

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: <https://doi.org/10.1086/716015>
Copyright 2021 The University of Chicago Press.

Title: Ecological filtering and exaptation in the evolution of marine snakes

Short Running Title: Marine snake convergence

Keywords: Serpentes, habitat transition, body mass, reproduction, biome occupancy, historical biogeography

Article Type: Major Article

Authors:

William Gearty,^{1,2} Elsie Carrillo,^{1,3} and Jonathan L. Payne¹

¹ Department of Geological Sciences, Stanford University, Stanford, CA 94305, USA

² School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

³ Department of Ecology and Evolutionary Biology, University of California-Santa Cruz, Santa Cruz, CA 95064, USA

Abstract

Convergent evolution is often attributed to adaptation of form to function, but it can also result from ecological filtering, exaptation, or nonaptation. Testing among these possibilities is critical to understanding how and why morphological similarities emerge independently in multiple lineages. To address this challenge, we combined multiple preexisting phylogenetic methods to jointly estimate the habitats and morphologies of lineages within a phylogeny. We applied this approach to the invasions of snakes into the marine realm. We utilized a dataset for 1,243 extant snake species consisting of newly compiled biome occupancy information and preexisting data on reproductive strategy, body mass, and environmental temperature and elevation. We find evidence for marine clades arising from a variety of aquatic and terrestrial habitats. Furthermore, there is strong evidence of ecological filtering for non-marine ancestors that were already viviparous, had slightly larger-than-average body sizes, and that lived in environments with higher-than-average temperatures and lower-than-average elevations. In aggregate, similarities among independent lineages of marine snakes result from a combination of exaptation and strong ecological filtering. Strong barriers to entry of new habitats appear to be more important than common adaptations following invasions for producing similarities among independent lineages invading a shared, novel habitat.

Introduction

Convergent evolution of phenotypes is common and often attributed to shared biotic or abiotic constraints (Losos 2017). Classic examples include the evolution of wings across insects, bats, birds, and pterosaurs (Alexander 2015); eyes in molluscs, chordates, and arthropods (Land and Nilsson 2012); and echolocation in cetaceans and bats (Parker et al. 2013). Such convergence is often taken as evidence for adaptation of these ancestrally diverse organisms to their new, shared abiotic and biotic conditions (Losos 2011); however, convergence can result from many different processes (Fig. 1).

The most commonly proposed mechanism driving convergence is *adaptation* from different ancestral trait values to similar descendant trait values for some functional purpose. In such a scenario, the trait or condition of interest evolves in tandem with the occupancy of a new niche, or a set of biotic and abiotic factors, in such a way that increases fitness (Fig. 1, purple path). The repeated evolution of leglessness in squamates as an adaptation for burrowing is an example of such adaptive convergence (Wiens et al. 2006).

Ecological filtering of preexisting traits or conditions could also be the cause of phenotypic similarity across disparate lineages that have invaded similar niches (Fig. 1, teal path). Such a trait may have been previously shaped by natural selection for a different use or conferred no effect on fitness in the ancestor. In the new niche, this trait or condition may be integral to the survival or success of the lineage. Therefore, while many lineages may be capable of dispersing to this new niche, ecological filtering (also known as environmental filtering or habitat filtering) limits the success of potential invaders to those that possess this trait or condition (Emerson and Gillespie 2008; Pardi and Smith 2012; Kraft et al. 2015; Cadotte and Tucker 2017). For example, intermediate sizes evolved in lizards in response to local

competition in multispecies communities, but only lizard species with these intermediate sizes were then able to invade island environments and survive as solitary species (Poe et al. 2007). Similarly, horses of the middle Miocene possessed many different dentition and limb morphologies, but only those with extreme hypsodonty and cursorial limb morphology were able to survive the increasing aridity and expansion of grasslands of the late Miocene, resulting in the limited diversity of horses today (Hulbert 1993).

Alternatively, the trait or condition could be coopted for another function upon invasion of the new niche, a process known as *exaptation* (Gould and Vrba 1982). Examples of such exaptations include the cooption of feathers, which initially evolved for heat regulation and/or display, for flight in birds (Gould and Vrba 1982); the cooption of lungs, which originally evolved to supply oxygen to the heart, as swim bladders in bony fish (Farmer 1997); the cooption of jaw bones, which initially evolved to provide a greater number of hinges in the jaw, as components of the ear in mammals (Anthwal et al. 2013); and the cooption of many physiological mechanisms, including ectothermy, low metabolic rate, and uric acid excretion, which initially evolved as various reptilian adaptations, for survival in desert environments (Bradshaw 1988). In many cases, ecological filtering and exaptation work in tandem, whereby the former results in colonization of new niches by lineages with specific traits or conditions that are then coopted for new functions. In either case, acquisition of the trait or condition prior to invading the new niche is fundamental to enabling the invasion.

Finally, convergence could occur via *nonadaptation* (or adaptation for unrelated functions) across these disparate lineages in similar niches (Fig. 1, yellow path). Following the colonization of the new niche by many different lineages, these colonizers could all evolve a similar trait that is unrelated to their convergent invasions. Such evolutionary changes could be due to constraints

related to the new physical or chemical conditions that these lineages now inhabit, similar biotic interactions, similar changes in climate, or merely coincidence (Stayton 2008; Losos 2011). Such nonadaptive convergence is well-known in developmental and genetic systems (Sackton and Clark 2019) but has also been demonstrated in phenotypic systems. For example, the loss of hind-limb digits in salamanders is developmentally linked to reductions in body size and increases in cell or genome size, resulting in phenotypic convergence across distantly related species but without adaptive benefit (Wake 1991; Losos 2011).

Despite these different pathways in convergent evolution, little attention has been given to disentangling their individual contributions to phenotypic similarity across the tree of life, with most attention being given to the contributions of adaptation (Gould and Lewontin 1979; Losos 2011). Such an attitude may stem from an adaptationist paradigm, but it may also be due to the difficulty in assessing the timing of the acquisition of traits or conditions in comparison to the timing of the colonization of new niches.

Here, we adapt multiple preexisting phylogenetic methods to address this problem with respect to the repeated invasions of the marine realm by snakes. At least four of these invasions exist within crown snakes, and these marine snake lineages share numerous morphological traits, many of which have been identified as convergent adaptations to living in the ocean (Murphy 2012). These traits include the modification of body shape and length (Brischoux and Shine 2011), the presence of specialized salt-secreting glands (Dunson and Dunson 1973; Babonis and Brischoux 2012), nostril valves (Heatwole 1978), lengthening of the vascular lung (Lillywhite et al. 2012; Perez et al. 2019), paddle-shaped tails (Sanders et al. 2012), and viviparity (live-birth) (Neill 1964).

Snakes evolved more than 100 million years ago and have since diversified to reach an extant diversity of over 3,500 species that, together, inhabit every continent except Antarctica (Hsiang et al. 2015; Burbrink et al. 2020). In addition to their geographic breadth, snakes exhibit substantial morphological and ecological disparity. They range in size from 0.6 g to 59,000 g (Fredriksson 2005; Hedges 2008; Feldman and Meiri 2013). Many species are arboreal, spending most of their time in trees; many others are fossorial, digging through leaf litter or burrowing; and yet others are aquatic, living in lakes, streams, or mangroves (Lawing et al. 2012). This ecological and taxonomic diversity, combined with well-known examples of convergent phenotypes within these independent marine lineages, makes snakes a suitable clade to study the contributions of adaptation, ecological filtering/exaptation, and nonadaptation to convergence and phenotypic similarity within lineages that have invaded similar environments.

Ancestral state estimation may be a useful tool for disentangling the roles of these pathways in generating convergent phenotypes. Ancestral state estimation has been used extensively to answer evolutionary questions, but it has also generated concerns related to its assumptions and limitations. Many of these limitations arise from the exclusive use of modern data, which represents an incomplete and nonrandom sampling of the history and diversity of a clade (Hunt and Slater 2016). By integrating fossil data with data from extant species, many of these concerns are resolved or diluted (Soul and Wright 2021). For many clades that have both well-preserved fossil records and robust molecular phylogenies [e.g., Carnivora (Law et al. 2019) and Echinodermata (Wright et al. 2021)], new frontiers in phylogenetic macroevolutionary methods may remove the need for older phylogenetic methods such as ancestral state estimation. However, for clades like snakes with a limited fossil record (Gauthier et al. 2012) but a wealth of

phylogenetic data (Pyron et al. 2013; Tonini et al. 2016), such approaches cannot be employed and ancestral state estimation remains a useful tool for reconstructing evolutionary history.

Here, we employ a new strategy to combine two previously developed ancestral-state-estimation methods. First, we jointly estimate the ancestral states of several traits and abiotic conditions. Multivariate approaches like this one have more statistical power than univariate estimation approaches, particularly when the pairwise trait covariances are taken into account (Zheng et al. 2009; Adams and Collyer 2018). Second, we use a Markov model previously developed for ancestral biogeographic estimation to estimate biome occupancy through time. This model, in contrast to many other ancestral-state-estimation methods, is able to accommodate the large, complex state space that represents the evolution of biome tolerance through time and across snakes, particularly when species can, and do, occupy multiple biomes at any one time (Ree and Smith 2008). The combination of these two methods allows us to determine the relative timing of trait acquisition compared to the timing of invasions into the marine realm, enabling us to discriminate among adaptation, ecological filtering/exaptation, and nonaptation.

To accomplish this discrimination among evolutionary modes, we compiled new data on biome occupancy and combined them with previously gathered data on other biotic and abiotic characteristics. Traits were chosen based on the number and familial diversity of species that have been coded, previous use in the snake literature, and relevance to marine lifestyles. We then inferred which of these traits and conditions facilitated the transition to a fully marine lifestyle (either through exaptation or ecological filtering), which traits evolved as adaptations during these transitions, and which traits arose after the invasions as nonaptations or adaptations for other functions.

Methods and Materials

Data Collection

We collected biome occupancy information from primary literature sources, data aggregation and encyclopedia websites (e.g., IUCN Red List, Encyclopedia of Life), and regional guides to construct a habitat database of extant snakes (<https://purl.stanford.edu/pd926wg6266>). We focused on sampling species that are included in the most recent molecular phylogeny of 1,262 snake species (Pyron et al. 2013; Tonini et al. 2016). Our database includes information on all 1,262 of these species, covering 36% of the ~3,500 described extant species.

