Evolutionary convergence classically occurs when species independently evolve similar phenotypes, often in association with similar selective environments (Losos et al. 1998; Schluter 2000; Losos 2011; Futuyma 2013). This process of convergence is most striking when it produces similar communities of species in different locations, with each community composed of species with divergent adaptations that are similar across locations. Well-known examples include the adaptive radiation of Anolis lizards on different islands in the Greater Antilles (Losos et al. 1998; Mahler et al. 2013) and of cichlid fishes in different lakes in East Africa (Clabaut et al. 2007; Muschick et al. 2012). Such strong convergence suggests that phenotypic evolution is frequently repeatable (Losos et al. 1998; Mahler et al. 2013), rather than subject to historical contingencies that may be caused by various factors, including different genomic architectures, developmental programs, or biotic interactions (Losos 2010). Furthermore, many-to-one mapping of morphology onto performance (Wainwright et al. 2005) can inhibit morphological convergence (Wainwright 2007; Collar et al. 2014). Similarly, the macroevolutionary landscape may be complex, with no convergent peaks (Arnold et al. 2001). In both of these latter cases, each lineage’s history of adaptation will bring it to different adaptive solutions to shared selective problems and thus prevent convergence. On the other hand, there might instead be more convergence at deeper time scales, as it may take considerable time to adapt to a new environment.
explicitly phylogenetic (Winemiller 1991). Furthermore, Melville et al. 2006; Moen et al. 2013) or have not been large-scale studies have included few communities (e.g. Grundler and Rabosky 2014), and community-focused evolution at large scales. Most studies of this topic have focused on particular phenotypes but not entire communities at the global scale (e.g. Wiens et al. 2006; Grundler and Rabosky 2014), and community-focused large-scale studies have included few communities (e.g. Melville et al. 2006; Moen et al. 2013) or have not been explicitly phylogenetic (Winemiller 1991). Furthermore, most studies only test convergence, not whether current phenotypes reflect convergence, history, or both (see instead Stayton 2006; Revell et al. 2007). One reason for this paucity of studies is that few approaches allow one to test the relative importance of history and convergence across scales in producing the phenotypes we see today (but see Langerhans and DeWitt 2004; Langerhans et al. 2006). Furthermore, no such approaches utilize Ornstein–Uhlenbeck (OU) models (Hansen 1997), which explicitly model the macroevolutionary process of adaptation to distinct selective environments (Butler and King 2004). Here we introduce such an approach and use it to examine convergence and history in anuran amphibians (frogs, including toads), a major vertebrate clade.

Our approach has two parts, centered on two questions that relate to how two types of historical effects can influence phenotypic evolution. First, when independent lineages colonize similar environments (e.g., microhabitats), do they converge toward the same phenotype? Or do their different evolutionary histories (and more specifically, different constraints shared by members of different clades) prevent such convergence? We compare “fully convergent” models, where microhabitat is the only factor that influences a species’s phenotype, to those in which its clade membership also influences its phenotype, corresponding to convergence at a shallower phylogenetic scale. Second, even if convergent adaptation has been important across clades, have shallower historical effects (such as limited time for adaptation) prevented species from reaching their inferred adaptive optimum? The estimated adaptive “optimum” in OU models is a statistical concept, not based on biomechanical or selection studies of what phenotype would function best in a given environment. Instead, it reflects a single phenotype toward which individual lineages evolve (Hansen 1997, 2012). Each lineage can have its own optimal phenotype due to its idiosyncratic evolutionary history and constraints, but the adaptive optimum estimated in the models (also called the primary optimum; Hansen 2012) is assumed to reflect selection due to a common factor shared by a set of species (in our case, those sharing a given microhabitat). It can take time to overcome constraints (e.g., genetic correlations, pleiotropy) and the impact of past environments on the phenotype. Both of these factors can cause species’ phenotypes to differ from the adaptive optimum. Here, we introduce a method to decompose the variation in species’ phenotypes around their inferred adaptive optimum. Systematic deviation from the current optimum (toward the optima of ancestral environments) would indicate that historical factors have prevented species from reaching the same, convergent adaptive optimum. In contrast, random deviations around the optimum would suggest that history is unimportant.

Using this two-pronged approach, we examine phenotypic evolution in frogs from around the world and over a time span of ~160 million years (myr). Anuran species use similar microhabitats across different communities globally, with arboreal, burrowing, terrestrial, and aquatic or semiaquatic species found in many local assemblages (Inger et al. 1984; Duellman 2005). However, anurans in different regions often belong to different clades (e.g., Hylidea dominates the New World and Ranoida the Old World; Roelants et al. 2007; Wiens 2007). Thus, different species across continents seem to have independently evolved to use these microhabitats (Bossuyt and Milinkovitch 2000; Young et al. 2005), potentially offering many independent evolutionary replicates for testing convergence. Furthermore, a recent study established the functional importance of the morphological traits that we study here and showed phenotypic similarity in frog species across three global sites, but explicitly tested convergence at only one site (Moen et al. 2013).

We studied 167 frog species from 10 sites around the world, analyzing data on microhabitat use and phylogeny from the literature and obtaining new morphological data from museum specimens from these localities. Using our two-part approach, we find that frogs show strong morphological convergence even at these large scales, with repeated convergence leading to similar ecomorphs in communities across the globe. We also find that lineages that have colonized their current microhabitat relatively recently differ from lineages that have been in the same microhabitat for much longer. But these differences are associated with their past ecology, and not the clades to which they belong. Overall, our results elucidate the relationships between evolution, ecology, and morphology in a major clade of vertebrates and reveal the insights that can be gained from this two-tiered phylogenetic approach to studying the relative importance of convergence and history.

