

Abstract.—Striking evolutionary convergence can lead to similar sets of species in different locations, such as in cichlid fishes and *Anolis* lizards, and suggests that evolution can be repeatable and predictable across clades. Yet most examples of convergence involve relatively small temporal and/or spatial scales. Some authors have speculated that at larger scales (e.g. across continents), differing evolutionary histories will prevent convergence. However, few studies have compared the contrasting roles of convergence and history, and none have done so at large scales. Here we develop a two-part approach to test the scale over which convergence can occur, comparing the relative importance of convergence and history in macroevolution using phylogenetic models of adaptive evolution. We apply this approach to data from morphology, ecology, and phylogeny from 167 species of anuran amphibians (frogs) from ten local sites across the world, spanning ~160 million years of evolution. Mapping ecology on the phylogeny revealed that similar microhabitat specialists (e.g. aquatic, arboreal) have evolved repeatedly across clades and regions, producing many evolutionary replicates for testing for morphological convergence. By comparing morphological optima for clades and microhabitat types (our first test), we find that convergence associated with microhabitat use dominates frog morphological evolution, producing recurrent ecomorphs that together encompass all sampled species in each community in each region. However, our second test, which examines whether and how much species differ from their inferred optima, shows that convergence is incomplete: that is, phenotypes of most species are still somewhat distant from the estimated optimum for each microhabitat, seemingly because of insufficient time for more complete adaptation (an effect of history). Yet these effects of history are related to past ecologies, and not clade membership. Overall, our study elucidates the dominant

drivers of morphological evolution across a major vertebrate clade and shows that evolution can be repeatable at much greater temporal and spatial scales than commonly thought. It also provides an analytical framework for testing other potential examples of large-scale convergence.

Keywords: Anura; amphibians; communities; constraint; convergence; ecomorphology; macroevolution

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 radiations 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 lead to different morphologies in different clades (Gould and Lewontin 1979; Gould 1989, 2002). The repeated evolution of entire communities of diverse ecomorphs through convergence has been proposed at large scales (e.g. placental and marsupial mammals, Mediterranean-climate plants; Cody and Mooney 1978; Futuyma 2013). Yet rigorous testing has found this pattern mostly in younger clades (Schluter and McPhail 1993; Clabaut et al. 2007; Muschick et al. 2012) and/or in smaller geographic regions (Losos et al. 1998; Mahler et al. 2013). For example, *Anolis* lizards show spectacular convergence in the Greater Antilles (Losos et al. 1998; Mahler et al. 2013), but not in the rest of the Neotropics, where most species occur (Irschick et al. 1997). Similarly, haplochromine cichlids show remarkable convergence among East African rift lakes, but not throughout Africa (Martin and Wainwright 2013).

Thus, an important but unresolved issue in evolutionary biology is whether convergent phenotypic evolution can still dominate at deeper temporal scales. For example, do evolutionary patterns become more clade-specific at deeper scales,

with historical differences among clades dominating their phenotypic evolution (Stayton 2006; Losos 2011)? Such historical differences 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 common selective problems and thus prevent convergence. On the other hand, there might instead be more convergence at larger time scales, as it may take considerable time to adapt to a new environment (Hansen 1997). In particular, many factors that limit adaptation may diminish with time (e.g. limited genetic variation, pleiotropy, correlated selection; Hansen 1997; Futuyma 2010). Taken one step further, given a sufficiently long timescale, can such convergence dominate entire communities across regions (Cody and Mooney 1978; Melville et al. 2006; Losos 2010), with most species in each community belonging to convergent ecomorphs? Or are most communities instead dominated by clade-specific phenotypes?

Few studies have quantitatively tested convergent 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 across scales (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 selection or biomechanical 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 (towards 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 semi-aquatic species found in many local assemblages (e.g. Inger et al. 1984; Duellman 2005). However, anurans in different regions often belong to different clades (e.g. Hyloidea dominates the New World and Ranoidea 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 ten 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 paper, 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 ten 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 ten 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 (determined from the complete list of species for each site; online Table S1; doi:10.5061/dryad.8vv63) and we measured almost all genera at each site. Furthermore, morphological and habitat-use diversity within genera was low relative to diversity across genera (online Appendix S1). In addition, we supplemented our community-focused sampling by measuring seven additional species found in the same general region as some of the sampled communities (online Table S1). We did this to include additional microhabitat types in each biogeographic region that are often more geographically restricted, particularly burrowing.

The sites we selected were: (i) Australasian region: Middle Point, Northern Territory, Australia (14 species measured of 17 species known for the site; Tyler and Knight 2009); (ii) Africa: Bwindi Impenetrable Forest, Uganda (14 of 28; Drewes and Vindum 1994); (iii) Madagascar: Ranomafana National Park, Madagascar (22 of 40; Andreone 1994); (iv) South Asia: Ponmudi, Kerala, India (11 of 26; Inger et al. 1984);

(v) Southeast Asia: Nanga Tekalit, Borneo, Malaysia (21 of 54; Field Museum collection records); (vi) Palearctic: Cádiz Province, Spain (7 of 7; Carnegie Museum collection records); (vii) Nearctic: southeastern Virginia, United States of America (14 of 17; Smithsonian Institution National Museum of Natural History [USNM] collection records); (viii) Middle America: Purulhá, Baja Verapaz, Guatemala (9 of 16; Natural History Museum of the University of Kansas collection records); (ix) South America (Amazonia): Explorer's Inn, Río Tambopata, Perú (26 of 86; USNM records); and (x) South America (Atlantic Forest): Boracéia, São Paulo State, Brazil (21 of 63; Heyer et al. 1990).