There is a lack of consistency across the biome schemes that have been used for snake field guides in the past. Therefore, we defined a new biome scheme that a) focuses on the differences between aquatic and non-aquatic environments for the purposes of this project, b) has broad enough categories for applicability to future projects, and c) has a manageable number of biomes for computationally complex dispersal analyses. We ultimately defined eight biomes for coding habitat occupancy. We split terrestrial environments into four biomes: a) forest, which includes woodland, tropical, temperate, and deciduous forests; b) grassland, which includes temperate grasslands, meadows, savanna, and chaparral; c) desert; and d) montane, which includes alpine and tundra. We split aquatic environments into three biomes: a) marine, which includes open ocean and coral reefs; b) brackish, which includes tidal flats and estuaries; and c) freshwater, which includes ponds, swamps, lakes, and rivers. Finally, we defined a semiaquatic biome, which includes coastal environments, mangroves, and salt marshes. We coded each species as occupying or not occupying each of these eight biomes (i.e., any given species must inhabit at least one biome but can inhabit multiple biomes).

We included reproductive mode because almost all marine snakes and almost all other marine tetrapod groups are viviparous (Neill 1964; O'Keefe and Chiappe 2011). We coded reproductive mode as either oviparous (egg-laying) or viviparous (live-bearing) using the data of Feldman et al. (2015) and Pyron and Burbrink (2014) and were able to code 1,260 of the species in our database. The reproductive strategies of the remaining two species (*Thermophis baileyi* and *Thermophis zhaoermii*) are unknown (Hofmann et al. 2012). Next, body size is an important factor in the evolution of marine vertebrates (Gearty et al. 2018; Gearty and Payne 2020), so we also included body mass estimates for 1,256 of the species in our database from other previous work by Feldman et al. (2016).

We included habitat elevation because it likely plays a key role in the opportunity to invade the oceans and temperature because it is another important factor in the evolution of marine vertebrates (Gearty et al. 2018; Gearty and Payne 2020). Feldman et al. (2015) used distribution maps and the climate variables from WorldClim (Hijmans et al. 2005) to calculate mean elevation and mean spring temperature (considered the breeding season; April to August in the Northern Hemisphere and October to February in the Southern Hemisphere) for 1,048 and 1,049 of these species, respectively, which we also added to our database. Feldman et al. limited their data collection to terrestrial alethinophidians (all snakes minus blind and thread species), so we repeated their approach when coding the remaining snakes in our database. We downloaded occurrences of any species lacking temperature and elevation data from the Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>) using the *rgbif* R package (Chamberlain et al. 2015). For species with any occurrences with geospatial coordinates, we calculated the mean spring temperature (as in Feldman et al. 2015) for each occurrence using the climate variables from WorldClim 2 (Fick and Hijmans 2017) at a resolution of 10 arc minutes for terrestrial

occurrences and the sea surface temperature data from the World Ocean Atlas (Locarnini et al. 2013) at a resolution of 0.25 degrees for marine occurrences. We used the Amazon Web Services Terrain Tiles (<https://aws.amazon.com/public-datasets/terrain/>) via the *elevatr* R package (Hollister et al. 2017) to determine the elevation of these occurrences (with an assumed elevation of zero for marine occurrences). We then calculated the mean elevation and mean temperature across the occurrences for each species, providing data for 201 and 200 additional species, respectively, and resulting in 1,249 species with elevation and temperature data. Of the 1,262 snake species in our dataset, 1,243 are coded for all the above traits and environmental variables.

In order to estimate evolutionary history, we used a set of previously published, dated squamate phylogenies (Pyron et al. 2013; Tonini et al. 2016). We calculated the maximum clade credibility tree using 1000 trees from the posterior of the Tonini et al. Bayesian analysis. We then pruned any species that are not coded for all of our traits of interest, leaving a time-scaled phylogeny with 1,243 tips. All of these species were included in the Pyron et al. backbone of the Tonini et al. analysis, and therefore none of them were placed in the phylogeny based on taxonomy.

Ancestral Biome Occupancy Estimation

We estimated the ancestral character states and biome occupancies across the phylogeny to assess how the timing of marine invasions compares to the timing of biotic and abiotic trait acquisition. To estimate ancestral biome occupancy, we used the dispersal-extinction-cladogenesis (DEC) model as implemented in the R package *BioGeoBEARS* (Ree and Smith 2008; Matzke 2013). The DEC model is a Markov model with an expanded state space that considers possible combinations of states (Ree and Smith 2008). It was originally designed and

has commonly been used for historical biogeography (i.e., the estimation of ancestral geographic ranges on a phylogeny). However, the DEC model is also appropriate for estimating biome occupancy across a phylogeny. This is because biome occupancy, as we have defined it here, also has a complex state space that allows for combinations of biomes to be occupied or tolerated by an individual species and is assumed to be inherited from ancestor to descendant. In this case, “dispersal” along a branch would indicate the addition of a new biome to a lineage’s biome occupancy or tolerance relative to that of its ancestor. In our model framework, we do not prevent “dispersal” between any particular biomes; species may theoretically add any biomes to their tolerance regardless of their ancestral biome occupancy. However, for the purposes of computational tractability, we limited ancestors to a maximum of five biomes, which corresponds to the maximum number of biomes occupied by any extant species. Given this limit, each node has a potential for one of 219 different biome ranges, or combinations of biomes. “Extinction” along a branch would indicate the removal of a biome from a lineage’s biome occupancy or tolerance. Finally, during “cladogenesis”, descendants may inherit a portion or the entirety of their ancestor’s biome occupancy, and individual descendants of the same ancestor may inherit the same or different biome occupancies.

We ran two different versions of the DEC model that estimate the occupancy of all eight biomes across the phylogeny: a) the standard DEC model, an implementation of Lagrange’s model (Ree and Smith 2008) and b) the DEC+J model, the DEC model with an additional parameter for allopatric founder-event speciation (Matzke 2014). In this non-biogeographic case, founder-event speciation represents cladogenetic events where descendants occupy or tolerate a biome or biomes that their ancestor did not. It should be noted that there has been some controversy over the “J” parameter (Ree and Sanmartín 2018); however, these claims have been

rebutted by several authors (McDonald-Spicer et al. 2019; Klaus and Matzke 2020). We used a corrected version of the Akaike information criterion (AICc) to compare the fit of the two models (Burnham and Anderson 2002). We then extracted the marginal probabilities for all of the potential biome ranges at each node of the phylogeny for the single best model. Finally, we calculated the probabilities of each biome being within the occupancy for each node as the sum of the probabilities for the biome ranges that include that biome.

Joint Ancestral State Estimation

We then jointly estimated the ancestral states of reproduction, body mass, elevation, and temperature using a modified version of *threshml* (Felsenstein 2012). This software package uses the threshold model (Wright 1934) and a Markov chain Monte Carlo (MCMC) algorithm to infer the evolutionary covariances of binary and continuous characters. The threshold model assumes that binary character states are the result of a threshold applied to an underlying continuous character (“liability”). During the MCMC, a Gibbs sampler assigns continuous values for each continuous trait and a continuous liability for each binary trait to interior nodes of the tree under a Brownian motion model (Felsenstein 2005, 2012), while a Metropolis algorithm assigns continuous liability values to the tips of the tree (the continuous traits are already observed for the tips) (Felsenstein 2012). An expectation-maximization algorithm is employed to adjust and reduce the sampling from one chain to the next based on a provisional covariance matrix that is inferred at the end of each chain (Felsenstein 2012). Provided sufficient sampling is performed over the course of each chain and sufficient chains are run, the final covariance matrix should approach the maximum likelihood estimate. *threshml* is useful in this scenario because it can accommodate discrete and continuous characters simultaneously. Furthermore, as opposed to

individual ancestral state estimation processes, *threshml* considers the covariances of the characters of interest whilst jointly estimating the character states at nodes across the phylogeny. By default, the *threshml* software outputs a covariance matrix and matrices describing the transformations inferred by the threshold model. However, we have modified the software to take the node estimates that are generated during the Monte Carlo process and output them in addition to this standard output of the software (see Supplemental Code). For this analysis, we ran an initial burn-in of 1,000 steps, then ran 30 sequential chains with 1,500,000 steps each. We used a Metropolis update proposal size of 10, which resulted in an acceptance fraction of ~0.3, which is within the range (0.1 to 0.5) suggested in the documentation (see Supplemental Code). Node estimates for the four characters were reported every 100 steps during the final two chains. We confirmed convergence of these values based on effective sample sizes (>200) and Gelman Rubin statistics (<1.1) using the R package *coda* (Gelman and Rubin 1992; Plummer et al. 2006; see Figs. S1 and S2). The estimates for the reproduction character were reported by *threshml* as continuous liabilities. We converted these to the probability of the character state being viviparous by calculating $1 - \Phi(x)$, where $\Phi(x)$ is the integral of the standard normal distribution from negative infinity to x , the liability value.

Inferring Convergence Mechanisms and Histories

By matching up the character estimates from *threshml* with the biome occupancy probabilities from *BioGeoBEARS*, we are able to estimate the biotic and abiotic states of the ancestors of the various marine snake groups before, during, and after their invasions. We projected the snake phylogeny in various phyloecospaces where the x-axis represents the trait of interest and the y-axis represents the probability of the marine biome being occupied (see Fig.1).

In this visual framework, we can qualitatively assess whether the biotic and abiotic traits were exaptations, ecological filters, adaptations, or nonaptations. If the evolution of the trait (motion along the x-axis) occurs before the invasion of the marine biome (motion along the y-axis), we would infer that the trait was evolved for some other purpose and was then coopted as an exaptation for living in the marine realm (Fig. 1, teal line). Alternatively, if there is no evolution of a trait (no motion along the x-axis) before the invasion of the marine biome, but only a subset of trait values is associated with invasions, we would infer that ecological filtering of the trait is at work (Fig. 1, teal line). Next, if the evolution of the trait occurs during the invasion of the marine biome, we would infer that the trait is an adaptation for living in the marine realm (Fig. 1, purple line). Finally, if the evolution of the trait occurs after the invasion of the marine biome, we would infer that the trait evolved as a nonadaptation (Fig. 1, yellow line).

Furthermore, in order to quantify whether traits evolved significantly differently during invasions as opposed to outside of invasions (before or after), we calculated the estimated amount of net change of each trait along each branch of the phylogeny. We also calculated the estimated rate of change (assuming time homogeneous rates of phenotypic evolution) of each trait along each branch by dividing these net changes by the length of the branches (in millions of years). We performed Mann-Whitney U tests to assess whether the average net changes or rates of change during invasions (net changes of marine biome occupancy probability > 0) were significantly different from those outside of invasions.