Note that throughout the article, we use “convergence” to simply refer to the repeated evolution of similar phenotypes. We recognize that in the phylogenetics literature, convergence may specifically refer to the case in which similar phenotypes evolve from different ancestral states, whereas parallelism refers to the case in which the same ancestral state gives rise to the same derived state (see review in Wiens
et al. 2003). We also recognize that there is a separate literature on convergence as a type of homoplasy in phylogenetic analyses, which we do not address here.

**Materials and Methods**

**Species Sampling**

We measured anuran species from 10 local sites distributed around the world. Sites were chosen to represent major biogeographic regions based on Wallace’s evolutionary areas (i.e., species within each location are often distantly related to those from other locations; Holt et al. 2013), which allowed us to study species that use similar microhabitats but are in different clades. We focused on sampling local sites (rather than sampling random species within higher taxa) because we expect interactions between co-occurring species to ultimately drive diversification to utilize different microhabitats, following from the ecological theory of adaptive radiation (Schluter 2000), and because of our interest in testing for convergence across communities. However, we recognize that species that are presently sympatric in local communities may have evolved their phenotypic differences in allopatry.

Most specimens we measured had been collected at these 10 sites (707 of 754 total individuals). We were not able to measure every species documented from every site because many were insufficiently represented in museum collections. However, we chose species that collectively represented all microhabitat types at each site because many were insufficiently represented in museum collections than females. From each specimen (Online Supplementary Table S2) we first measured body length (snout-to-vent length), forelimb length, hindlimb length, head length, and head width using calipers (Online Appendix S1). We then took photographs of the hands and feet of each specimen and measured the area of finger and toe tips, area of interdigital foot webbing, and area of the inner metatarsal tubercle using ImageJ (ver. 1.42; Rasband 1997) to trace the circumference of each structure and calculate its area. We used the sums of individual webbing or digit tips across the entire foot or hand as data for analysis. Finally, we dissected out the two major muscle groups of the legs (those associated with the femur and the tibia/fibula) to calculate hindlimb muscle mass. We chose these 10 morphological variables given their demonstrated functional importance in swimming, jumping, clinging, and burrowing (Emerson 1976, 1991; Marsh 1994; Nauwelaerts et al. 2005; Moen et al. 2013). For all variables we used species means (mean $n = 4.49$ individuals per species; Online Supplementary Table S2) for statistical analyses. See Online Appendix S1 for full details of morphological data collection, including precise variable definitions.

We conducted principal components analysis (PCA) on the correlation matrix of the ln-transformed morphological data across all species to account for size-related redundancy in our variables (Jolicoeur 1963). We conducted both standard and phylogenetic PCA (Revell 2009) using the package phytools version 0.2-1 (Revell 2012) in R ver. 2.15 (R Core Team 2012) but only considered the phylogenetic PC scores from our second phylogeny (see below) for further analyses, given similar scores across phylogenies and methods (see Online Appendix S3). We retained all PC axes for further analyses to fully characterize variation among species and microhabitat categories (Monteiro 2013; Adams 2014). We obtained similar results in our OU model comparisons when only analyzing axes 2–4 (see Online Appendix S3).

**Microhabitat Use**

We placed each species into one of the five broad microhabitat categories that are standard in the literature on anuran ecology (Bossuyt and Milinkovitch...
We used three approaches to obtain a phylogeny and branch lengths for the 167 focal species. All three approaches were based upon the phylogeny and/or molecular data from Pyron and Wiens (2011), who used 12 genes (3 mitochondrial, 9 nuclear; up to 12,712 bases per taxon) and maximum likelihood to estimate a phylogeny of 2871 species of extant amphibians. We pruned the full data set to the 167 species in our study, with some straightforward taxon substitutions (see Online Appendix S2 for full details of substitutions). First, we used the (pruned) phylogeny and branch lengths (in substitutions per site) from Pyron and Wiens (2011). Second, we estimated a time-calibrated phylogeny using the Bayesian uncorrelated lognormal approach (in BEAST; Drummond and Rambaut 2007), the molecular data of Pyron and Wiens (2011), and nine fossil calibration points. For this analysis, we constrained the topology to that of Pyron and Wiens (2011) to reduce potential errors in topology associated with limited taxon sampling. Third, we used the same data and method (BEAST) to simultaneously estimate the phylogeny and divergence times and utilized three secondary calibration points (i.e., without extensively constraining the topology, the nine fossils could not be assigned to nodes). Full phylogeny and branch-length estimation methods are in the Online Appendix S2. We found that the choice of topologies had little impact on our main results (see below). For brevity, the main results presented are based on the second set of analyses.

We next estimated the history of microhabitat changes across the phylogeny of the sampled species to identify instances where similar microhabitat use has independently evolved. This step was necessary for both testing convergence (i.e., we expect convergence when multiple lineages independently colonize the same environment) and for setting up the subsequent models of morphological evolution (see next section). In these analyses we were not trying to estimate the actual number of microhabitat transitions across all frogs, given our sampling of only 167 of ~6500 described species (AmphibiaWeb 2014). Rather, we wanted to test whether microhabitat change has been frequent (relative to the number of species in our data set) and whether these changes have generated many independent evolutionary replicates of potential convergent morphological evolution. Furthermore, we wanted to test the direction of microhabitat changes (e.g., terrestrial to arboreal vs. arboreal to terrestrial) and the relative frequency of different types of changes among the sampled species. We note that undersampling species might underestimate the number of trait origins and extent of convergence, but it cannot overestimate them.