Morphology

To reduce potential differences among conspecific individuals due to ontogeny and sexual dimorphism, we focused on measuring adult males, which (in our experience) are better represented in museum collections than females. From each specimen (online Table S2) we first measured body length (snouth-to-vent length), forelimb length, hindlimb length, head length, and head width using calipers (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 tibiofibula) to calculate hindlimb muscle mass. We chose these ten 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 Table S3) 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 (Jolicouer 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 (e.g. Bossuyt and Milinkovitch 2000; Young et al. 2005; Bossuyt et al. 2006): arboreal, aquatic/semi-aquatic, terrestrial, burrowing, and torrent (i.e. dwelling in high-gradient streams). For some analyses, we treated aquatic and semi-aquatic habitat use as separate categories (see Results). We categorized species based on adult activity outside the breeding season. In most cases, literature sources also categorized these species using these same categories. A full description and justification of these categories is given in the

online Appendix S1. Data on microhabitat use and supporting references are listed in online Table S1. Microhabitats were treated as character states in subsequent analyses.

Phylogeny

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 (three mitochondrial, nine 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 dataset 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.

History of Changes in Microhabitat Use

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 ~6,500 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 ($AICc_{\text{symmetric}} = 358.06 < AICc_{\text{all-rates-different}} = 368.46 < AICc_{\text{equal-rates}} = 378.94$). 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 this 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 1,000 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. Using alternative options, we found quantitatively similar results (e.g. mean number of transitions, shape of posterior distributions for parameter estimates). These alternative options included sampling Q with MCMC, using various prior distributions for Q when MCMC sampling, extending the burnin to 10,000

generations (default = 1,000), and sampling the MCMC chain less frequently than default (i.e. every 1,000 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 in order 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 Ornstein-Uhlenbeck 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 Ornstein-Uhlenbeck (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 α . 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 α) 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' 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 α 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 α 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 (e.g. 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 the 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 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. α and σ^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, Hyloidea + Myobatrachidae, Ranoidea, Pelobatoidea, Pipoidea) with no microhabitats included; (iv) standard microhabitat model with five optima (one per microhabitat); (v) same as previous model but 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 13 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 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 $AICc = 1466.13 < OU\ 5\text{-state}\ AICc = 1469.21$) and our results were nearly identical when including PC1 in our overall 12-model comparison (online 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 Akaike Information Criterion with a correction for finite sample sizes (AICc) 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 deviation towards the ancestral, terrestrial optimum (Fig. 3; online Fig. S1).

To start, we define the total sum of squared deviations (TSS) around the adaptive optimum, depending on which environment species currently inhabit, as:

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

where x_{ij} is the phenotype of species j in environment i and θ_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 k microhabitats. However, in principle, this could be done separately for each microhabitat category.

We can break down this sum of squares into three components:

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

$$= \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^k n_i (\bar{x}_i - \theta_i)^2 + 2 \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(\bar{x}_i - \theta_i) \quad (3)$$

where \bar{x}_i , is the mean of species' values for each environment, based on which environment the species currently inhabit. The first term of this sum represents an overall measure of phenotypic variation across species from the same environment, while the second term represents the distance between species' means and the evolutionary optima (i.e. systematic deviation from the optima). The latter is ostensibly explained by the time needed to adapt to successive new environments, therefore reflecting an effect of history: if species have (on average) spent little time in their current environment or if the rate of adaptation (α) is low, their phenotypes will not be near the optimum for their current environment and this sum of squares will be large. This will be clearest if a single ancestral environment dominates the history of a group, such that most species' values will be biased in the direction of that single ancestral environment (see Results). The variation expressed in the first

term, in contrast, is a combination of random variation due to drift in the OU model, differences across species in terms of the time spent in their current habitat (which can be considered somewhat random), and measurement error (Hansen and Bartoszek 2012) (for brevity we refer to this as a random term, as compared to the aforementioned historical term). The final term represents the covariance between these first two differences and cancels to zero upon rearrangement of its sums.

We can thus use these sums of squares to estimate the portion of the species' deviance from the optimum of their current environment due to random factors (SSE; first term of equation 3) and the portion due to history (SSH; second term in equation 3). The sums of squares can be divided by their respective degrees of freedom (number of environments – 1 for SSH, number of species – number of environments for SSE) to obtain mean squares (Sokal and Rohlf 1995), which form the basis for the variances. The estimate of variance 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 $(MS_{\text{hist}} - \text{MSE}) / n$, 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_h^2 and s^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 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 on 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 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 α 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 of 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 α and time-for-adaptation, we note that estimating multiple α values (e.g. for different adaptive regimes) would strongly influence the results of this second test. In particular, if α 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 α) and others having a larger component (lower α). Unfortunately, we were not able to accurately estimate multiple- α models (a different α for each OU regime), as we always found unrealistic log-likelihoods (e.g. positive values in the thousands) and parameter estimates (e.g. optima 3 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- α models have more statistical support than models with a global α).

RESULTS

General Patterns of Microhabitat Use and Morphology

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

Phylogenetic principal components (PC) analysis 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, 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 nine 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 (e.g. 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. Hyperoliidae), 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 (Ptychadenidae, Pyxicephalidae), Asia (Dicroglossidae, Ranidae), Australia (pelodyadine Hylidae), Madagascar (Mantellidae), and South America (hyline Hylidae [*Pseudis*, *Lysapsus*], Leptodactylidae). Intriguingly, some geographically isolated regions contain multiple origins of different ecomorphs in the same clade (Mantellidae on Madagascar, pelodyadine 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, Eurasia, and Australia).

