information about tree shape from the evolutionary distances between nodes (Lewitus & Morlon 2016). These profiles include three summary statistics (λ*, ψ, and η) that describe different aspects of tree shape. The ln-transformed principal eigenvalue (λ*) describes the distance between nodes such that speciation-poor and speciation-rich regions of the tree correspond with longer and shorter distances, respectively. Therefore, smaller eigenvalues depict regions of the tree undergoing phylogenetic expansion. Skewness (ψ) indicates the stem-to-tip distribution of branching events. In other words, if the eigenvalues are disproportionately small or large, the phylogeny is comprised of primarily stemmy events. In other words, if the eigenvalues are disproportionately small or large, the phylogeny is comprised of primarily stemmy events. Peak height (η) depicts the heterogeneity of branch lengths in the phylogeny such that low and high values indicate heterogeneous and homogeneous branch lengths, respectively.

We evaluated the structure of the 27 focal clades across phylogenetic space using Ward’s minimum variance clustering. Clusters were evaluated using confidence intervals derived from 10,000 multiscale bootstrap resampling. Cluster analysis and P values were calculated using pvclust package (Suzuki & Shimodaira 2011). Approximately unbiased P values are between 0 and 1 and depict how strongly the cluster is supported by the data (Suzuki & Shimodaira 2011). Clusters were evaluated using λ*, ψ, and η.

Lineages often explore phylogenetic space over time (Lewitus & Morlon 2016); therefore, we tested relationships between clade age and λ*, ψ, and η. Because these parameters describe the shape of regions of the cichlid phylogeny (i.e., the focal clades) rather than characteristics of the tips themselves and do not assume a model of evolution, we evaluated these relationships using a simple linear model. We determined if models were statistically significant by calculating effect sizes (correlation coefficient) and determining if 0 fell outside of the 95% confidence interval. For any significant models, we then evaluated the influence of ecological opportunity by testing for an interaction between ecological opportunity and clade age using analysis of covariance (ANCOVA). If present, we evaluated if interactions significantly improved the model using a drop test via the analysis of variance (ANOVA) function in R.

**Results**

**THE PHYLOGENY OF CICHLIDS**

We inferred the most densely sampled hypothesis to date of the cichlid phylogeny based on seven mitochondrial and 12 nuclear loci (Fig. S1). Consistent with previous family-wide analyses, we found poor node support along the backbone of the African clade as well as among species within the East African Great Lakes (Genner et al. 2007; McMahan et al. 2013; Friedman et al. 2013; Fig. S1). The major tribes were recovered as monophyletic; however, most of the East African tribes had poor support (Fig. S1). The relationships among tribes were generally consistent with previous analyses of extant cichlids (Genner et al. 2007; McMahan et al. 2013; Friedman et al. 2013).

**THE DISTRIBUTION OF RATE SHIFTS ACROSS THE PHYLOGENY**

The distribution of rate shifts across the phylogeny was highly skewed such that most rates were relatively slow and fast rates occurred mostly in isolated bouts (Fig. 1D). The distribution of diversification rates across the phylogeny was highly skewed such that most rates were relatively slow and fast rates occurred mostly in isolated bouts (Fig. 1D).

**THE RELATIVE RATES OF DIVERSIFICATION OF THE FOCAL CLADES**

The distribution of core rate shifts across the phylogeny corresponds with colonization events (Fig. 1E) and the rate shifts with the highest marginal odds ratios are conspicuously associated with a handful of lineages, namely the Geophagini, Heroini, Oreochromini, and the Malawian and Victorian super flocks (Fig. 1F). By far, the fastest and most exponentially increasing diversification rates occur within the lineages that arose with the African Lakes Malawi, Victoria, and Barombi Mbo (Fig. S4). Three river-dwelling clades also exhibit diversification rates that markedly increase over time, albeit at distinctly slower rates than their lake-dwelling counterparts. These include the heroine clade distributed throughout Middle America, the *Crenicichla lacustris* clade distributed throughout subtropical South America, and the *Serranochromis* clade distributed throughout Southern
Africa (Fig. S4). In contrast, all Lake Tanganyikan clades (i.e., Bathybatini and allies, Ectodini, Lamprologini, and Tropheini) exhibit diversification rates that distinctly slow through time (Fig. S4).