Lastly, we also tested whether reproduction, in particular, caused changes in the probability of dispersal between biomes. We performed a separate DEC analysis using the R package *BioGeoBEARS* that allows for an evolving binary character to affect dispersal probability (Matzke 2013; Nicolaï and Matzke 2019; Klaus and Matzke 2020). For each of the

two different models that we ran for our previous DEC analyses, we ran one version of the model where the reproduction and biome occupancy evolved independently. This added two additional trait transition rate parameters to the models ($t12$ and $t21$). We also ran a version of the model where the dispersal was dependent on reproduction state, making the evolution of biome occupancy dependent on the evolution of reproduction. The dispersal multiplier for oviparity ($m1$) was set to 1, while the dispersal multiplier for viviparity ($m2$) was estimated as an additional parameter. We ran these four models on the same eight biome occupancy dataset as before. We then also ran the same four models on a biome occupancy dataset where we collapsed the eight biomes to three biomes (Aquatic, Semiaquatic, and Terrestrial) and on a biome occupancy dataset where we collapsed the eight biomes to two biomes (Marine and Non-Marine). We again used AICc to compare the fit of the four models for each occupancy dataset (Burnham and Anderson 2002). To account for the varying degrees of support among the different models, we calculated AICc weights for each of the sets of four models and used these to calculate model-averaged dispersal multipliers using the formulae of Burnham and Anderson (2002).

Results

Ancestral Habitats of Marine Snakes

Of the 1,240 modern species with complete data, 519 of them occupy only a single biome, 476 occupy two biomes, 204 occupy three biomes, 41 occupy four biomes, and 3 occupy five biomes. No coded species occupies more than five biomes. We identified 38 species that occupy the marine biome, 17 that occupy the brackish biome, 92 that occupy the freshwater biome, 267 that occupy the semiaquatic biome, 921 that occupy the forest biome, 691 that

occupy the grassland biome, 221 that occupy the desert biome, and 15 that occupy the montane biome (Table S1).

The DEC+J model was overwhelmingly supported over the normal DEC model for biome occupancy estimation (Table S2). The results of this model indicate a history of biome occupancy evolution with far more “dispersal” ($d = 0.007$) than “extinction” ($e < 0.001$). This implies that ancestral snakes had narrower biome tolerances, and their average tolerance has gradually increased over the evolutionary history of snakes. This model also includes the process of founder-event speciation (+J), implying that some cladogenetic events are associated with the invasion of biomes that were not included in their ancestral biome ranges, although the model infers that this occurs less frequently than normal dispersal ($j = 0.002$). The ancestral state estimates based on this model indicate that four major groups of living snakes have invaded the ocean: “sea snakes”, including *Hydrophis*, *Hydrelaps*, *Aipysurus*, and *Emydocephalus*; “sea kraits” or *Laticauda*; “water snakes” or Homalopsidae (possibly up to 4 separate invasions); and “file snakes” or *Acrochordus*. According to the model estimates, all of these invasions occurred 10 to 20 million years ago. This number of marine invasions is far fewer than the 15 or more independent clades of snakes that have invaded the freshwater biome and 20 or more that have invaded the semiaquatic biome (Fig. 2).

The occupancy of the freshwater and semiaquatic biomes in comparison to the occupancy of the marine biome across the phylogeny indicates a wide range of invasion histories (Fig. 3). Some ancestors appear to have occupied these freshwater or semiaquatic environments well before invading the ocean (e.g., water snakes, Fig. 3A), others did not occupy freshwater or semiaquatic environments before invading the ocean (e.g., sea snakes, Fig. 3), and still others invaded the freshwater or semiaquatic biomes nearly simultaneously with their invasions of the

marine biome (e.g., sea kraits, Fig 3B). A small number of lineages even entered the freshwater or semiaquatic biomes following their occupancy of the marine biome (e.g., sea snakes, Figs. 2 and 3).

Occupancy of the forest biome is widespread across the snake phylogeny under the DEC+J model, with a very high occupancy probability at the root (Fig. 2). Further, under this model the ancestors of sea snakes, sea kraits, and file snakes occupied the forest biome (Fig. S3A). The grassland biome is also inferred to be widespread across the tree, although the occupancy probability at the root is not as high (Fig. 2). However, none of the immediate ancestors of marine snakes appear to have occupied the grassland biome, nor did they occupy the desert or montane biomes (Fig. S3B-D).

Ancestral Biotic and Abiotic Traits of Marine Snakes

The *threshml* analysis identified a variety of covariations between the abiotic and biotic traits of interest (Table S3). As expected, the model infers a negative covariation between elevation and temperature. The model also indicates a strong negative covariation between reproduction liability and temperature and a slight negative covariation between reproduction and body mass. All other estimated covariations appear to be negligible.

Body mass estimates vary across the snake phylogeny, with many clades containing lineages that independently evolved toward extreme body sizes (Fig. 4). Modern marine snakes evolved from non-marine ancestors within a narrower range of body sizes than that of modern non-marine snakes, and no marine invasions appear to have occurred in clades with extreme (small or large) body sizes (Fig. 5A). Furthermore, the average size of these ancestors of these invasions appears to be slightly larger than the average modern non-marine snake. Some

invading lineages increased in size over the course of their invasions, while others decreased in size or stayed the same size. On average, we find that these evolutionary changes in body size are not significantly different from those of non-invaders (Table S4).

Viviparity appears to have evolved independently more than 20 times over the evolutionary history of snakes (Fig. 4). In the case of snakes that live in the marine realm, three of the four lineages adopted viviparity before they invaded the ocean (Fig. 5B). In the case of sea snakes and water snakes, viviparity seems to have evolved several nodes back in the tree, whereas in file snakes, viviparity may have only evolved one or two nodes back (although there is a very long branch leading to this clade). The one marine clade that is not viviparous is the semiaquatic sea kraits, *Laticauda*.

Snake species that live at high elevations or extreme temperatures exist in numerous clades that are scattered across the phylogeny (Fig. 4). However, we infer that marine snakes evolved exclusively from ancestors that inhabited environments with significantly lower-than-average elevations and significantly warmer-than-average temperatures before invading the ocean (Fig. 6). We find no statistical evidence for a trend in average temperature over the course of these invasions, but we do find strong statistical support for further decrease in elevation over the course of these invasions (Table S4).

Reproduction's Effect on Snake Dispersal

The best-fitting models always included separate dispersal multipliers for oviparity and viviparity, regardless of how biomes were grouped, indicating that reproductive mode has a strong influence on the rate at which biome occupancy or tolerance can expand (Tables S5 – 7).

The viviparity dispersal multiplier decreases with increasing separation of biomes, although the multiplier was always greater than the multiplier of 1 that was set for oviparity (Figure 7).

Discussion

The physical and chemical differences between land and sea present many barriers to entry by terrestrial vertebrates. Such barriers include higher salinity, greater rates of heat loss, greater viscosity, and increased hydrostatic pressure in seawater. Marine snakes have converged on body plans and emerged from environments that have made such invasions possible. However, until now, identifying whether these traits converged as adaptations, ecological filters or exaptations, or as nonaptations has been poorly constrained. While such a question could be tackled from a paleontological perspective, the lack of an adequate fossil record makes this difficult for snakes, and we have therefore opted to use a unique combination of ancestral state estimation methods. We have assembled a biome occupancy dataset for 1,243 modern snakes. We then used joint ancestral state estimation and coopted a Markov model originally designed for modelling historical biogeography to estimate when marine invasions occurred and when traits were acquired across the snake phylogeny. This strategy makes it possible to theorize which of these traits or conditions permitted successful invasions of the marine realm via ecological filtering or exaptation, which evolved as adaptations to life in the ocean over the course of the invasions, and which, if any, evolved after the invasions as nonaptations.

Exaptations and Ecological Filtering

Nearly all extant species of marine snake are viviparous (Fig. 5). The only oviparous genus of marine snake, *Laticauda*, lays its eggs on land by necessity, a behavior akin to that of

sea turtles (Heatwole 1999). All other marine species evolved viviparity well prior to their invasions of the ocean (Fig. 5). Furthermore, trait-dependent biome occupancy models indicate that viviparity promoted increased biome dispersal across the evolutionary history of snakes, particularly between terrestrial and aquatic biomes (Fig. 7).

Viviparity is present in many other extant and extinct marine reptiles, such as mosasaurs, ichthyosaurs, plesiosaurs, and sauropterygians (Boettcher 1990; Caldwell and Lee 2001; Cheng et al. 2004; O'Keefe and Chiappe 2011), but is less common in terrestrial reptiles (including snakes) (Motani et al. 2014). In terrestrial groups that have evolved viviparity, this trait has been proposed as an adaptation to cold or unpredictable climates (Tinkle and Gibbons 1977; Shine 1983, 2014; Feldman et al. 2015), which is supported by the strong negative correlation between reproductive mode and temperature inferred here (Table S3). We estimate that the viviparous non-marine ancestors of modern marine snakes lived in warmer-than-average environments. It is possible that these terrestrial snakes experienced periods of unpredictable climate in the subtropics, possibly in association with droughts (Feldman et al. 2015). Alternatively, viviparous snakes in warm climates could be descendants of snakes from colder climates in which viviparity was adaptive (Shine 2005; Feldman et al. 2015). Viviparity serves a vital function in terrestrial reptiles by allowing for climate control during an embryo's development. In marine snakes, viviparity most likely still serves this purpose, but has also been coopted to prevent suffocation during development, permitting completely marine lifestyles. The amphibious *Laticauda*, which have retained morphological adaptations to moving on land such as laterally expanded ventral scales, would still be able to rely on their terrestrial habitats to lay their eggs, whereas other marine snakes have lost such adaptations in favor of fully marine lifestyles (Heatwole 1999).

Therefore, viviparity has been both ecologically filtered for and used as an exaptation in the terrestrial lineages leading to modern snakes that live their entire lives in the ocean.

Viviparity alone did not guarantee successful marine invasion; most viviparous lineages have not invaded the oceans (Fig. 4; e.g., Boidae, Viperidae, and Natricidae). In addition, modern marine clades appear to have evolved from a smaller range of body sizes than exists in non-marine snakes today, with an average size slightly larger than that of modern non-marine snakes (Fig. 5). Furthermore, we find no signal for increases in body size during marine invasions (Table S4). Taken together, these findings indicate a degree of ecological filtering with respect to body size, where larger body sizes are more viable for invading the marine realm. This finding mirrors results for mammals and crocodiles that indicate an increase in evolutionary pressures upon invading the marine realm and a narrowing of viable body sizes (Gearty et al. 2018; Gearty and Payne 2020). An increase in the minimum body size may be related to a need to dive for extended periods of time. Indeed, based on allometric work in crocodiles (Gearty and Payne 2020), the lungs of extant marine snakes should be able to support dives for at least 45 minutes. This relationship suggests that these snakes are all large enough to support ecologically relevant diving times (e.g., time needed for scouting and probing burrows for eel prey or waiting for their venom to take effect in their prey). Furthermore, many aquatic snakes (e.g., Hydrophiidae and Acrochordidae) have larger lungs than those of similarly-sized terrestrial snakes (Wood and Lenfant 1976; Heatwole 1978), and some marine snakes (e.g., *Hydrophis*) are even capable of cutaneous respiration that can meet up to 20% of their total oxygen demand, in contrast to just 2% in terrestrial snakes (Heatwole and Seymour 1978). These adaptations tend to enable longer dives than would be expected based on allometry alone. Consistent with this prediction, some marine snakes dive for up to two hours at a time (Heatwole 1999).