We estimated the history of microhabitat use in two ways. First, we used likelihood to estimate the relative support for each state at each node (Schluter et al. 1997). We used diversiTree version 0.9-6 (FitzJohn 2012) to choose the optimal model of discrete character evolution (here meaning the number of different possible transition rates between pairs of states). We assumed no effect of character states on diversification (Maddison et al. 2007; FitzJohn 2012), a necessary assumption given the limited species sampling. Comparisons of these models using AICc (small sample-size adjusted Akaike information criterion; Burnham and Anderson 2002) revealed the symmetric model as the best compromise between model fit and parameter number (AICc-symmetric = 358.06 <AICc-all-rates-different = 368.46 <AICc-equal-rates = 379.94). Note that this model fit was for five microhabitat states. In Appendix S5 we provide a more detailed description of model fitting when setting up the diverse OU models, which necessitate specifying microhabitat states at internal nodes. We then used the optimal model (which differed depending on the OU model; see below) to estimate likelihood support for each state at each node. We then set all nodes to the state with the highest probability and with significant support (i.e., a single state that was at least 7.39 times more likely than the next most likely state; Pagel 1999). We used these strongly supported nodes to estimate the number of origins for each microhabitat state and for labeling internal branches for OU analyses (see below).

Second, we conducted Bayesian stochastic character mapping (Nielsen 2002; Huelsenbeck et al. 2003) with phytools version 0.2-1 (Revell 2012) to better estimate the number of times each microhabitat state evolved and to examine whether transitions were more likely in one direction versus another. We generated 1000 stochastic maps and we constrained the transition-rate matrix (Q) to be symmetric (the same as the symmetric model above, which had the best AICc support). We estimated and fixed Q with likelihood (obtaining nearly the same rate estimates as in diversiTree), and we used the estimated stationary distribution as a prior on the root state. We found quantitatively similar results (e.g., mean number of transitions, shape of posterior distributions for parameter estimates) using alternative options. These options included sampling Q with Markov chain Monte Carlo (MCMC), using various prior distributions for Q when MCMC sampling, extending the burn-in to 10,000 generations (default = 1000), and sampling the...
MCMC chain less frequently than default (i.e., every 1000 generations instead of 100).

Because Bayesian stochastic character mapping simulates histories of transitions along branches (and thus can include multiple changes along a single branch), we expected that this method would estimate a larger number of changes than likelihood given high transition rates (Nielsen 2002). We did this Bayesian analysis only for the five broad microhabitat states to examine general patterns of evolution in microhabitat use. In contrast, in addition to the simplest, five-state model that we used to estimate number of microhabitat transitions, we also did the likelihood reconstructions for all other models of convergence and history, each with a different number of states, so as to properly specify the OU models (see below). See Online Appendices S4 and S5 for full details and justification of methods for characterizing the evolution of microhabitat use.

Models of Phenotypic Evolution, Adaptive Convergence, and History

Our approach to analyzing the roles of adaptive convergence and history on morphological evolution involved two tests, both utilizing OU models of adaptive evolution (Hansen 1997; Butler and King 2004). In these models, one estimates evolutionary optima in a continuous character (in our case morphology) for discrete selective environments or adaptive peaks (in our case microhabitat use), with the rate of approach to these optima controlled by a parameter $a$. These two parameters of OU models (i.e., optimum, rate) make them especially useful for the present study. First, different clades adapting to the same environment may nonetheless have different evolutionary optima in that environment due to clade-specific differences in, for example, development, response to selection, body architecture, or complex relationships between form and function (Gould and Lewontin 1979; Gould 2002; Losos 2010; Collar et al. 2014). Thus, it is possible to evaluate whether clade history is important by testing whether the optimal model incorporates clade membership only, a combination of clade membership and microhabitat, or microhabitat alone (the first two would show the impact of clade history). Second, the rate at which a lineage moves toward the optimum (the parameter $a$) can determine whether lineages have been in a selective environment (e.g., microhabitat) long enough to overcome historically different starting points (Hansen et al. 2008). In other words, a species’s morphology is the consequence of its history of adaptation to its past and current environments. How closely its morphology fits the adaptive optimum of its current environment will depend on (among other things) how long it has been in this environment and on the macroevolutionary rate at which the species approaches the inferred optimum for that environment. Thus, the nature of the OU model allows one to test various aspects of the importance of adaptive convergence (same optima for all species in the same environment; high $a$ and rapid evolution to the optimum phenotype for a given environment) versus the importance of history (different optima for different clades in the same environment for deeper historical constraints; low $a$ and more time to reach the optimum phenotype). We used these properties of the OU model to design two new complementary tests of the two types of historical limitation to convergence.

We note that a recent statistical approach also tests for convergence, but we do not use it here because it does not allow one to compare the relative importance of convergence and history. SURFACE (Ingram and Mahler 2013; Mahler et al. 2013) uses phenotypic data and phylogeny to search for OU adaptive optima instead of specifying them a priori, as in typical OU tests (Hansen 1997; Butler and King 2004; Bartoszek et al. 2012; Beaulieu et al. 2012). This approach is well suited to estimating the total number of convergent events because the same optimum may be found in different parts of a phylogeny. However, such a search algorithm does not allow one to test for possible historical differences. For example, if we know from ecological data that two distantly related species use the same microhabitat but we find that they are different morphologically, our two approaches described below can test whether this difference is due to history (whether clade history or time-for-adaptation). In contrast, SURFACE might find different adaptive optima for the two species but would not address why they are different, because the fact that species use the same microhabitat is ignored by the analysis.