THE DISTRIBUTION OF CLADES ACROSS PHYLOGENETIC SPACE

The focal clades were widely distributed across phylogenetic space defined by $\lambda^*$, $\psi$, and $\eta$ (Fig. 2A) and varied in their degree of phylogenetic expansion, the relative stem-to-tip distribution of branching events, and heterogeneity of branch lengths. Most notably, the Lake Malawi super flock occupied a distinct region of phylogenetic space defined by high $\eta$ values (Fig. 2A). Clades varied much more continuously along the $\lambda^*$ and $\psi$ axes. Clades associated with smaller $\lambda^*$ values depicting rapid phylogenetic expansion include the clades of Lakes Victoria and Barombi Mbo, whereas clades associated with higher $\lambda^*$ values depicting slow phylogenetic expansion include the Tanganyikan Lamprologini and the Neotropical Cichlasomatini, Geophagini, and Heroini (Fig. 2A). The frequency distribution of spectral curves based on $\lambda^*$ values indicates high degrees of overlap among most clades, except for the handful of clades at each extreme (Fig. 2B). Clades associated with smaller $\psi$ values depicting branching events skewed toward the stem include the Lamprologini, Tropheini, and Chromidotilapiini, whereas clades associated with larger $\psi$ values depicting branching events skewed toward the tips include the Ptychochrominae, Bathybates and allies, and Orthochromis clades (Fig. 2A).

Of the parameters that describe phylogenetic space, only $\lambda^*$ values were significantly correlated with clade age (Fig. 3C) such that younger clades tend to exhibit more phylogenetic expansion (Fig. 2D). The focal clades were well distributed across $\lambda^*$ and clade age values (Fig. 2D). Cluster analysis revealed well-supported spatial structure in the distribution of the focal clades across phylogenetic space; however, there were no intuitive cases where clades from similar regions or ecosystems were clustered together (Fig. 2E).

THE INFLUENCE OF ECOLOGICAL OPPORTUNITY ON THE RATES AND SHAPE OF ADAPTIVE RADIATION

The probability of a shift occurring at one of the 11 colonization nodes was significantly higher than the probability of a shift at another node (Fig. 3A). Likewise, the marginal odds ratios of rate shifts were 45.9 times higher, on average, for colonization nodes than for noncolonization nodes (Fig. 3B; Table S3). The shift with the highest marginal odds ratio, by far, was associated with the colonization of Lake Malawi (Fig. 1E and F; Table S4). In fact, there were highly probable shifts on the colonization node and
Figure 2. The distribution of the focal clades across phylogenetic space defined by $\lambda^*$, $\psi$, and $\eta$ (A). The distributions of ln-transformed eigenvalues ($\lambda^*$) among the focal clades (B). Effect sizes ($r \pm 95\%$ confidence interval; Cl) calculated from the correlations between clade age and the parameters that describe phylogenetic space (C). The distribution of the focal clades across the relationship between clade age and ln-transformed eigenvalues (D). Hierarchical cluster analysis of the focal clades based parameters ($\lambda^*$, $\psi$, and $\eta$) that describe phylogenetic shape (E). See Table 1 for the clades that correspond to the abbreviations used in panel (E).

Figure 3. The role of ecological opportunity in the timing and shape of adaptive radiation. The probability of rate shifts at colonization-associated nodes and other nodes (A). The marginal odds ratios of rate shifts at colonization-associated nodes and other nodes (B). The relationship between clade age and ln-transformed eigenvalues ($\lambda^*$) among colonization-associated clades and other clades (C). Effect sizes ($r \pm 95\%$ CI) for these correlations are inset.
descendant nodes (Table S4). Rate shifts associated with the colonization of Lakes Victoria, Tanganyika, Barombi Mbo, and the La Plata Basin also had high marginal odds ratios (Table S4). The highest marginal odds ratios were coincident with the colonization node for Lake Barombi and the La Plata Basin; however, in Lakes Victoria and Tanganyika, the highest marginal odds ratios occurred on a descendant node (Table S4). The rate shift coincident with the colonization of Middle America also had a high marginal odds ratio; however, the rate shift with the highest marginal odds ratio occurred on a descendant node following the colonization of South America and colonization of the Congo River by Nanochromis and Steatocranus (Table S4). In contrast, the secondary colonization of Lake Tanganyika by the Trophini and colonization of the Malagarasi River by Orthochromis were not associated with a well-supported shift at the colonization nodes or their immediately descendant nodes (Table S4).

There was a significant positive relationship between In-transformed eigenvalues and clade age for both colonization-associated and noncolonization-associated clades (Fig. 3C) such that younger clades exhibit more phylogenetic expansion than older clades (Lewitus & Morlon 2016); however, we found that there was a significant interaction between ecological opportunity (i.e., colonization- vs. noncolonization-associated clades) and their λ* values ($F = 7.629; P = 0.014$; Fig. 3C; Table S5) such that clades that arose following colonization events explored phylogenetic space faster than other clades. A drop test indicated that including the interaction significantly improved the model ($F = 7.629; P = 0.011$; Tables S6 and S7).