Tests of Adaptive Convergence and History

Clade history.—In our first test of the relative importance of 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 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.

The estimated adaptive optima (θ) of our best-fit model revealed the morphology of microhabitat specialists and the differences among them (Fig. 2; online 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 PC3 (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 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; $r_s = -0.229$; $P = 0.012$), 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 towards the ancestral, terrestrial phenotype. Hence, patterns of morphological variation in a 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.

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 ten 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; Grundler 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 towards 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 microhabitat, 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 myr 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 Pipoidea (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 (e.g. 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 ~6,500 species (AmphibiaWeb 2014) and over >150 myr, whereas it often seems more localized in other groups (e.g. 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 sympatry (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 (e.g. 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 Philippines (Blackburn et al. 2013), but are represented by only one species here. Other clades appear to have ecologically diversified in New Guinea (Microhylidae; Menzies 2006), the Solomon Islands (Dicroglossidae: Platymantinae; Menzies 2006; Brown et al. 2015), and the Caribbean (*Eleutherodactylus*; Hedges 1989), though explicit analyses are lacking. Thus, some ecological and morphological diversity was doubtless missed. However, our broad sampling suggests that most species of frogs should fit within our general categories.

In summary, we develop a two-part phylogenetic approach to test the roles of adaptive convergence and history in morphological evolution. We then apply this approach to a new dataset for frogs. We find rampant convergence in frogs over unexpectedly large geographic (global) and temporal scales (over 150 million years). Yet, this approach also shows that at more recent time scales most species have morphology that lags behind the "optimal" morphology 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

Supplementary material, including data files and/or online-only appendices, can be found in the Dryad data repository (doi:10.5061/dryad.8vv63).

FUNDING

Funding was provided by a Smithsonian Predoctoral Fellowship to DSM, an NSF Doctoral Dissertation Improvement Grant to DSM and JJW (DEB-1110704), the CNRS, and grant ECOEVOBIO-CHEX2011 from the French National Research Agency (ANR) awarded to HM.

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 (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. McDiarmid, K. de Queiroz, K. Tighe, R. Wilson, A. Wynn, and G. Zug (U.S. 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

Adams D.C. 2014. A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution* 68:2675–2688.

- AmphibiaWeb. 2014. AmphibiaWeb: Information on amphibian biology and conservation. (<http://amphibiaweb.org>). Accessed 12 May 2014.
- Andreone F. 1994. The amphibians of Ranomafana rain forest, Madagascar - preliminary community analysis and conservation considerations. *Oryx* 28:207–214.
- Arnold S.J., Pfrender M.E., Jones A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112–113:9–32.
- Bartoszek K., Pienaar J., Mostad P., Andersson S., Hansen T.F. 2012. A phylogenetic comparative method for studying multivariate adaptation. *J. Theor. Biol.* 314:204–215.
- Beaulieu J.M., Jhvueng D.C., Boettiger C., O'Meara B.C. 2012. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* 66:2369–2383.
- Blackburn D.C., Siler C.D., Diesmos A.C., McGuire J.A., Cannatella D.C., Brown R.M. 2013. An adaptive radiation of frogs in a southeast Asian island archipelago. *Evolution* 67:2631–2646.
- Blankers T., Townsend T.M., Pepe K., Reeder T.W., Wiens J.J. 2013. Contrasting global-scale evolutionary radiations: phylogeny, diversification, and morphological evolution in the major clades of iguanian lizards. *Biol. J. Linn. Soc.* 108:127–143.
- Bossuyt F., Brown R.M., Hillis D.M., Cannatella D.C., Milinkovitch M.C. 2006. Phylogeny and biogeography of a cosmopolitan frog radiation: late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. *Syst. Biol.* 55:579–594.

- Bossuyt F., Milinkovitch M.C. 2000. Convergent adaptive radiations in Madagascan and Asian ranid frogs reveal covariation between larval and adult traits. *Proc. Natl. Acad. Sci. U.S.A.* 97:6585–6590.
- Brakefield P.M. 2006. Evo-devo and constraints on selection. *Trends Ecol. Evol.* 21:362–368.
- Brown R.M., Siler C.D., Richards S.J., Diesmos A.C., Cannatella D.C. 2015. Multilocus phylogeny and a new classification for Southeast Asian and Melanesian forest frogs (family Ceratobatrachidae). *Zool. J. Linn. Soc.* 174:130-168.
- Burnham K.P., Anderson D.R. 2002. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. New York (NY) Springer-Verlag.
- Butler M.A., King A.A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164:683–695.
- Clabaut C., Bunje P.M.E., Salzburger W., Meyer A. 2007. Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. *Evolution* 61:560–578.
- Cody M.L., Mooney H.A. 1978. Convergence versus nonconvergence in Mediterranean climate ecosystems. *Ann. Rev. Ecol. Syst.* 9:265–321.
- Collar D.C., Reece J.S., Alfaro M.E., Wainwright P.C., Mehta R.S. 2014. Imperfect morphological convergence: variable changes in cranial structures underlie transitions to durophagy in moray eels. *Am. Nat.* 183:E168–184.
- Drewes R.C., Vindum J.V. 1994. Amphibians of the Impenetrable Forest, southwest Uganda. *J. Afr. Zool.* 108:55–70.
- Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214–221.