Testing the importance of clade history.—We first designed a series of models (Fig. 1) in which we compared strictly clade-based models (no adaptive convergence, morphology determined by clade membership) to strictly microhabitat-based models (a single adaptive optimum for all lineages in a microhabitat; i.e., convergence). We also included models that reflected both clade and microhabitat differences, in which each clade had its own, separate optimum for each microhabitat state present in that clade. These latter models included those in which all independent origins of a given microhabitat state (e.g., 13 arboreal origins) had different optima, as well as major clade-specific models in which the same microhabitat state in different major clades had a different morphology (e.g., arboreal ranoids were different from arboreal hyloids). Models differed only in how optima were specified (i.e., $a$ and $\sigma^2$ were the same for all optima; see below for further discussion; Beaulieu et al. 2012).

We compared the following 12 models: (i) Brownian motion (BM), a model of random evolution along the tree; (ii) single-peak OU model; (iii) a different optimum for each of five major clades (Discoglossoidea, Myobatrachidae + Myobatrachidae, Ranoida, Pelobatoidea, Pipidae) with no microhabitats included; (iv) standard microhabitat model with five optima (one per microhabitat); (v) same as previous model but
Brownian motion (random)  
Single optimum  
(all frogs the same)  
Convergence only  
(5 or 6 microhabitats)  
Only clades (history)  
Clade-specific arboreal  
Clade-specific burrowing  
Clade-specific aquatic  
All origins  
by clade  
All origins  
by all three types  
All origins  
Aquatic

**FIGURE 1.** Graphical comparison of models used in the first test in this article, showing how the models are conceptually nested. Complexity of models increases from top to bottom. For example, the "all origins arboreal" model collapses into the "clade-specific arboreal" model (because multiple independent origins are assigned to a single clade in the latter model). This then collapses into the five-microhabitat model (all arboreal origins within clades are given a single, arboreal optimum), which collapses into the single-optimum OU model when all microhabitat states are given the same optimum.

with separate aquatic and semi-aquatic species, for a total of six microhabitat optima. All other models were based on model iv but with one or two changes; (vi) aquatic/semi-aquatic species assigned a different optimum based on the major clade to which they belong (with clades as in model iii); (vii) aquatic/semi-aquatic species assigned a separate optimum for each of their 11 independent origins of this microhabitat type (among the species sampled here); (viii) arboreal species divided by major clades; (ix) arboreal species divided by their 15 independent origins; (x) burrowing species divided by clades; (xi) burrowing species divided by their nine origins; and (xii) aquatic/semi-aquatic, arboreal, and burrowing all divided by major clades. We did not subdivide terrestrial nor torrent taxa because of difficulties determining their independent origins (see below).

We used PC scores as response variables to compare these 12 models. We only analyzed PC2–10, as PC1 represented overall size (see section "Results") and we excluded size because species of all sizes occur in all types of microhabitats (Moen and Wiens 2009; Van Bocxlaer et al. 2010; Moen et al. 2013). Furthermore, different microhabitats were not strongly supported as different in PC1 (OU single-optimum AICc = 1466.13 <OU five-state AICc = 1469.21) and our results were nearly identical when including PC1 in our overall 12-model comparison (Online Supplementary Table S4). We ran OU models with the R package _ouch_ ver. 2.8-4 (Butler and King 2004), assuming each PC axis was independent (Mahler et al. 2013). Given this assumption, we summed likelihoods of all PC axes to obtain the overall likelihood of each model across all traits. Full details of OU model implementation are provided in Online Appendix S5.

We compared models based on the AICc with a correction for finite sample sizes and its associated weight (the relative strength of a model with respect to all others estimated, with values ranging from 0 to 1; Burnham and Anderson 2002). We considered a model as strongly supported if it contained most of the AICc weight. However, AICc weights can be used to measure the relative support of shared factors across models. For example, if the fully convergent models (models iv and v) together contained most of the AICc weight, then regardless of which model is stronger, our results overall would support a strong role of convergence in explaining frog morphological diversity. However, if the clade-only model alone (model iii) contained most of the weight, history would have a large role in explaining frog morphology. The models that combine convergence and history (e.g., clade-specific microhabitat models vi-xii) would support an intermediate role of each factor.

**Testing the importance of historical changes in microhabitat use.**—Even if all species in a given microhabitat have a similar optimum, their current morphology might still reflect traces of adaptation to an ancestral environment (Hansen 1997; Hansen et al. 2008). Therefore, as our second test, we estimated the deviation of species’ phenotypes from their estimated adaptive optima for each microhabitat. We then compared the proportion of this deviation associated with random deviation to the proportion associated with an average systematic
To start, we define the total sum of squared deviations (TSS) around the adaptive optimum, depending on which environment species currently inhabit, as:

$$\text{TSS} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i - \theta_i)^2$$

where $x_{ij}$ is the phenotype of species $j$ in environment $i$ and $\theta_i$ is the OU adaptive optimum for environment $i$. Given that we are interested in estimating the influence of history overall across all species and microhabitats (the environments in this study), we sum across all of history overall across all species and microhabitats. Given that we are interested in estimating the influence of history overall across all species and microhabitats, we can use the harmonic mean of $n$ per regime (Sokal and Rohlf 1995). Once we have the variance due to history and that due to random factors is simply the mean-square error (MSE or $s^2$). Moreover, if we expect that each environment’s realized historical effect is similar and not of interest, we can calculate an overall variance due to history as $(\text{MS}_\text{hap} - \text{MSE})/n$, where $n$ is the number of species in each environment. In cases where $n$ is not the same for all environments (as here), we can use the harmonic mean of $n$ per regime (Sokal and Rohlf 1995). Once we have the variance due to history and that due to random factors ($s_1^2$ and $s_2^2$, respectively), we can simply compare the contribution of each to the total variance of all species from the estimated optimum value given their environment.