**Discussion**

Ecological opportunity is considered one of the prominent features of adaptive radiation that may alter adaptive landscapes and therefore, evolutionary trajectories (Rabosky & Lovette 2008; Rabosky 2009a; Losos 2010; Yoder et al. 2010; Mahler et al. 2010). Ecological opportunity has been invoked to explain the rates and modes of diversification among adaptive radiations ranging from fishes, lizards, and birds, among others (Rabosky & Lovette 2008; Mahler et al. 2010; Burress et al. 2016); however, the extent to which ecological opportunity influences broader adaptive radiations remains unclear (Slater 2015). We demonstrate that colonization-associated ecological opportunity has influenced the rates and modes of diversification and has led to the uneven distribution of extant taxa across the cichlid phylogeny.

The colonization of lakes has long been considered a major source of ecological opportunity for fishes by providing access to novel resources and physical dimensions relative to the marine or riverine conditions from which the colonists dispersed (See hausen 2015; Burress 2015). Ecological diversification following lake colonization results in a few prodigious patterns across disparately related groups. For example, repeated diversification along the benthic-to-pelagic habitat axis by stickleback and by cichlids is a major theme among glacial, crater, and rift lakes (Rundle et al. 2000; Hulsey et al. 2013; Elmer et al. 2014; Machado-Schiaffino et al. 2014). We found that bursts in diversification occurred following the colonization of Lakes Tanganyika, Malawi, and Victoria. Although Lake Tanganyika may have served as the cradle from which the entire East African radiations arose, their geographic dispersal involved a series of complex colonization events. For example, there were multiple invasions of Lake Tanganyika from riverine ancestors. The first gave rise to the Tanganyikan species flock and a subsequent colonization gave rise to the Trophini (Salzburger et al. 2002). In contrast, Lake Malawi may have been colonized directly via dispersal from Lake Tanganyika (Salzburger et al. 2002; Joyce et al. 2011). Lastly, the Lake Victoria species flock may have arisen via hybridization between divergent riverine lineages (Meier et al. 2017). We also found evidence of a burst in diversification following the colonization of Lake Barombi Mbo, which is located near the Cameroonian coast. Similar to the resident assemblages of the East African Great Lakes, the Lake Barombi Mbo species flock also arose from riverine ancestors and may have a complex history of hybridization (Martin et al. 2015). There are clear themes associated with the colonization of lakes by cichlids. Namely, colonization-associated ecological opportunity results in bursts in diversification, ecological differentiation along the benthic-to-pelagic habitat axis, and complex instances of repeated gene flow (Joyce et al. 2011; Hulsey et al. 2013; Martin et al. 2015; Meier et al. 2017). Ecosystem size may ultimately constrain the degree to which in situ diversification unfolds such that larger lakes may provide more ecological opportunity (Wagner et al. 2014).

The evolutionary implications of colonizing rivers are less clear than lakes. In general, riverine species flocks are far less common (Seehausena 2015). This pattern suggests that diversification tends to occur via vicariance among rivers rather than in situ. We assessed four colonization-associated riverine clades that represent all known riverine species flocks among cichlids (Salzburger et al. 2002; Schwarzer et al. 2011; Piálek et al. 2012); however, it is noteworthy that the Orthochromis clade may have arose in an ancient paleolake that subsequently dried (Joyce et al. 2005). We found support for rate shifts associated with three of these four clades. Also in contrast with lake-dwelling clades, there are fewer themes across the riverine clades. For example, Crenicichla exhibit a burst in diversification rates following the colonization of the La Plata Basin, particularly within the Paraná and Uruguay Rivers, and also exhibit dramatic trophic diversity, including the rise of parallel ecomorphs (Piálek et al. 2012; Burress et al. 2013; Piálek et al. 2015). Steatocranus and Nanochromis exhibit some ecomorphologic diversity associated with specialization among different microhabitats in the lower
Congo River (Schwarzer et al. 2011), whereas Orthochromis have not ecomorphologically diversified much within the Malagarasi River (Seegers and De Vos 1988). Therefore, the nature of adaptive radiation following the colonization of rivers is variable and may be strongly constrained by the history of each lineage. For example, Crenicichla already colonized a novel functional region of the adaptive landscape characterized by high velocity ram feeding (López-Fernández et al. 2013), and thus may have been predisposed to exploiting any subsequent trophic-associated ecological opportunities. Alternatively, rivers themselves are heterogeneous and factors such as depth and habitat heterogeneity may constrain ecological opportunities for resident lineages. Disentangling the relative importance of intrinsic characteristics of the lineages and characteristics of the ecosystems in moderating or constraining adaptive radiations within rivers may be a fruitful topic for future research.