- Duellman W.E. 2005. *Cusco Amazónico*. Ithaca (NY): Comstock Publishing Associates, Cornell University Press.
- Duellman W.E., Trueb L. 1986. *Biology of Amphibians*. Baltimore (MD): Johns Hopkins University Press.
- Emerson S.B. 1976. Burrowing in frogs. *J. Morphol.* 149:437–458.
- Emerson S.B. 1991. The ecomorphology of Bornean tree frogs (family Rhacophoridae). *Zool. J. Linn. Soc.* 101:337–357.
- Felsenstein J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125:1–15.
- FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 33:1084–1092.
- Friedman M., Keck B.P., Dornburg A., Eytan R.I., Martin C.H., Hulsey C.D., Wainwright P.C., Near T.J. 2013. Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20131733.
- Futuyma D.J. 2010. Evolutionary constraint and ecological consequences. *Evolution* 64:1865–1884.
- Futuyma D.J. 2013. *Evolution*. Sunderland (MA): Sinauer Associates.
- Gould S.J. 1989. *Wonderful Life*. New York (NY): Norton.
- Gould S.J. 2002. *The Structure of Evolutionary Theory*. Cambridge (MA): Belknap Press, Harvard University Press.
- Gould S.J., Lewontin R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B Biol. Sci.* 205:581–598.

- Grundler M.C., Rabosky D.L. 2014. Trophic divergence despite morphological convergence in a continental radiation of snakes. *Proc. R. Soc. Lond. B Biol. Sci.* 281:20140413.
- Hansen T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51:1341–1351.
- Hansen T.F. 2012. Adaptive landscapes and macroevolutionary dynamics. In: Svensson E., Calsbeek R. editors. *The Adaptive Landscape in Evolutionary Biology*. London: Oxford University Press, p. 205–226.
- Hansen T.F., Bartoszek K. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. *Syst. Biol.* 61:413–425.
- Hansen T.F., Pienaar J., Orzack S.H. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62:1965–1977.
- Hedges S.B. 1989. An island radiation: allozyme evolution in Jamaican frogs of the genus *Eleutherodactylus* (Leptodactylidae). *Carib. J. Sci.* 25:123–147.
- Heyer W.R., Rand A.S., Gonçalves da Cruz C.A., Peixoto O.L., Nelson C.E. 1990. *Frogs of Boracéia*. *Arq. Zool.* 31:231–410.
- Holt B.G., Lessard J.P., Borregaard M.K., Fritz S.A., Araujo M.B., Dimitrov D., Fabre P.H., Graham C.H., Graves G.R., Jonsson K.A., *et al.* 2013. An update of Wallace's zoogeographic regions of the world. *Science* 339:74–78.
- Huelsenbeck J.P., Nielsen R., Bollback J.P. 2003. Stochastic mapping of morphological characters. *Syst. Biol.* 52:131–158.
- Inger R.F., Shaffer H.B., Koshy M., Bakde R. 1984. A report on a collection of amphibians and reptiles from the Ponmudi, Kerala, South India. *J. Bombay Nat. Hist. Soc.* 81:406–427, 551–570.

- Ingram T., Mahler D.L. 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods Ecol. Evol.* 4:416–425.
- Irschick D.J., Vitt L.J., Zani P.A., Losos J.B. 1997. A comparison of evolutionary radiations in mainland and caribbean *Anolis* lizards. *Ecology* 78:2191–2203.
- Jolicouer P. 1963. The multivariate generalization of the allometry equation. *Biometrics* 19:497–499.
- Langerhans R.B., DeWitt T.J. 2004. Shared and unique features of evolutionary diversification. *Am. Nat.* 164:335–349.
- Langerhans R.B., Knouft J.H., Losos J.B. 2006. Shared and unique features of diversification in Greater Antillean *Anolis* ecomorphs. *Evolution* 60:362–369.
- Losos J.B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175:623–639.
- Losos J.B. 2011. Convergence, adaptation, and constraint. *Evolution* 65:1827–1840.
- Losos J.B., Jackman T.R., Larson A., de Queiroz K., Rodriguez-Schettino L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279:2115–2118.
- MacArthur R.H., Wilson E.O. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17:373–387.
- Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* 56:701–710.
- Mahler D.L., Ingram T., Revell L.J., Losos J.B. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341:292–295.

- Marsh R.L. 1994. Jumping ability of anuran amphibians. In: Jones J.H. editor. *Advances in Veterinary Science and Comparative Medicine*. New York (NY): Academic Press, p. 51–111.
- Martin C.H., Wainwright P.C. 2013. On the measurement of ecological novelty: scale-eating pupfish are separated by 168 my from other scale-eating fishes. *PLoS One* 8:e71164.
- Melville J., Harmon L.J., Losos J.B. 2006. Intercontinental community convergence of ecology and morphology in desert lizards. *Proc. R. Soc. Lond. B Biol. Sci.* 273:557–563.
- Menzies J. 2006. *The Frogs of New Guinea and the Solomon Islands*. Sofia-Moscow (ID): Pensoft.
- Moen D.S., Irschick D.J., Wiens J.J. 2013. Evolutionary conservatism and convergence both lead to striking similarity in ecology, morphology and performance across continents in frogs. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20132156.
- Moen D.S., Wiens J.J. 2009. Phylogenetic evidence for competitively driven divergence: body-size evolution in Caribbean treefrogs (Hylidae: *Osteopilus*). *Evolution* 63:195–214.
- Monteiro L.R. 2013. Morphometrics and the comparative method: studying the evolution of biological shape. *Hystrix* 24:25–32.
- Muschick M., Indermaur A., Salzburger W. 2012. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr. Biol.* 22:2362–2368.
- Nauwelaerts S., Stamhuis E.J., Aerts P. 2005. Propulsive force calculations in swimming frogs I. A momentum-impulse approach. *J. Exp. Biol.* 208:1435–1443.