For this test we used optima estimated in our best-supported model from the previous test (six-microhabitat OU model; see section “Results”). We then compared the relative magnitudes of the variance due to history and that due to random factors. We did this for each PC score individually and also summed across all axes (i.e., axes 2–10).

Finally, we more explicitly tested the idea of limited time for adaptation by testing whether species that had more recently colonized their current microhabitat were farther from that microhabitat’s optimum. For this test we estimated a Spearman rank correlation of the multivariate distance of species’ values from the optimum of their current microhabitat and the natural log of the time since the most recent transition into that microhabitat. Time was log-transformed because in OU models the importance of past environments is modeled as decreasing exponentially over time (Hansen 1997). To account for phylogenetic non-independence, we conducted this analysis on phylogenetic independent contrasts (Felsenstein 1985) calculated with the R package ape version 3.1-1 (Paradis et al. 2004). We did this analysis for all microhabitats together because of low within-microhabitat sample sizes (see section “Results”), but in future studies it may be interesting to compare results across different environments to see if limited time for adaptation is important in some environments but not others.

As we mention at beginning of this section, the effect of a lag time for adaptation is closely tied to the estimated value of alpha for a given character. In particular, the OU framework already has a related concept, the phylogenetic half-life, which is the amount of time a lineage needs to move half the distance to an optimum and is calculated as $t_{1/2} = \ln(2)/\alpha$ (Hansen 1997; Hansen et al. 2008). Thus, as $\alpha$ increases it takes
less time to move toward an adaptive optimum and thus the lag time for adaptation will be shorter. This concept is useful for understanding the rate of approach to an adaptive peak in general (the phylogenetic half-life compared to the total length of the tree) or in specific cases (the half-life compared to the length of time a specific lineage has been in its current adaptive regime). Yet, we have developed the above method as a complementary approach because it allows one to compare two quantities directly—the variance among species due to time-for-adaptation versus that due to idiosyncratic differences among lineages, including unmeasured factors (Hansen 1997).

Because of this direct tie between $\alpha$ and time-for-adaptation, we note that estimating multiple $\alpha$ values (e.g., for different adaptive regimes) would strongly influence the results of this second test. In particular, if $\alpha$ is stronger in some regimes than others, the overall average deviation of species from their inferred adaptive optima (the history effect above) would vary across regimes, with some regimes potentially having a smaller historical component (higher $\alpha$) and others having a larger component (lower $\alpha$). Unfortunately, we were not able to accurately estimate multiple $\alpha$ models (a different $\alpha$ for each OU regime), as we always found unrealistic log-likelihoods (e.g., positive values in the thousands) and parameter estimates (e.g., optima three orders of magnitude larger than observed data). Running multiple searches with different starting points and likelihood search strategies did not improve the estimates, consistent with high complexity of the likelihood surfaces of these models (Beaulieu et al. 2012). Regardless of our inability to estimate these models, future research should give strong consideration to this potentially confounding factor (i.e., if multi-$\alpha$ models have more statistical support than models with a global $\alpha$).

**RESULTS**

**General Patterns of Microhabitat Use and Morphology**

All sites had arboreal, aquatic/semi-aquatic, and terrestrial species (Supplementary Fig. S2). Many sites also had burrowing and/or torrent species (Supplementary Fig. S2). All species from these locations fit into one of these five microhabitat categories (Online Supplementary Table S1).

Phylogenetic PCA of the morphological data allowed us to visualize morphological variation among species and showed that species cluster in morphological PC space according to their microhabitat use (Fig. 2). PC1 represented overall size and accounted for much variation (85.5%; Table 1), as expected. PC2 primarily represented the size of toe and fingertips (Table 1). PC3 largely represented foot webbing size. Finally, PC4 primarily showed negative weights for head and leg length, contrasted with a large positive weight for metatarsal tubercle size (Table 1). These results were robust across phylogenetic topologies and branch lengths estimated three different ways (see above). For brevity, all results shown here are based on the maximum-likelihood topology of Pyron and Wiens (2011) with branch lengths in units of time based on fossil calibrations (Online Appendices S2 and S3).

**History of Microhabitat Use in Frogs**

Terrestriality appears to be the ancestral microhabitat for frogs (Online Fig. S1). From this state, likelihood analyses suggest that frogs have independently evolved arboreal microhabitat usage a minimum of 13 times, burrowing 9 times, and aquatic/semi-aquatic microhabitat 11 times (Fig. 3), at least among the species sampled here. There have been relatively few clear reversions to terrestriality. The number of origins of the torrent-inhabiting state was ambiguous, with two clear independent origins but possibly many more (Fig. 3). Bayesian stochastic character mapping gave qualitatively similar results, but with many more inferred changes than likelihood (Supplementary Fig. S3). In particular, stochastic mapping inferred a mean of 78.0 independent transitions among all five microhabitat states (95% credibility interval = 66–92). As in likelihood analyses, most states originated repeatedly from terrestriality (Table 2).