Tetrapod marine invaders must also overcome extreme thermoregulatory differences in order to survive for extended periods of time in the open ocean. Many tetrapod groups have addressed this thermoregulatory challenge by evolving to larger body sizes (Gearty et al. 2018; Godoy et al. 2019; Gearty and Payne 2020), which provide lower surface area-to-volume ratios, and/or developing insulation from fat (Iverson 2009; Lindgren et al. 2018), both of which reduce the rate of heat-loss. While this selective pressure may also help explain the increase in minimum body size in marine snakes, extant marine snakes do not come close to achieving the gigantic body sizes of other marine reptiles (Godoy et al. 2019; Gearty and Payne 2020). This pattern may be resulting from other constraints on the maximum viable size for marine snakes, such as limited food availability or reduced food acquisition rates in these snakes' habitats (Gearty et al. 2018; Gearty and Payne 2020).

Nonetheless, particularly given their narrow, elongate shapes and, hence, high surface area-to-volume ratios, heat-loss would occur rapidly in marine snakes in the average marine environment. However, modern marine snakes live only in environments with warmer-than-average temperatures, significantly warmer than those of all other snakes (Fig. 6), reducing the selective pressure from heat loss. Some marine snake individuals have been observed outside of tropical and subtropical water, but colder waters inhibit their ability to reproduce (Heatwole 1999; Heatwole et al. 2012), which is even more problematic for marine snakes that already have smaller clutch sizes than other snakes (Lemen and Voris 1981; Ford and Seigel 1989). Here we infer that there has been no significant change in their environmental temperatures since these snakes invaded the oceans (Table S4). Rather, the terrestrial ancestors of modern marine snakes also lived exclusively in significantly warmer-than-average environments (Fig. 6). These snakes appear to have coopted their association with warmer environments to counteract the increased

heat-loss that they would otherwise experience upon invading the oceans. This strong ecological filter has possibly prevented other snake lineages in colder environments, including many which are larger and/or viviparous, from invading the oceans.

Finally, we estimate that modern marine snakes evolved from terrestrial ancestors that lived at elevations that are significantly lower than the average modern terrestrial snake, and their average elevations further decreased significantly over the course of their invasions (Fig. 6, Table S4). Taken together, these findings indicate that elevation has also served as a very strong ecological filter for marine invasions in snakes. This filtering is to be expected, as a snake would need to live in a habitat adjacent to the ocean in order to invade it, and that adjacent habitat would inevitably be at low elevation.

Adaptations, Nonaptations, and Other Factors

Of the traits and conditions included in our analyses, none appear to have evolved or been acquired as adaptations during or nonaptations after marine invasions. However, it is possible that traits we did not include in our analyses evolved as adaptations during these marine invasions, particularly those that might provide fitness benefits in the ocean but are not required for survival or invasion opportunity, such as paddle-shaped tails. Furthermore, ancestral habitat appears to have played a minimal role in determining the viability of such invasions. Snakes have been able to invade the ocean from several different ancestral habitats. For example, many ancestors of marine snakes invaded the ocean from freshwater or semiaquatic environments, but others never invaded freshwater environments or invaded them in tandem with or after they invaded the marine biome (Fig. 3).

Freshwater has been argued as a key transitional habitat for animals moving between marine and terrestrial environments (Vermeij and Dudley 2000; Lee 2016). For example, paleontological evidence shows that the first cetaceans and the ancestors of modern insects lived in freshwater systems (Thewissen and Bajpai 2001; Thewissen and Williams 2002; Glenner et al. 2006). Such an intermediate step could facilitate later invasion of the ocean because some selective pressures are common to all aquatic environments. However, freshwater does not appear to be a required intermediate habitat for snakes invading the marine realm. Instead, many of their ancestors appear to have invaded the ocean directly or nearly directly from forest environments (Fig. S3).

Finally, here we treat elevation and temperature as species traits that evolve. While this approach is partially realistic, given that species most certainly have elevation and temperature tolerance ranges that change over evolutionary time, using ancestral state estimation likely does not capture the entirety of a lineage's environmental history. This may be because these environmental variables also change over time independent of evolutionary pressures. Ideally, future methods would account for both the biological evolution of environmental tolerances and the physical evolution of the environment itself.

Broader Implications

Evolutionary convergence occurs via multiple processes. Ecological filtering constrains solitary island lizards to intermediate body sizes, whereas adaptation following these island colonization events has resulted in greater sexual size dimorphism (Poe et al. 2007). In flowering plants, exaptations have played major roles in the evolution of unique defense and attraction systems, whereas pollination adaptations have resulted in new interactions with herbivores

(Armbruster et al. 2009). In rocky intertidal fishes, the reduction of sexual and agonistic displays in males may represent an adaptation to the strong wave-action in the intertidal zone, while behavioral guarding was ecologically filtered for in the colonization of the intertidal by these fishes (Almada and Santos 1995). Finally, acanthocephalan and trematode parasitic worms have induced behaviors in their bird and amphipod hosts that are the results of both adaptations and exaptations (Beisel and Médoc 2010). Despite these cases, most reports of convergence focus solely on adaptations. In most of these cases, details of timing are not reported, limiting the inferences that can be made about the processes that have resulted in convergence.

Therefore, estimation of ancestral traits and biome occupancy using phylogenetic comparative methods provides an empirical solution to the problem of estimating the relative timing of trait acquisition versus habitat transition (or simply, the relative timing of acquisition of two different traits). By estimating the timing of trait acquisition relative to habitat transition, phenotypic modifications can be classified as adaptations, ecological filters, exaptations, or nonaptations. Within the scope of biome occupancy, this approach becomes a powerful tool for understanding the required traits and conditions for successful invasions and the traits and conditions that may increase fitness thereafter, ensuring longevity within the new environment.

This approach also has the potential to be extended to incorporate fossil data. In the case of snakes illustrated herein, the fossil record was used to inform the dating of the phylogeny. Time-scaling the tree is a key component of assessing the timing of trait and condition acquisition. Unfortunately, few fossil snakes have enough of their skeletons preserved to estimate body size or reproductive mode. Other clades have better fossil records, and the information from fossils in these clades could be used to validate joint estimation results, add certainty to ancestral state estimations, inform transitional character states, or even apply this

method to fully extinct clades. In the case of snakes, exceptionally preserved snake fossils found in the future could be used to confirm or reject various aspects of the results presented herein. For example, the presence of viviparity in a non-marine ancestor of *Acrochordus* would confirm that the clade possessed viviparity before invading the ocean, whereas the absence of viviparity in the earliest marine ancestor of *Acrochordus* would reject this result. However, habitat transitions can be geologically rapid; in such cases, even a modestly complete fossil record may not record transitional states of the invading taxa. Regardless, the addition of fossil information to any evolutionary biological model should improve both accuracy and precision (Finarelli and Flynn 2006; Hunt and Slater 2016).

Conclusion

Most phenotypic convergence is attributed to adaptation to similar environments, but these morphological similarities among lineages are more often a result of a complex combination of adaptations, ecological filtering, exaptations, and nonaptations. Modern snakes exhibit many evolutionary novelties and acquired conditions that appear to have helped them survive and ultimately prosper in the sea, such as viviparity, larger body sizes, and greater association with warmer, lowland environments. All of these traits and environmental associations appear to have evolved prior to the invasions of the marine realm and were required or nearly so to successfully invade the ocean. Living at lower elevations provided proximity to viable ocean environments. Viviparity and warmer temperatures permitted the ancestors of these marine lineages to survive in an environment with significantly higher rates of heat-loss than on land. Finally, slightly larger body sizes may have been required for permitting sufficiently long dive times. The terrestrial lineages that led to modern marine snakes emerged from many

different biomes, such as forests and other aquatic habitats. In aggregate, a combination of exaptation and strong ecological filtering resulted in morphological similarities between independent groups of marine snakes. Convergent evolution is not always driven by adaptation; rather, the strict requirements of new habitats may enforce strong barriers and filters on the species that can successfully invade them.

Acknowledgments

We are indebted to Christianne Ormsby for collecting snake biome data from the primary literature. We thank Joe Felsenstein for assistance with *threshml*, Nick Matzke for assistance with *BioGeoBEARS*, and Graham Slater and two anonymous reviewers for feedback and discussion on earlier versions of this manuscript. This research was supported by a grant from the U.S. National Science Foundation (EAR-1151022 to JLP). WG was supported by the ARCS Foundation Scholarship and the Population Biology Program of Excellence Postdoctoral Fellowship from the University of Nebraska-Lincoln School of Biological Sciences.

Statement of Authorship

WG and EC conceived of the study and carried out the data collection. WG performed the statistical analyses and drafted the manuscript. JLP and EC participated in the design of the study and helped revise the manuscript. All authors gave final approval for publication.

Data and Code Accessibility Statement

The code and data for all analyses are archived in the Stanford Digital Repository (<https://purl.stanford.edu/pd926wg6266>, Gearty et al. 2021).

Literature Cited

- Adams, D. C., and M. L. Collyer. 2018. Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations. *Systematic Biology* 67:14–31.
- Alexander, D. E. 2015. On the Wing: Insects, Pterosaurs, Birds, Bats and the Evolution of Animal Flight. Oxford University Press, Oxford.
- Almada, V. C., and R. S. Santos. 1995. Parental care in the rocky intertidal: a case study of adaptation and exaptation in Mediterranean and Atlantic blennies. *Reviews in Fish Biology and Fisheries* 5:23–37.
- Anthwal, N., L. Joshi, and A. S. Tucker. 2013. Evolution of the mammalian middle ear and jaw: adaptations and novel structures. *Journal of Anatomy* 222:147–160.
- Armbruster, W. S., J. Lee, and B. G. Baldwin. 2009. Macroevolutionary patterns of defense and pollination in *Dalechampia* vines: Adaptation, exaptation, and evolutionary novelty. *Proceedings of the National Academy of Sciences of the USA* 106:18085–18090.
- Babonis, L. S., and F. Brischoux. 2012. Perspectives on the Convergent Evolution of Tetrapod Salt Glands. *Integrative and Comparative Biology* 52:245–256.
- Beisel, J. N., and V. Médoc. 2010. Bird and amphipod parasites illustrate a gradient from adaptation to exaptation in complex life cycle. *Ethology Ecology and Evolution* 22:265–270.
- Boettcher, R. 1990. New information on the reproductive biology of ichthyosaurs (Reptilia). *Stuttgarter Beitraege zur Naturkunde Serie B (Geologie und Palaeontologie)* 1–51.
- Bradshaw, S. D. 1988. Desert reptiles: a case of adaptation or pre-adaptation? *Journal of Arid Environments* 14:155–174.
- Brischoux, F., and R. Shine. 2011. Morphological adaptations to marine life in snakes. *Journal of Morphology* 272:566–572.
- Burbrink, F. T., F. G. Grazziotin, R. A. Pyron, D. Cundall, S. Donnellan, F. Irish, J. S. Keogh, et al. 2020. Interrogating Genomic-Scale Data for Squamata (Lizards, Snakes, and

Amphisbaenians) Shows no Support for Key Traditional Morphological Relationships. *Systematic Biology* 69:502–520.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference. Springer New York, New York, NY.