Many organisms have successfully colonized the Americas via trans-Atlantic dispersal from Africa, including freshwater fishes, reptiles, mammals, and plants (Lundberg 1993; de Queiroz 2005; Oaks 2011; Beaulieu et al. 2013). For cichlids, the active nature of this colonization event is a relatively recent paradigm shift from the long-held hypothesis of Gondwanan vicariance (Genner et al. 2007; Friedman et al. 2013; Matschiner et al. 2017). Cichlids likely used a combination of island hopping, dilute surface water from the outflow of a large river system such as the present-day Congo River, as well as some salinity tolerance to successfully colonize freshwaters separated by marine environments (Friedman et al. 2013 and references therein). Likewise, the subsequent colonization of Middle America from South America likely occurred via a combination of land bridges and island hopping through the Greater Antilles (Rich et al. 2013; Tagliacollo et al. 2017). Both of these continental scale colonization events corresponded with bursts in diversification rates, albeit generally less probable shifts than those associated with ecosystem-scale colonization events.

There were few unifying themes across the cichlid phylogeny such that diversification rates were heterogeneous and clade-specific phylogenetic shape was highly variable, even among clades that have classically been associated (i.e., the East African Great Lakes; Kocher 2004; Seehausen 2015). One apparent unifying theme is the role that colonization-associated ecological opportunity has played during the evolutionary history of cichlid fishes. Colonization events, ranging from continental to ecosystem scale elicited bursts in diversification rates and altered the movement of clades across phylogenetic space. There were several other emergent patterns: (1) lake- and river-associated clades do not occupy distinct regions of phylogenetic space and (2) the most probable rate shifts are associated with ecosystem-scale colonization events rather than continental scale events. Colonization-associated ecological opportunities may be responsible for much of the dramatic diversification rate heterogeneity and phylogenetic imbalance that arose during the evolutionary history of cichlid fishes and has served as a catalyst for numerous adaptive radiations of cichlid fishes.

Cichlids have colonized and flourished within incredibly diverse and sometimes harsh environments, including hot, alkaline, and hypersaline conditions (Trewavas 1983; Ford et al. 2015; 2016; Kavembe et al. 2016), torrential streams (Alter et al. 2017), and abyssal conditions (Schobert et al. 2013; Hahn et al. 2017). Ecological opportunities that lineages capitalize upon (i.e., successful colonization events) are likely only a subset of those that arise. The strong correspondence between colonization events and bursts in diversification not only emphasizes the critical role that colonization-associated ecological opportunities played during the evolutionary history of cichlids, but also emphasizes the importance of any intrinsic traits that permitted cichlids to successfully colonize given the opportunity. Most notably, colonizing alkaline lakes, often from comparatively acidic rivers, or colonizing freshwater ecosystems after dispersing through marine environments, points toward a clear demand upon cichlid physiological systems. Hence, we pose the hypothesis that physiological plasticity and adaptability predisposed cichlids to capitalizing upon colonization opportunities that arose throughout their evolutionary history may be the principal factor in driving their bewildering diversity.

AUTHOR CONTRIBUTIONS

EDB conceived the study, performed analyses, and wrote the paper. MT performed analyses and provided revisions of the paper.

ACKNOWLEDGMENTS

We thank B. Morgan and the High Performance Computing oversight committee for access and assistance with the Center for Advanced Science Innovation and Commerce (CASIC) supercomputer at Auburn University, and R. A. Petit III and T. D. Read for access and assistance with computing at Emory University. J. Oaks kindly provided feedback on an earlier draft of this manuscript. Insight from two anonymous reviewers and an Editor improved this paper.

DATA ARCHIVING

Data are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.nh539

LITERATURE CITED


Associate Editor: I. Lovette
Handling Editor: M. Noor

Supporting Information
Additional Supporting Information may be found in the online version of this article at the publisher’s website:

Table S1. Sequences that were manually included as user sequences in the alignment using the PHLAWD.
Table S2. Bayes factors comparing the probabilities of six to 25 rate shifts across the cichlid phylogeny.
Table S3. Summary of marginal odds ratios for colonization and non-colonization associated nodes.
Table S4. Marginal odds ratios for colonization nodes and immediate descendant nodes.
Table S5. ANCOVA results testing if the slope of the eigenvalue–clade age regression is different because of a clade age–EO (ecological opportunity) interaction.
Table S6. Model without interaction term for comparison.
Table S7. Results from drop test evaluating if including the interaction significantly improves the model.
Figure S1. Phylogeny of cichlid fishes used in the analyses for the article.
Figure S2. Prior and posterior distributions of the number of rate shifts based on priors of 1, 5, and 10 shifts.
Figure S3. The top-nine rate shift configurations from the 95% credible shift set.
Figure S4. Rate through time plots (mean ± 95% credible interval) for major clades discussed in the main text.