Examining patterns within clades and geographic regions shows repeated origins of each microhabitat type across regions, providing many independent replicates for testing the impact of history and convergence on morphology. For example, there are separate origins of arboreal frogs in Africa (e.g., Hyperoliiidae), Asia (Microhylidae, Rhacophoridae), and Madagascar (Mantellidae). Sites in South America have many arboreal Hylidae, but also have arboreal species that evolved independently in other clades (e.g., Centrolenidae, Craugastoridae, Hemiphractidae). Similarly, there were multiple geographic origins of aquatic frogs, in Africa (Pyxicephalidae), Asia (Dicroglossidae), and Madagascar (Mantellidae), and South America (Hyline Hylidae [Pseudis, Lynapsis], Leptodactylidae). Intriguingly, some geographically isolated regions contain multiple origins of different ecomorphs in the same clade (Mantellidae on Madagascar, pelodyridine Hylidae in Australia). In other cases, a single origin of a microhabitat specialist appears to have spread across multiple continents (e.g., arboreal Hylidae in the New World, Asia, and Australia).

**Tests of Adaptive Convergence and History**

**Clade history.**—In our first test of the relative importance of adaptive convergence and history on frog morphology, the best-fitting model was one dominated by adaptive convergence, in which each microhabitat had a single morphological optimum independent of clade, including separate optima for aquatic, arboreal, burrowing, semi-aquatic, terrestrial, and...
Figure 2. Principal components scores for species’ morphology, plotted for PC2–4, which show the greatest amount of variation beyond general size among the 167 species (Table 1). Loadings of raw variables on PC axes are indicated with their labels. Colors indicate microhabitat use of each species, while symbol shape indicates from which site the species comes. Species that do not occur in a specific assemblage (e.g., *Laliostoma labrosum*) were plotted with the symbol of the geographically closest assemblage (e.g., Ranomafana, Madagascar). Inferred evolutionary optima from OU analyses are indicated by large circles. The figure illustrates how extant species are slightly offset from the optima. This is an effect of history, because most species can be envisioned as evolving from a terrestrial ancestral value (central optimum) toward the optimum for their current microhabitat. Values for extant species that are intermediate between the optimal value for their current microhabitat and the optimum for the ancestral, terrestrial microhabitat are considered to reflect insufficient time to adapt to the current microhabitat, and thus the imprint of history (Hansen 1997; Hansen et al. 2008). The torrent optimum in the upper plot has nearly the same value as the arboreal optimum and thus is almost entirely obscured by it. Photos represent examples of each ecomorph and are color-coded as such (from top to bottom): *Phyllomedusa tomopterna*, *Platyplectrum ornatum*, *Babina pleuraden*, *Litoria tornieri*, and *Amolops tuberodepressus*. All photos by DSM.

Torrent-dwelling species (Table 3). AICc weights showed that no models with different optima in different clades had strong statistical support (Table 3). Thus, frog species were more similar (on average) to other species in the same microhabitat than to related species in different microhabitats, and clade membership had no detectable impact on the similarity of species within these categories.
2016 MOEN ET AL.—TESTING CONVERGENCE IN FROGS 155

Table 1. Results of phylogenetic PCA

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
<th>PC8</th>
<th>PC9</th>
<th>PC10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvectors</td>
<td>8.549</td>
<td>0.653</td>
<td>0.384</td>
<td>0.187</td>
<td>0.081</td>
<td>0.054</td>
<td>0.034</td>
<td>0.029</td>
<td>0.017</td>
<td>0.011</td>
</tr>
<tr>
<td>Proportion of total variation</td>
<td>0.855</td>
<td>0.165</td>
<td>0.038</td>
<td>0.019</td>
<td>0.008</td>
<td>0.005</td>
<td>0.003</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Notes: The phylogeny used for this analysis was the time-calibrated tree from BEAST with fossil calibrations and constrained topology. The estimated adaptive optima (θ) of our best-fit model revealed the morphology of microhabitat specialists and the differences among them (Fig. 2; Online Supplementary Table S5). Arboreal species had a strongly negative optimum for PC2, indicating that their finger and toe tips are enlarged into pads, which increase clinging performance on vertical surfaces (Emerson 1991; Moen et al. 2013). Aquatic and semi-aquatic species differed from other ecomorphs in having large optima for PC3, an axis primarily related to foot webbing; large foot webbing appears to increase swimming performance (Nauwelaerts et al. 2005). Aquatic species primarily differed from semi-aquatic species by having a more extreme optimum on PC5 (i.e., more foot webbing and more pointed toe tips). Burrowing frogs had a large, positive optimum for PC4, which mainly reflects short head length, short legs, and large metatarsal tubercle size. The latter two characteristics are known to improve burrowing performance (Emerson 1976). Torrent species had both large finger and toe tips (strongly negative PC2 optimum) and extensive foot webbing (high optimum value for PC3), consistent with the fact that they inhabit fast-flowing streams (where strong swimming ability seems to be necessary) and cling to rocks and vegetation in or near streams (seemingly requiring large digital pads). Terrestrial species were characterized by intermediate values for all PC optima.

The estimated adaptive optima (θ) of our best-fit model revealed the morphology of microhabitat specialists and the differences among them (Fig. 2; Online Supplementary Table S5). Arboreal species had a strongly negative optimum for PC2, indicating that their finger and toe tips are enlarged into pads, which increase clinging performance on vertical surfaces (Emerson 1991; Moen et al. 2013). Aquatic and semi-aquatic species differed from other ecomorphs in having large optima for PC3, an axis primarily related to foot webbing; large foot webbing appears to increase swimming performance (Nauwelaerts et al. 2005). Aquatic species primarily differed from semi-aquatic species by having a more extreme optimum on PC5 (i.e., more foot webbing and more pointed toe tips). Burrowing frogs had a large, positive optimum for PC4, which mainly reflects short head length, short legs, and large metatarsal tubercle size. The latter two characteristics are known to improve burrowing performance (Emerson 1976). Torrent species had both large finger and toe tips (strongly negative PC2 optimum) and extensive foot webbing (high optimum value for PC3), consistent with the fact that they inhabit fast-flowing streams (where strong swimming ability seems to be necessary) and cling to rocks and vegetation in or near streams (seemingly requiring large digital pads). Terrestrial species were characterized by intermediate values for all PC optima.