Cadotte, M. W., and C. M. Tucker. 2017. Should Environmental Filtering be Abandoned? *Trends in Ecology and Evolution* 32:429–437.

Caldwell, M. W., and M. S. Y. Lee. 2001. Live birth in Cretaceous marine lizards (mosasauroids). *Proceedings of the Royal Society B* 268:2397–2401.

Chamberlain, S., K. Ram, V. Barve, and D. Mcglinn. 2015. rgbif: Interface to the Global Biodiversity Information Facility API.

Cheng, Y., X. Wu, and Q. Ji. 2004. Triassic marine reptiles gave birth to live young. *Nature* 432:383–386.

Dunson, W. A., and M. K. Dunson. 1973. Convergent evolution of sublingual salt glands in the marine file snake and the true sea snakes. *Journal of Comparative Physiology* 86:193–208.

Emerson, B. C., and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution* 23:619–630.

Farmer, C. 1997. Did lungs and the intracardiac shunt evolve to oxygenate the heart in vertebrates? *Paleobiology* 23:358–372.

Feldman, A., A. M. Bauer, F. Castro-Herrera, L. Chirio, I. Das, T. M. Doan, E. Maza, et al. 2015. The geography of snake reproductive mode: a global analysis of the evolution of snake viviparity. *Global Ecology and Biogeography* 24:1433–1442.

Feldman, A., and S. Meiri. 2013. Length-mass allometry in snakes. *Biological Journal of the Linnean Society* 108:161–172.

Feldman, A., N. Sabath, R. A. Pyron, I. Mayrose, and S. Meiri. 2016. Body sizes and diversification rates of lizards, snakes, amphisbaenians and the tuatara. *Global Ecology and Biogeography* 25:187–197.

- Felsenstein, J. 2005. Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society B* 360:1427–1434.
- _____. 2012. A comparative method for both discrete and continuous characters using the threshold model. *American Naturalist* 179:145–56.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37:4302–4315.
- Finarelli, J. A., and J. J. Flynn. 2006. Ancestral State Reconstruction of Body Size in the Caniformia (Carnivora, Mammalia): The Effects of Incorporating Data from the Fossil Record. (T. Oakley, ed.) *Systematic Biology* 55:301–313.
- Ford, N. B., and R. A. Seigel. 1989. Relationships among body size, clutch size, and egg size in three species of oviparous snakes. *Herpetologica* 45:75–83.
- Fredriksson, G. M. 2005. Predation on sun bears by reticulated python in East Kalimantan, Indonesian Borneo. *Raffles Bulletin of Zoology* 53:165–168.
- Gauthier, J. A., M. Kearney, J. A. Maisano, O. Rieppel, and A. D. B. Behlke. 2012. Assembling the Squamate Tree of Life : Perspectives from the Phenotype and the Fossil Record Assembling the Squamate Tree of Life : Perspectives from the Phenotype and the Fossil Record 53:3–308.
- Gearty, W., C. R. McClain, and J. L. Payne. 2018. Energetic tradeoffs control the size distribution of aquatic mammals. *Proceedings of the National Academy of Sciences of the USA* 115:4194–4199.
- Gearty, W., E. Carrillo, and J. L. Payne. 2021. Data for "Ecological filtering and exaptation in the evolution of marine snakes". Stanford Digital Repository. Available at: <https://purl.stanford.edu/pd926wg6266>
- Gearty, W., and J. L. Payne. 2020. Physiological constraints on body size distributions in Crocodyliformes. *Evolution* 74:245–255.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7:457–472.

- Glenner, H., P. F. Thomsen, M. B. Hebsgaard, M. V. Sørensen, and E. Willerslev. 2006. The Origin of Insects. *Science* 314:1883–1884.
- Godoy, P. L., R. B. J. Benson, M. Bronzati, and R. J. Butler. 2019. The multi-peak adaptive landscape of crocodylomorph body size evolution. *BMC Evolutionary Biology* 19:167.
- Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society B* 205:581–598.
- Gould, S. J., and E. S. Vrba. 1982. Exaptation—a Missing Term in the Science of Form. *Paleobiology* 1:4–15.
- Heatwole, H. 1978. Adaptations of Marine Snakes: Unusual physiological strategies have enabled some snakes to live in a saltwater environment. *American Scientist* 66:594–604.
- Heatwole, H. 1999. *Sea Snakes*. Krieger Publishing Company.
- Heatwole, H., A. Grech, J. F. Monahan, S. King, and H. Marsh. 2012. Thermal biology of sea snakes and sea kraits. Pages 257–273 in *Integrative and Comparative Biology* (Vol. 52). Narnia.
- Heatwole, H., and R. Seymore. 1978. Cutaneous oxygen uptake in three groups of aquatic snakes. *Australian Journal of Zoology* 26:481.
- Hedges, S. B. 2008. At the lower size limit in snakes: two new species of threadsnakes (Squamata: Leptotyphlopidae: Leptotyphlops) from the Lesser Antilles. *Zootaxa* 30:1–30.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- Hofmann, S., P. Fritzsche, T. Solhøy, T. Dorge, and G. Miehe. 2012. Evidence of Sex-biased Dispersal in *Thermophis baileyi* Inferred from Microsatellite Markers. *Herpetologica* 68:514–522.
- Hollister, J., T. Shah, and A. L. Robitaille. 2017. elevatr: Access Elevation Data from Various APIs.

- Hsiang, A. Y., D. J. Field, T. H. Webster, A. D. Behlke, M. B. Davis, R. A. Racicot, and J. A. Gauthier. 2015. The origin of snakes: Revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. *BMC Evolutionary Biology* 15:87.
- Hulbert, R. C. 1993. Taxonomic evolution in North American Neogene horses (subfamily Equinae): the rise and fall of an adaptive radiation. *Paleobiology* 19:216–234.
- Hunt, G., and G. Slater. 2016. Integrating Paleontological and Phylogenetic Approaches to Macroevolution. *Annual Review of Ecology, Evolution, and Systematics* 47:189–213.
- Iverson, S. J. 2009. Blubber. Pages 115–120 in *Encyclopedia of Marine Mammals*. Elsevier Ltd.
- Klaus, K. V., and N. J. Matzke. 2020. Statistical Comparison of Trait-Dependent Biogeographical Models Indicates That Podocarpaceae Dispersal Is Influenced by Both Seed Cone Traits and Geographical Distance. (V. Savolainen, ed.) *Systematic Biology* 69:61–75.
- Kraft, N. J. B., P. B. Adler, O. Godoy, E. C. James, S. Fuller, and J. M. Levine. 2015. Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* 29:592–599.
- Land, M. F., and D.-E. Nilsson. 2012. *Animal Eyes*. Oxford University Press, Oxford.
- Law, C. J., G. J. Slater, and R. S. Mehta. 2019. Shared extremes by ectotherms and endotherms: Body elongation in mustelids is associated with small size and reduced limbs. *Evolution* 73:735–749.
- Lawing, A. M., J. J. Head, and P. D. Polly. 2012. The Ecology of Morphology: The Ecometrics of Locomotion and Macroenvironment in North American Snakes. Pages 117–146 in *Paleontology in Ecology and Conservation*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Lee, C. E. 2016. Evolutionary mechanisms of habitat invasions, using the copepod *Eurytemora affinis* as a model system. *Evolutionary Applications* 9:248–270.
- Lemen, C. A., and H. K. Voris. 1981. A Comparison of Reproductive Strategies Among Marine Snakes. *The Journal of Animal Ecology* 50:89.

- Lillywhite, H. B., J. S. Albert, C. M. Sheehy, and R. S. Seymour. 2012. Gravity and the evolution of cardiopulmonary morphology in snakes. Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology 161:230–242.
- Lindgren, J., P. Sjövall, V. Thiel, W. Zheng, S. Ito, K. Wakamatsu, R. Hauff, et al. 2018. Soft-tissue evidence for homeothermy and crypsis in a Jurassic ichthyosaur. Nature 564:359–365.
- Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, et al. 2013. World ocean atlas 2013. Volume 1, Temperature.
- Losos, J. 2017. Improbable destinies: Fate, chance, and the future of evolution. Riverhead Books, New York.
- Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.
- Matzke, N. J. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers of Biogeography 5.
- Matzke, N. J. 2014. Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. Systematic Biology 63:951–970.
- McDonald-Spicer, C., N. J. Knerr, F. Encinas-Viso, and A. N. Schmidt-Lebuhn. 2019. Big data for a large clade: Bioregionalization and ancestral range estimation in the daisy family (Asteraceae). Journal of Biogeography 46:255–267.
- Motani, R., D. Jiang, A. Tintori, O. Rieppel, and G. Chen. 2014. Terrestrial Origin of Viviparity in Mesozoic Marine Reptiles Indicated by Early Triassic Embryonic Fossils. (P. Dodson, ed.) PLoS ONE 9:e88640.
- Murphy, J. C. 2012. Marine invasions by non-sea snakes, with thoughts on terrestrial-aquatic-marine transitions. Pages 217–226 in Integrative and Comparative Biology (Vol. 52). Narnia.
- Neill, W. T. 1964. Viviparity in Snakes: Some Ecological and Zoogeographical Considerations. The American Naturalist 98:35–55.