- Nielsen R. 2002. Mapping mutations on phylogenies. *Syst. Biol.* 51:729–739.
- Pagel M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* 48:612–622.
- Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20:289–290.
- Pyron R.A., Wiens J.J. 2011. A large-scale phylogeny of Amphibian including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Mol. Phylogenet. Evol.* 61:543–583.
- R Core Team 2012. R: A Language and Environment for Statistical Computing. Vienna, R Foundation for Statistical Computing.
- Rasband W.S. 1997. ImageJ. Bethesda (MD): U. S. National Institutes of Health.
- Revell L.J. 2009. Size-correction and principal components for inerspecific comparative studies. *Evolution* 63:3258–3268.
- Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.
- Revell L.J., Johnson M.A., Schulte J.A., Kolbe J.J., Losos J.B. 2007. A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution* 61:2898–2912.
- Roelants K., Gower D.J., Wilkinson M., Loader S.P., Biju S.D., Guillaume K., Moriau L., Bossuyt F. 2007. Global patterns of diversification in the history of modern amphibians. *Proc. Natl. Acad. Sci. U.S.A.* 104:887–892.
- Schluter D. 2000. *The Ecology of Adaptive Radiation*. Oxford: Oxford University Press.

- Schluter D., McPhail J.D. 1993. Character displacement and replicate adaptive radiation. *Trends Ecol. Evol.* 8:197–200.
- Schluter D., Price T.D., Mooers A.O., Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51:1699–1711.
- Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. *Proc. R. Soc. Lond. B Biol. Sci.* 273:1987–1998.
- Sokal R.R., Rohlf F.J. 1995. *Biometry*. 3rd ed. New York, NY, W.H. Freeman.
- Stayton C.T. 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. *Evolution* 60:824–841.
- Tyler M.J., Knight F. 2009. *Field Guide to the Frogs of Australia*. Collingwood (Victoria, Australia): CSIRO Publications.
- Van Bocxlaer I., Loader S.P., Roelants K., Biju S.D., Menegon M., Bossuyt F. 2010. Gradual adaptation toward a range-expansion phenotype initiated the global radiation of toads. *Science* 327:679–682.
- Wainwright P.C. 2007. Functional versus morphological diversity in macroevolution. *Ann. Rev. Ecol. Evol. Syst.* 38:381–401.
- Wainwright P.C., Alfaro M.E., Bolnick D.I., Hulseley C.D. 2005. Many-to-one mapping of form to function: a general principle in organismal design? *Int. Comp. Biol.* 45:256–262.
- Wake D.B. 1991. Homoplasy: the result of natural selection or evidence of design limitations? *Am. Nat.* 138:543–567.
- Wiens J.J. 2007. Global patterns of diversification and species richness in amphibians. *Am. Nat.* 170:S86–S106.

- Wiens J.J. 2011. The niche, biogeography and species interactions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366:2336–2350.
- Wiens J.J., Brandley M.C., Reeder T.W. 2006. Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. *Evolution* 60:123–141.
- Wiens J.J., Chippindale P.T., Hillis D.M. 2003. When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Syst. Biol.* 52:501–514.
- Winemiller K.O. 1991. Ecomorphological diversification in lowland freshwater fish assemblages from five biotic regions. *Ecol. Monogr.* 61:343–365.
- Young J.E., Christian K.A., Donnellan S., Tracy C.R., Parry D. 2005. Comparative analysis of cutaneous evaporative water loss in frogs demonstrates correlation with ecological habits. *Phys. Biochem. Zool.* 78:847–856.

Figure Legends

FIGURE 1. Graphical comparison of models used in the first test in this paper, 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.

FIGURE 2. Principal components scores for 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 it 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 a microhabitat and the optimum for the ancestral, terrestrial microhabitat are considered to reflect insufficient time to adapt to the current microhabitat, and thus the impact of the past microhabitat on the

current phenotype and 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.

FIGURE 3. Phylogeny and microhabitat use in frogs. The topology is from Pyron and Wiens (2011) and branch lengths (in millions of years; myr) were estimated for this paper 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 7.39 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 Pipoidea are represented by the species *Discoglossus jeanneae* and *Xenopus wittei*, respectively. Finally, circular node labels indicate other important clades discussed in this paper: (1) Microhylidae, (2) Ranidae, (3) Rhacophoridae, (4) Mantellidae, (5) Bufonidae, (6) Hylidae, and (7) Pelodyadinae. See online Fig. S1 for the species name associated with each tip.

FIGURE 4. Relationship between the time since transition into a given microhabitat and the multivariate morphological distance from that microhabitat's estimated

optimum. Time is plotted on a logarithmic scale to reflect that the influence of past selective environments (e.g. microhabitat) decays exponentially over time in OU models (Hansen 1997). The distance from the optimum has no units because it was calculated in principal components space (across PC2–10). Since terrestriality was estimated to be the ancestral microhabitat for all frogs in this analysis (see Results), we only include terrestrial species here that secondarily became terrestrial after spending part of their history in another microhabitat (e.g. *Litoria tornieri*; online Fig. S1; Fig. 2). While there were not enough species in individual microhabitats to reach statistical significance separately, all microhabitat-specific correlations were negative except for terrestrial species (results not shown).