Testing the effects of historical microhabitat change.—We found that most frog species were not at the phenotypic optimum of their current microhabitat (Fig. 2). Our second test of adaptive convergence versus history showed that (on average) species’ values were displaced from the optima of their current microhabitats, with the historical variance higher than the variance due to random differences among species (Table 4). In the first three PC dimensions this displacement was toward the ancestral, terrestrial optimum (Fig. 2). Specifically, these time-for-adaptation effects were strongest in PCs 2, 3, and 5 (PCs which together explain 77% of the overall size-independent variation in morphology; Table 1). In contrast, more variation in many of the higher PCs (i.e., PC4, PC7, PC9) was due largely to random error around the optima rather than systematic deviation (Table 4). In the case of PC4, an axis that distinguishes burrowing species, the estimated value of alpha was higher than for all other PC axes (Online Supplementary Table S5), suggesting a rapid approach to the optimum once lineages become burrowing and/or once burrowing lineages transition to another microhabitat. Finally, lineages that have been in their current microhabitat less time were farther from the optimum (Fig. 4; $r_2 = -0.229, P = 0.012$), as expected if adaptation to a new environment takes considerable time.

Discussion

In this study, we demonstrate a two-part phylogenetic approach for testing the relative importance of adaptive convergence and history on species’ phenotypes, using a global-scale dataset for frogs. We found strong overall convergence, even at the deepest time scales. Specifically, the six microhabitat-related ecomorphs were similar around the world, no matter where or how many times they evolved. Yet, we also found that the phenotypes of species were generally not at the estimated phenotypic optimum for their microhabitat, and instead showed an imprint of history associated with a systematic bias toward the ancestral, terrestrial phenotype. Hence, patterns of morphological variation in this major vertebrate clade are explained by rampant but incomplete convergence. Importantly, our tests show that the imprint of history here is related to past ecology, and not differences in morphological optima between clades.
Figure 3. Phylogeny and microhabitat use in frogs. The topology is from Pyron and Wiens (2011) and branch lengths (in millions of years) were estimated for this article using the Bayesian uncorrelated lognormal approach in BEAST (Drummond and Rambaut 2007). Branch colors refer to ancestral states estimated by unordered, symmetric-rates maximum-likelihood (Schluter et al. 1997) in R with the package diversitree (Fitzjohn 2012). Dotted branches are those whose reconstruction was ambiguous (no single state was at least 3.9 times more likely than the next most likely state; Pagel 1999), their colors represent the most likely state. We label the clades we used in our clade-specific analyses (Hyloidea + Myobatrachidae, Pelobatoidea, and Ranoidea are identified), while Discoglossoidea and Pipidae are represented by the top terminal branch and second-from-top terminal branch (the species Discoglossus jeanneae and Xenopus witti, respectively; see Fig. S1). Finally, circular node labels indicate other important clades discussed in this article: (1) Microhylidae, (2) Ranidae, (3) Rhacophoridae, (4) Mantellidae, (5) Bufonidae, (6) Hylidae, and (7) Pelodryadinae. See Online Supplementary Figure S1 for the species name associated with each tip.
Our results are significant in showing repeated convergence at a large spatial scale (global) and temporal scale (>150 myr), involving all sampled species in all 10 communities (i.e., all species belong to only six distinguishable ecomorph categories). This community-wide convergence has thus far only been found in relatively isolated settings, such as nearby sets of islands (Losos et al. 1998; Mahler et al. 2013) and lakes (Seehausen 2006; Muschick et al. 2012). Furthermore, the clades in these examples are much younger than the focal clade of the present study (Greater Antillean Anolis ~40 myr [Blankers et al. 2013]; East African rift lake cichlids 2.3 myr [Friedman et al. 2013]). At the same time, some studies have found strong convergence globally, but only involving some species in each region, rather than most species considered (e.g., limb-reduced ecomorphs of lizards [Wiens et al. 2006]; convergence between taxa in Australia and other regions [Melville et al. 2006; Grünbler and Rabosky 2014]).

Our results also show that despite this widespread adaptive convergence, there is still a large time-for-adaptation effect on many morphological variables. Specifically, we found that species’ morphologies are often offset from the morphological optimum for their microhabitat, primarily toward the ancestral optimum (terrestriality; Fig. 2). This result might seem to overturn the importance of adaptive convergence found in the first set of analyses. However, these analyses show an effect of the past history of adaptation to other microhabitats, rather than evidence of clade-specific historical effects on morphology.

What might explain this lag time in adaptation? In frogs, lineages have moved frequently between microhabitats over their evolutionary history. As a consequence, many species may not have been in their current microhabitat long enough to reach its estimated morphological optimum. Indeed, most microhabitat transitions are relatively recent (i.e., 77.5% are less than 80-my-old, half the length of the entire phylogeny), and many of these ages may be overestimates because our taxon sampling is incomplete. For example, the long history of aquatic/semi-aquatic microhabitat use inferred for Discoglossoidea and Pipidae (Fig. 3) may be an artifact of missing the terrestrial and burrowing members of these two clades (Alytidae and Rhinophrynidae, respectively; Duellman and Trueb 1986; Pyron and Wiens 2011). Finally, although this may be the first study to quantitatively examine differences between species’ phenotypes and estimated adaptive optima, visual examination of figures in other studies (Ingram and Mahler 2013; Mahler et al. 2013) suggests that this pattern may be common.