- Nicolaï, M. P. J., and N. J. Matzke. 2019. Trait- based range expansion aided in the global radiation of Crocodylidae. *Global Ecology and Biogeography* 28:1244–1258.
- O'Keefe, F. R., and L. M. Chiappe. 2011. Viviparity and K-selected life history in a mesozoic marine plesiosaur (Reptilia, Sauropterygia). *Science* 333:870–873.
- Pardi, M. I., and F. A. Smith. 2012. Paleoecology in an era of climate change: How the past can provide insights into the future. Pages 93–116 in *Paleontology in Ecology and Conservation*. Springer Berlin Heidelberg.
- Parker, J., G. Tsagkogeorga, J. A. Cotton, Y. Liu, P. Provero, E. Stupka, and S. J. Rossiter. 2013. Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502:228–231.
- Perez, D., C. M. Sheehy, and H. B. Lillywhite. 2019. Variation of organ position in snakes. *Journal of Morphology* 280:1798–1807.
- Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* 6:7–11.
- Poe, S., J. R. Goheen, and E. P. Hulebak. 2007. Convergent exaptation and adaptation in solitary island lizards. *Proceedings of the Royal Society B* 274:2231–2237.
- Pyron, R. A., and F. T. Burbrink. 2014. Early origin of viviparity and multiple reversions to oviparity in squamate reptiles. *Ecology Letters* 17:13–21.
- Pyron, R. A., F. T. Burbrink, and J. J. Wiens. 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* 13:93.
- Ree, R. H., and I. Sanmartín. 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography* 45:741–749.
- Ree, R. H., and S. A. Smith. 2008. Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. (A. Baker, ed.) *Systematic Biology* 57:4–14.

- Sackton, T. B., and N. Clark. 2019. Convergent evolution in the genomics era: new insights and directions. *Philosophical Transactions of the Royal Society B* 374:20190102.
- Sanders, K. L., A. R. Rasmussen, and J. Elmberg. 2012. Independent Innovation in the Evolution of Paddle-Shaped Tails in Viviparous Sea Snakes (Elapidae: Hydrophiinae). *Integrative and Comparative Biology* 52:311–320.
- Shine, R. 1983. Reptilian viviparity in cold climates: testing the assumptions of an evolutionary hypothesis. *Oecologia* 57:397–405.
- . 2005. Life-History Evolution in Reptiles. *Annual Review of Ecology, Evolution, and Systematics* 36:23–46.
- . 2014. Evolution of an Evolutionary Hypothesis: A History of Changing Ideas about the Adaptive Significance of Viviparity in Reptiles. *Journal of Herpetology* 48:147–161.
- Soul, L. C., and D. F. Wright. 2021. Phylogenetic Comparative Methods: A User's Guide for Paleontologists. *Elements of Paleontology*.
- Stayton, C. T. 2008. Is convergence surprising? An examination of the frequency of convergence in simulated datasets. *Journal of Theoretical Biology* 252:1–14.
- Thewissen, J. G. M., and S. Bajpai. 2001. Whale origins as a poster child for macroevolution. *BioScience* 51:1037–1049.
- Thewissen, J. G. M., and E. M. Williams. 2002. The Early Radiations of Cetacea (Mammalia): Evolutionary Pattern and Developmental Correlations. *Annual Review of Ecology and Systematics* 33:73–90.
- Tinkle, D. W., and J. W. Gibbons. 1977. The distribution and evolution of viviparity in reptiles. *Miscellaneous Publications Museum of Zoology University of Michigan* 1–55.
- Tonini, J. F. R., K. H. Beard, R. B. Ferreira, W. Jetz, and R. A. Pyron. 2016. Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation* 204:23–31.
- Vermeij, G. J., and R. Dudley. 2000. Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? *Biological Journal of the Linnean Society* 70:541–554.

- Wake, D. B. 1991. Homoplasy: The Result of Natural Selection, or Evidence of Design Limitations? *American Naturalist* 138:543–567.
- Wiens, J. J., M. C. Brandley, and T. W. Reeder. 2006. Why does a trait evolve multiple times within a clade? Repeated evolution of snakeline body form in squamate reptiles. *Evolution* 60:123–141.
- Wood, S., and C. Lenfant. 1976. Respiration: mechanics, control and gas exchange. *Biology of the Reptilia* 5:225–274.
- Wright, A., P. Wagner, and D. Wright. 2021. Testing character-evolution models in phylogenetic paleobiology: a case study with Cambrian echinoderms. *Elements of Paleontology*.
- Wright, S. 1934. An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs. *Genetics* 19:506–36.
- Zheng, L., A. R. Ives, T. Garland, B. R. Larget, Y. Yu, and K. Cao. 2009. New multivariate tests for phylogenetic signal and trait correlations applied to ecophysiological phenotypes of nine *Manglietia* species. *Functional Ecology* 23:1059–1069.

Figure Legends

Figure 1: The three pathways to evolutionary convergence. Traits may evolve during environmental invasions as *adaptations* to the new environment (purple line). Alternatively, *ecological filtering* may prevent invasions by species without such evolved traits that are necessary for survival in the new environment, or traits may evolve before environmental invasions and be coopted for a new purpose in the new environment as *exaptations* (teal line). Finally, traits may evolve after invasions as *nonaptations* or adaptations that are not related to the initial invasion of the new environment (yellow line).

Figure 2: Ancestral state estimation of biome occupancy across the snake phylogeny. Occupancy of the marine, freshwater, semiaquatic, forest, grassland, and desert biomes are plotted as the summed posterior probabilities of each of these biomes being within the ancestral biome range as estimated by the DEC+J model in *BioGeoBEARS*. Black lines at tips indicate marine species. Lettered circles indicate most recent common ancestors of marine clades (S: “sea snakes”, including *Hydrophis*, *Hydrelaps*, *Aipysurus*, and *Emydocephalus*; K: “sea kraits” or *Laticauda*; W: “water snakes” or Homalopsidae; F: “file snakes” or *Acrochordus*).

Figure 3: Phyloecospace of snakes for marine biome occupancy by freshwater (A) and semiaquatic (B) biome occupancy. Lines and vertices represent the branches and nodes, respectively, of the phylogeny through time (color represents age) plotted by the probability of their estimated biome ranges including the marine biome and the probability of their estimated biome range including the freshwater (A) or semiaquatic (B) biomes. All probabilities come from the DEC+J model in *BioGeoBEARS*. Blue dots represent modern marine snakes with jitter

to show point density. Node annotations as in Fig. 2. Some ancestors of marine snakes avoided freshwater and semiaquatic biomes, some invaded them beforehand, and others invaded them at the same time.

Figure 4: Ancestral state estimations of biotic and abiotic variables across the snake phylogeny. Reproductive mode is plotted as the mean probability of being viviparous. Body mass, temperature, and elevation are plotted as the mean ancestral state estimate. All four traits were jointly estimated using *threshml*. Tip and node annotations as in Fig. 2.

Figure 5: Phyloecospace of snakes for marine biome occupancy by body mass (A) and reproductive mode (B). Lines and vertices represent the branches and nodes, respectively, of the phylogeny through time (color represents age) plotted by the probability of their estimated biome ranges including the marine biome as estimated by the DEC+J model in *BioGeoBEARS* and their mean estimated mass (A) or mean estimated probability of being viviparous (B) from *threshml*. Blue dots represent modern marine snakes (with jitter in B to show point density). Boxplots represent body mass distributions of modern snakes split by marine occupancy (marine: blue; non-marine: orange). Node annotations as in Fig. 2. A narrow range of ancestral body sizes appear to have been viable for marine invasions, and body size did not significantly change during marine invasions. Three out of four lineages of marine snakes evolved a viviparous reproductive mode before invading the ocean.

Figure 6: Phyloecospace of snakes for marine biome occupancy by elevation (A) and temperature (B). Lines and vertices represent the branches and nodes, respectively, of the

phylogeny through time (color represents age) plotted by the probability of their estimated biome ranges including the marine biome as estimated by the DEC+J model in *BioGeoBEARS* and their mean estimated elevation (A) or temperature (B) from *threshml*. Blue dots represent modern marine snakes. Boxplots as in Fig. 5. Node annotations as in Fig. 2. Ancestors of modern marine groups inhabited environments with significantly lower-than-average elevations before invading the ocean. Ancestors of modern marine groups also inhabited warmer-than-average environments before invading the ocean.

Figure 7: Model-averaged dispersal multipliers for reproductive mode based on different biome occupancy datasets. The multiplier for “Oviparous” was always fixed to 1. A value greater than 1 for “Viviparous” indicates that dispersal rates are relatively higher for viviparous species, a value less than 1 indicates that dispersal rates are relatively lower for viviparous species, and a value of 1 indicates that dispersal rates were not affected by reproductive mode.

Figure 1

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: <https://doi.org/10.1086/716015>
Copyright 2021 The University of Chicago Press.