TABLE 1. Results of phylogenetic principal components analysis.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Eigenvalues	8.549	0.653	0.384	0.187	0.081	0.054	0.034	0.029	0.017	0.011
Percent of total variation	0.855	0.065	0.038	0.019	0.008	0.005	0.003	0.003	0.002	0.001
<hr/>										
Eigenvectors										
Original variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Snout-to-vent (body) length	0.336	0.136	-0.098	-0.097	-0.054	0.248	-0.298	-0.323	-0.767	-0.055
Leg length	0.335	0.089	0.048	-0.254	-0.347	0.035	0.312	0.446	-0.040	-0.627
Head length	0.329	0.173	-0.135	-0.291	0.444	-0.317	0.363	0.296	-0.203	0.449
Head width	0.329	0.216	-0.187	-0.061	0.456	-0.234	-0.184	-0.402	0.379	-0.451
Arm length	0.335	0.031	-0.091	-0.153	0.076	0.629	-0.382	0.326	0.374	0.249
Leg mass	0.325	0.285	0.077	-0.127	-0.600	-0.131	0.165	-0.417	0.285	0.364
Tubercle area	0.308	0.216	-0.197	0.878	-0.052	-0.036	0.067	0.197	-0.032	0.027
Foot webbing area	0.279	-0.056	0.924	0.141	0.209	-0.013	-0.044	-0.016	-0.020	0.007
Toe tip area	0.297	-0.572	-0.098	-0.025	-0.219	-0.522	-0.468	0.176	-0.006	0.065
Finger tip area	0.284	-0.661	-0.139	0.085	0.107	0.310	0.500	-0.311	0.044	-0.006

Notes: The phylogeny used for this analysis was the time-calibrated tree from BEAST with fossil calibrations and constrained topology. The similar sign and magnitude of values for the PC1 eigenvector indicates that it represents overall size (Jolicouer 1963)

TABLE 2. SIMMAP-inferred transitions between states over the history of the anuran species we examined in this study.

	arboreal	burrowing	semi-aquatic	terrestrial	torrent
arboreal	-	1.1 (0 ,3)	0.5 (0, 2)	12.0 (6,21)	0.0 (0, 0)
burrowing	0.2 (0, 1)	-	3.0 (1, 6)	3.1 (0 ,8)	0.0 (0, 0)
semi-aquatic	0.1 (0, 1)	2.0 (0, 5)	-	4.4 (1, 9)	5.7 (2,10)
terrestrial	19.6 (14,25)	10.5 (6,15)	12.4 (8,17)	-	0.0 (0, 0)
torrent	0.0 (0, 0)	0.0 (0, 0)	3.4 (0,7)	0.0 (0, 0)	-

Notes: Cell values are the mean number of transitions (the average number across

1,000 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.

TABLE 3. Comparison of OU models with log-likelihood (lnL) and finite-sample-size AIC (AICc).

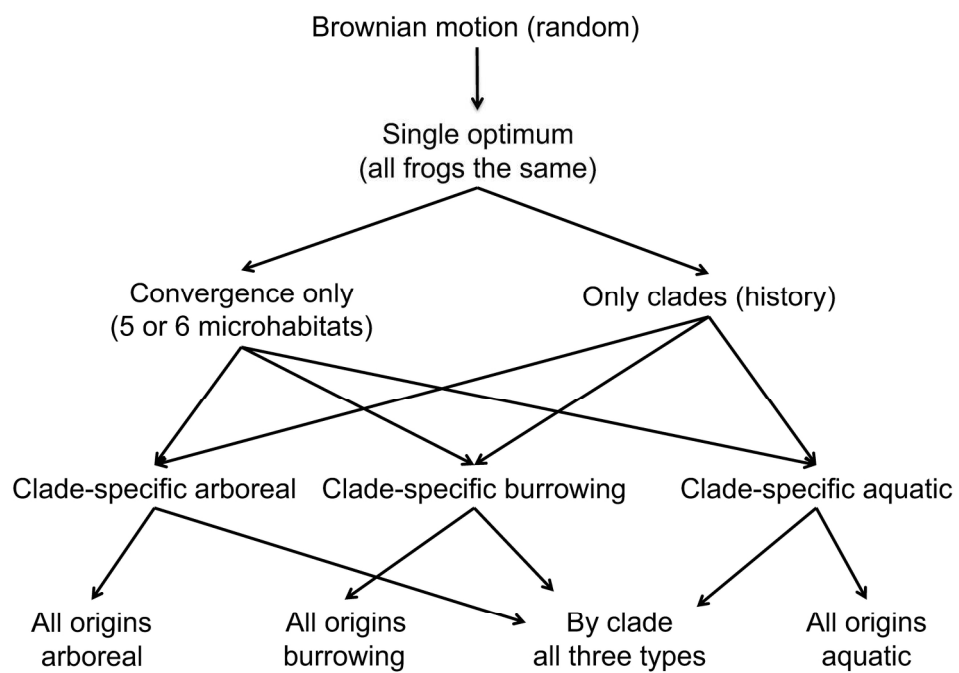
Class	Model	lnL	AICc	Δ AICc	w_i	K
Null model						
	Brownian motion	-3038.1	6112.9	209.8	0.00	18
	Single optimum OU	-3021.3	6097.9	194.8	0.00	27
History only						
	Only clades	-2995.2	6122.7	219.6	0.00	63
Convergence only						
	5 microhabitats	-2887.2	5906.8	3.7	0.14	63
	6 microhabitats	-2875.5	5903.1	0.0	0.84	72
Mixed convergence and history						
	Clade-specific aquatic	-2860.7	5914.1	11.0	0.00	90
	All aquatic origins	-2805.0	5952.9	49.8	0.00	153
	Clade-specific arboreal	-2882.8	5917.8	14.7	0.00	72
	All arboreal origins	-2803.5	5995.5	92.4	0.00	171
	Clade-specific burrowing	-2870.3	5912.8	9.7	0.01	81
	All burrowing origins	-2806.3	5911.1	8.0	0.02	135
	Aquatic, arboreal, and burrowing by clade	-2840.6	5936.5	33.4	0.00	117