Paradoxically, the effects of deep history on morphology may explain the exceptional convergence that we document across frogs. For example, all burrowing frogs may be fundamentally similar because they still retain the basic morphology shared by all frogs (as opposed to resembling burrowing caecilians or lizards). In other words, given shared developmental patterns, genomic architecture, and body form across all frogs (Duellman and Trueb 1986), there may be limited ways to respond to selection and this may lead to frequent convergence (Wake 1991; Brakefield 2006) instead of the common expectation that these factors would prevent convergence (Gould 1989, 2002). Thus, the evolutionary optimum is still a relative concept (Losos 2011). An important question for future studies is why in frogs this temporal zone of convergence seems to extend across nearly the entire group of ~6500 species (AmphibiaWeb 2014) and over >150 myr, whereas it often seems more localized in other groups (Martin and Wainwright 2013).

Our analysis also elucidates the history of microhabitat use and ecomorphs across frogs, with surprising results. Many classic studies of convergence show that an ecomorph arises in a region only once, with the idea that competition and niche filling would prevent multiple origins of the same ecomorph in the same place (e.g., Anolis ecomorphs on islands; Losos et al. 1998; snakelike ecomorphs in squamate reptiles; Wiens et al. 2006). Surprisingly, our analysis shows that similar ecomorphs can have multiple origins in each region and co-occur in sympathy (e.g., multiple origins of arboreal ecomorphs in South America; see also Wiens 2011), even if they possibly arose in allopatry within the region. More remarkably, these separately evolved ecomorphs are not distinguishable from each other based on our tests, despite their co-occurrence in our sampled communities. These results provide an intriguing counterpoint to the long-standing idea of saturation of communities with ecologically similar species (MacArthur and Wilson 1963), and possibly the role of limited ecological opportunity in constraining ecomorph evolution (Schluter 2000).

Our study also has some limitations. First, given the large number of separate origins of each microhabitat type, the number of possible models to explore is vast and our analyses may have lumped some distinctive ecomorphs. Nevertheless, our analyses clearly favor a very limited number of ecomorph categories. We also acknowledge that we have included only a small fraction of the total diversity of frogs. We included most major groups of frogs typically found at sites around the world, but not all radiations were included. For example, Kaloula has diversified ecologically and morphologically in the

<table>
<thead>
<tr>
<th>Table 2. SIMMAP-inferred transitions between states over the history of the anuran species examined in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arboreal Burrowing Semi-aquatic Terrestrial Torrent</td>
</tr>
<tr>
<td>Arboreal – 11 (0.3) 0.5 (0.2) 12.0 (6.21) 0.0 (0.0)</td>
</tr>
<tr>
<td>Burrowing 0.2 (0.1) – 3.0 (1.6) 3.1 (0.8) 0.0 (0.0)</td>
</tr>
<tr>
<td>Semi-aquatic 0.1 (0.1) 2.0 (0.5) – 4.4 (1.9) 5.7 (2.10)</td>
</tr>
<tr>
<td>Terrestrial 19.6 (14.28) 0.5 (6.15) 12.4 (8.17) – 0.0 (0.0)</td>
</tr>
<tr>
<td>Torrent 0.0 (0.0) 0.0 (0.0) 3.4 (0.7) 0.0 (0.0) –</td>
</tr>
</tbody>
</table>

Notes: Cell values are the mean number of transitions (the average number across 1000 draws from the posterior distribution of transitions) with their 95% credibility interval (the bounds of the central 95% of the posterior distribution of transitions). All changes refer to changes from the row state to the column state. The number across 1000 draws from the posterior distribution of transitions). All changes refer to changes from the row state to the column state. The number across 1000 draws from the posterior distribution of transitions). All changes refer to changes from the row state to the column state.
found rampant convergence in frogs over unexpectedly large geographic (global) and temporal scales (over 150 myr). Yet, this approach also showed that at more recent time scales most species have morphology that lags behind the morphological optimum for their
current microhabitat due to the imprint of their past ecology (but not their clade history). Thus, widespread but incomplete convergence explains patterns of morphological variation in one of the major clades of terrestrial vertebrates. More generally, using our approach in other groups could reveal the distinct roles that adaptive convergence, clade history, and time may play in explaining phenotypic diversity across the Tree of Life.

SUPPLEMENTARY MATERIAL

Data and computer code in R are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.8vv63.

FUNDING

Funding was provided by a Smithsonian Predoctoral Fellowship to D.S.M., an NSF Doctoral Dissertation Improvement Grant to D.S.M. and J.J.W. (DEB-110704), the CNRS, and grant ECOEVOBIO-CHEX2011 from the French National Research Agency (ANR) awarded to H.M.

ACKNOWLEDGMENTS

We greatly thank the many museum curators and collection managers who generously gave access to specimens and advice for this study, including A. Resetar and K. Kelly (The Field Museum); S. Rogers and S. Busack (Carnegie Museum); J. Losos and J. Rosado (Harvard Museum of Comparative Zoology); D. Blackburn, R. Drewes, and J. Vindum (California Academy of Sciences); R. Brown and A. Campbell (Univ. Kansas Museum of Natural History); and R. and M. Heyer, J. Jacobs, R. McDermid, K. de Queiroz, K. Tighe, R. Wilson, A. Wynn, and G. Zug (US National Museum, Smithsonian Institution). J. Mendelson and R. Heyer kindly shared unpublished ecological observations. This work was improved by comments from F. Anderson, R. Glor, two anonymous reviewers, and discussions with F. Condamine, J. Rolland, and the Behavior and Morphology discussion group at the University of Massachusetts, Amherst.

REFERENCES