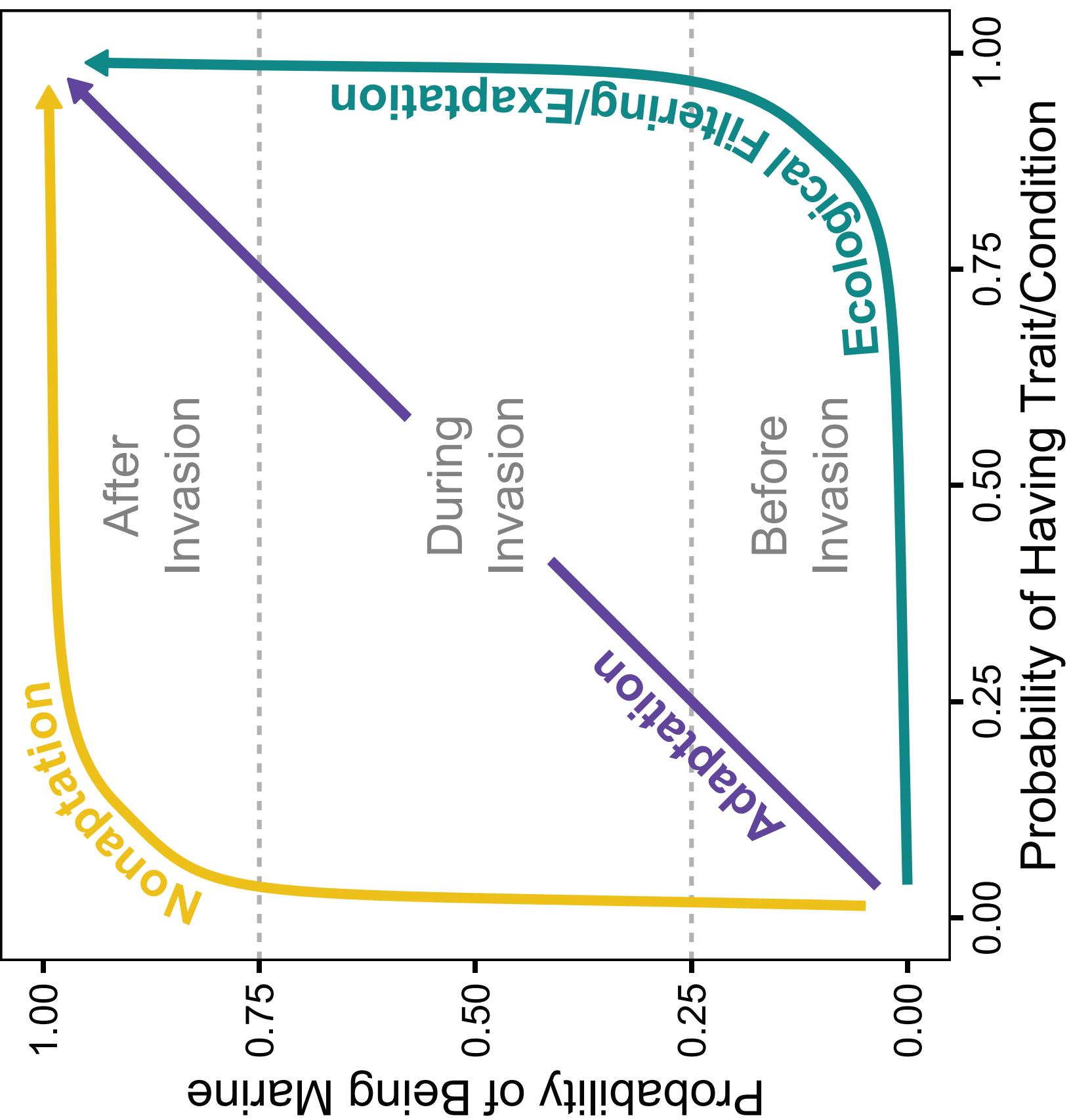


Figure 2

This is the author's accepted manuscript, without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: <https://doi.org/10.1086/716015>

Copyright 2021 The University of Chicago Press

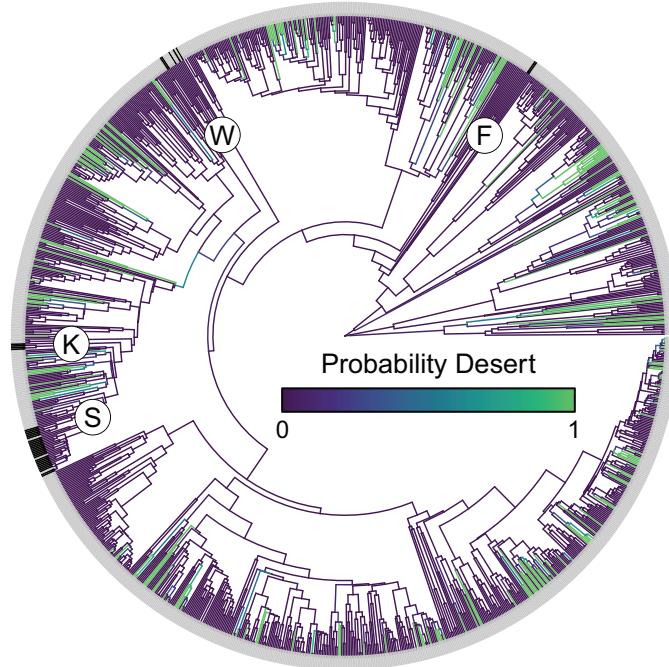
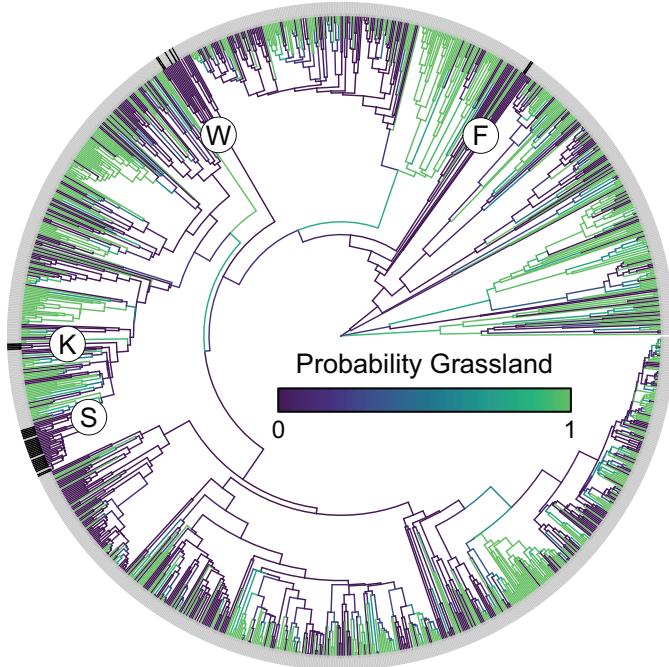
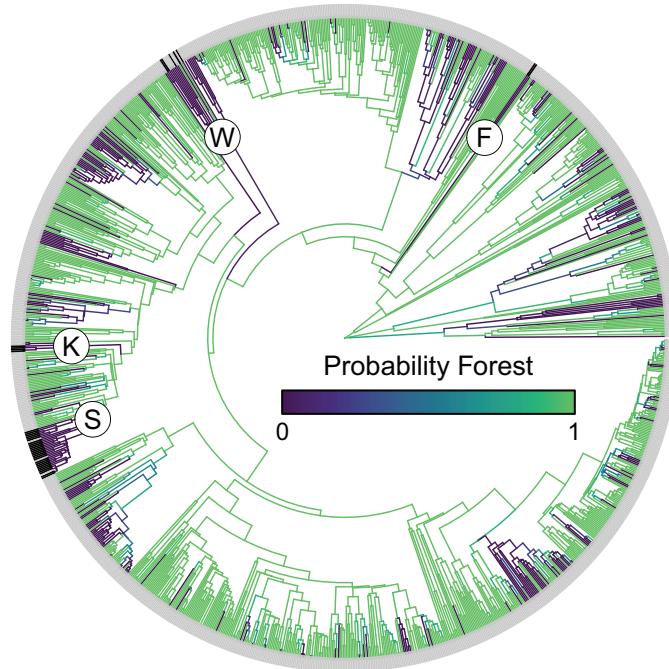
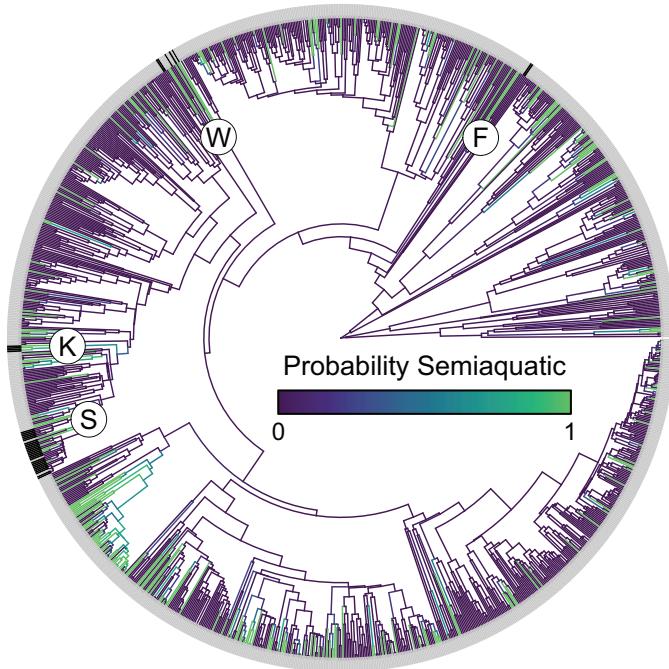
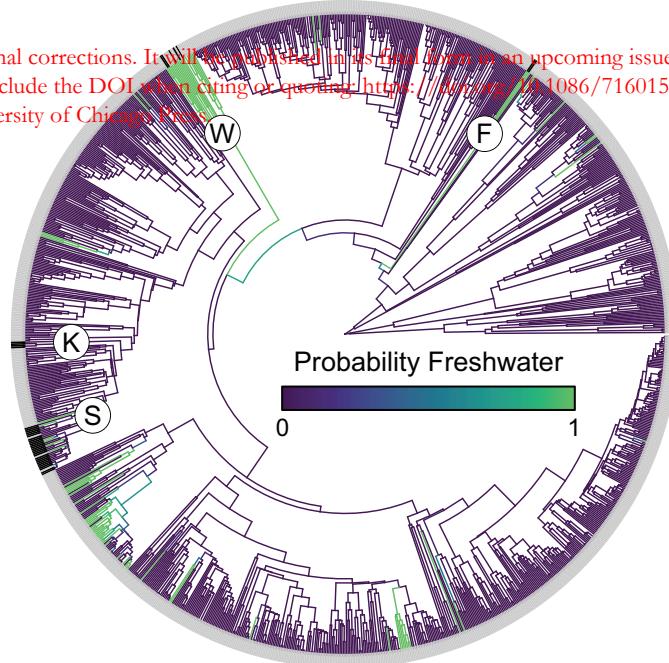
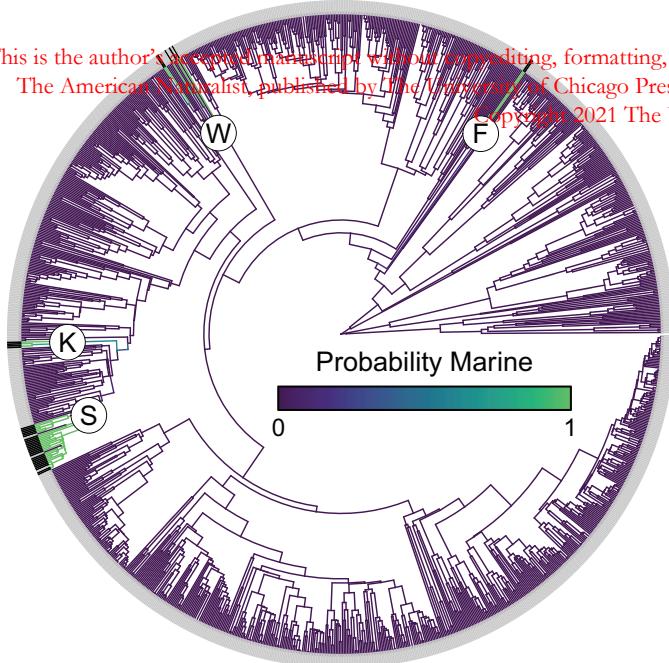


Figure 3

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press, include the DOI when citing or quoting: <https://doi.org/10.1086/716015>
 Copyright 2019 The University of Chicago Press.