Note: AICc shows strongest support for the model with separate morphological optima for aquatic, arboreal, burrowing, semi-aquatic, terrestrial, and torrent species (in bold). Definitions of the other models can be found in the Methods. Note that the parameter numbers (K) reflect the sum of parameters for each model fit across nine morphological characters (PC2–10); for example, the six-microhabitat model has alpha, sigma, and six optima (eight parameters) for nine characters, resulting in 72 total parameters (online Table S5). Δ AICc is the difference between a given model and the model with the lowest AICc, and w_i is the AICc weight, the proportional likelihood support for a given model relative to the others (Burnham and Anderson 2002).

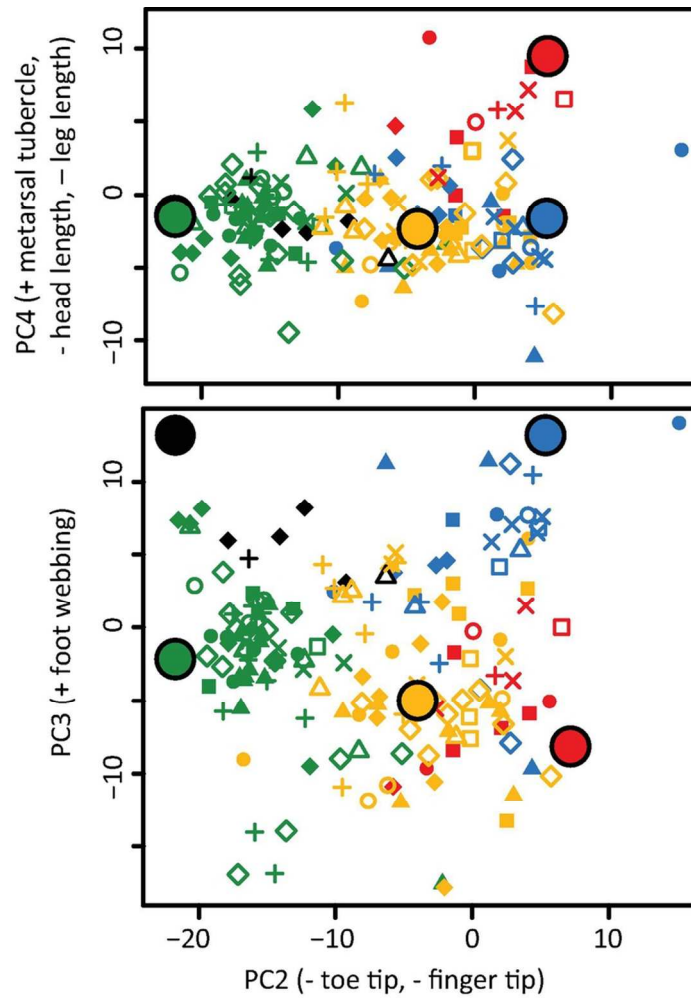
TABLE 4. Partitioned morphological variance of species around the optimum value (estimated by the OU analysis) for the microhabitat that they currently inhabit.

PC Axis	$\sigma^2_{\text{history}}$	σ^2_{rand}	σ^2_{total}	% history	% random
PC2	63.1	18.2	81.3	0.777	0.223
PC3	71.4	27.9	99.2	0.719	0.281
PC4	3.7	7.5	11.2	0.331	0.669
PC5	10.7	5.3	15.9	0.669	0.331
PC6	3.7	4.3	8.0	0.462	0.538
PC7	0.6	2.2	2.8	0.222	0.778
PC8	1.4	1.5	2.8	0.488	0.512
PC9	0.2	1.0	1.1	0.150	0.850
PC10	0.6	0.7	1.3	0.480	0.520
All PC2-10	155.4	68.4	223.8	0.694	0.306

Notes: σ^2 refers to estimated variance for history, random factors, and total (the sum of the first two variances; Sokal and Rohlf 1995). The percent attributed to each factor is simply that variance divided by the total variance for each PC axis.



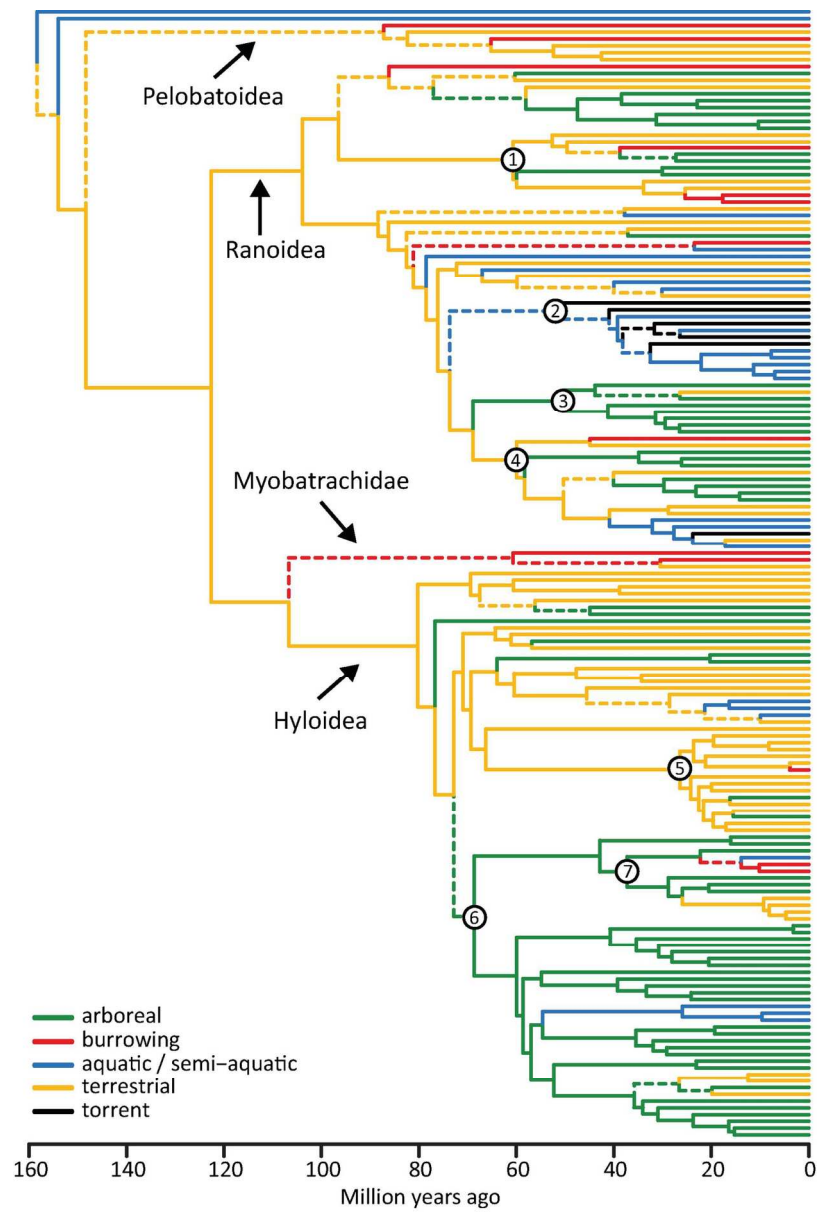
190x142mm (300 x 300 DPI)



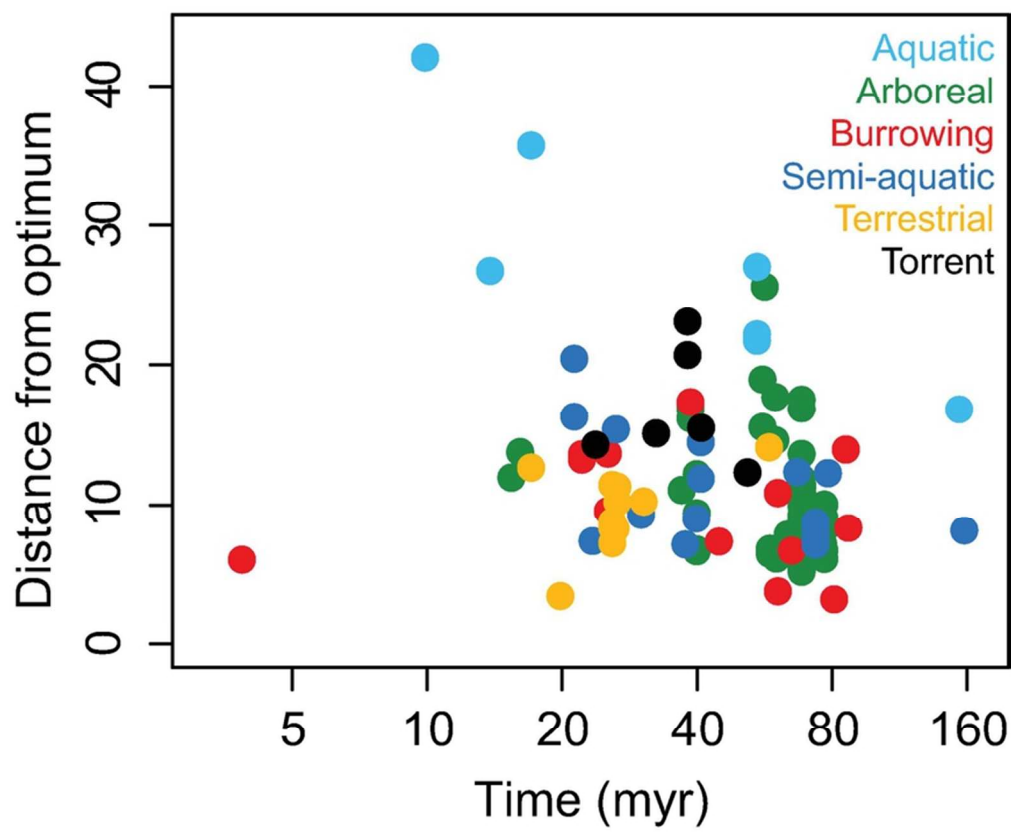
- | | | |
|-----------------------|---------------------------|--------------|
| ○ Purulhá, Guatemala | ● Bwindi Forest, Uganda | Arboreal |
| × SE Virginia, U.S.A. | + Ranomafana, Madagascar | Burrowing |
| ◇ Exporer's Inn, Peru | △ Ponmudi, India | Semi-Aquatic |
| ▲ Boraceia, Brazil | ◆ Nanga Tekalit, Borneo | Terrestrial |
| □ Cádiz, Spain | ■ Middle Point, Australia | Torrent |



100x119mm (300 x 300 DPI)



126x189mm (300 x 300 DPI)



72x61mm (300 x 300 DPI)