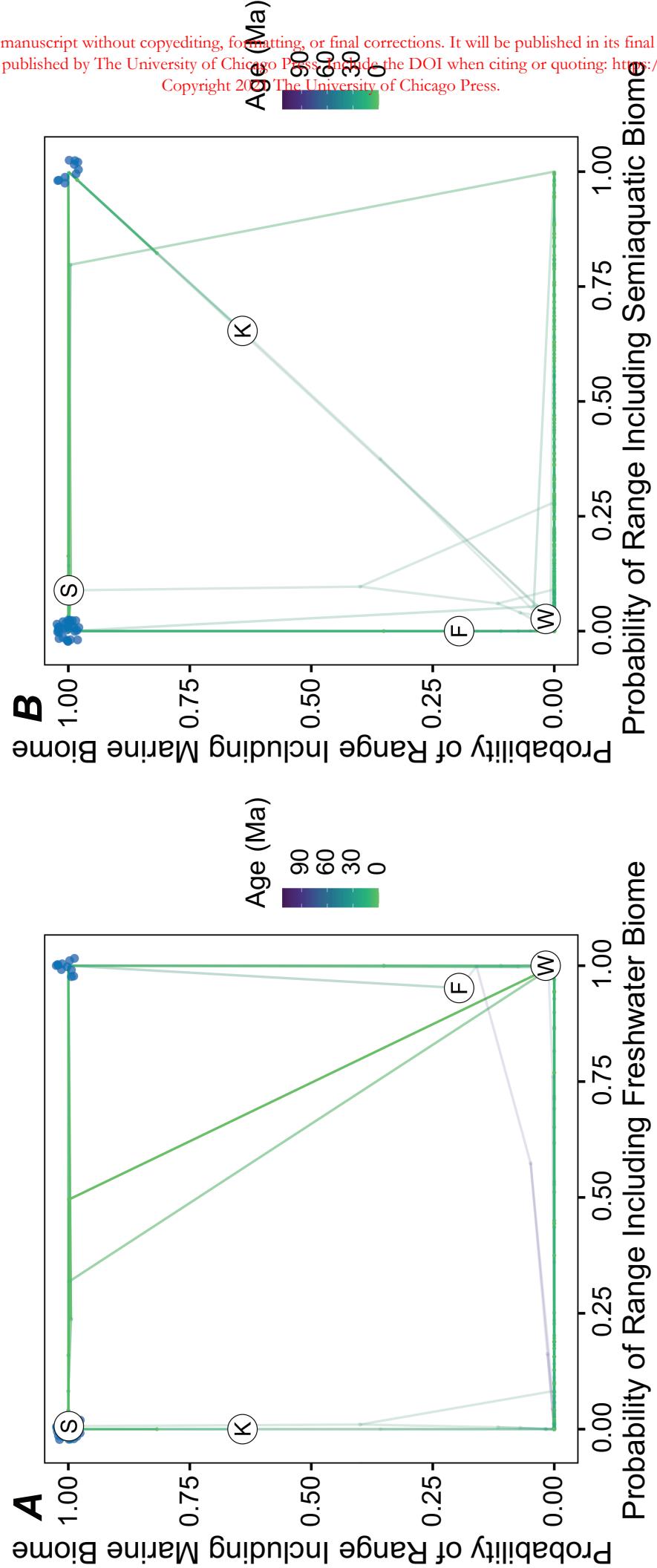


Figure 4

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: <https://doi.org/10.1086/716015>
Copyright 2021 The University of Chicago Press.

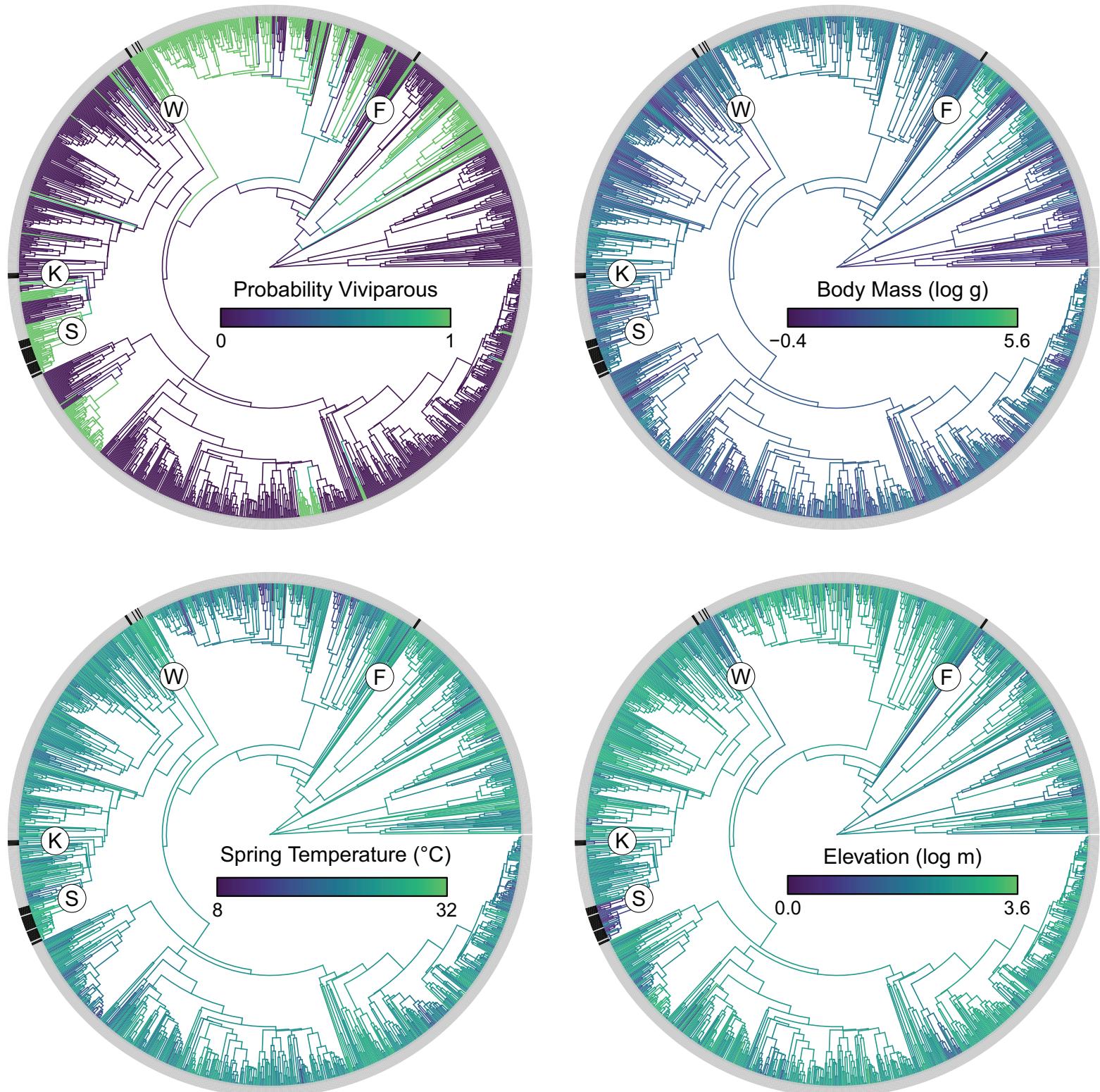


Figure 5

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press, include the DOI when citing or quoting: <https://doi.org/10.1086/716015>
 Copyright 2019 The University of Chicago Press.

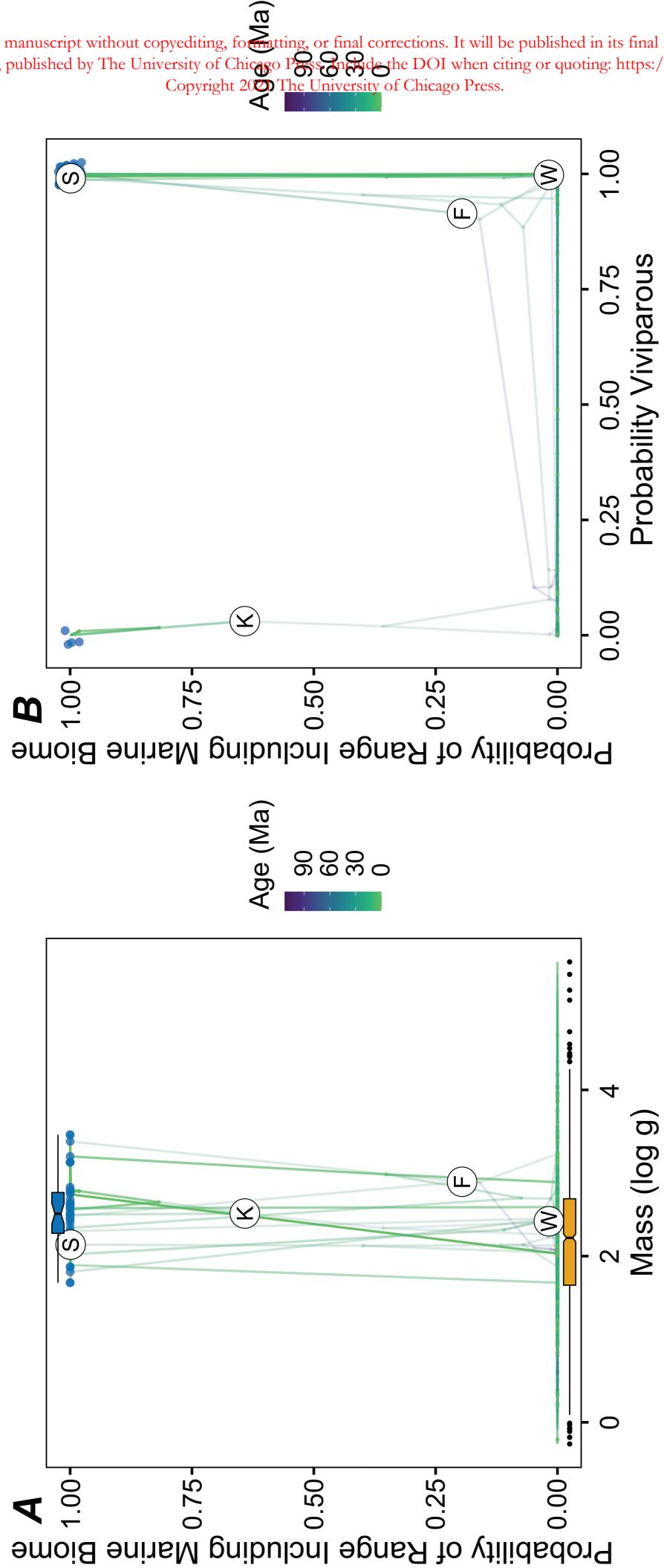


Figure 6

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Copyright 2015 The University of Chicago Press.

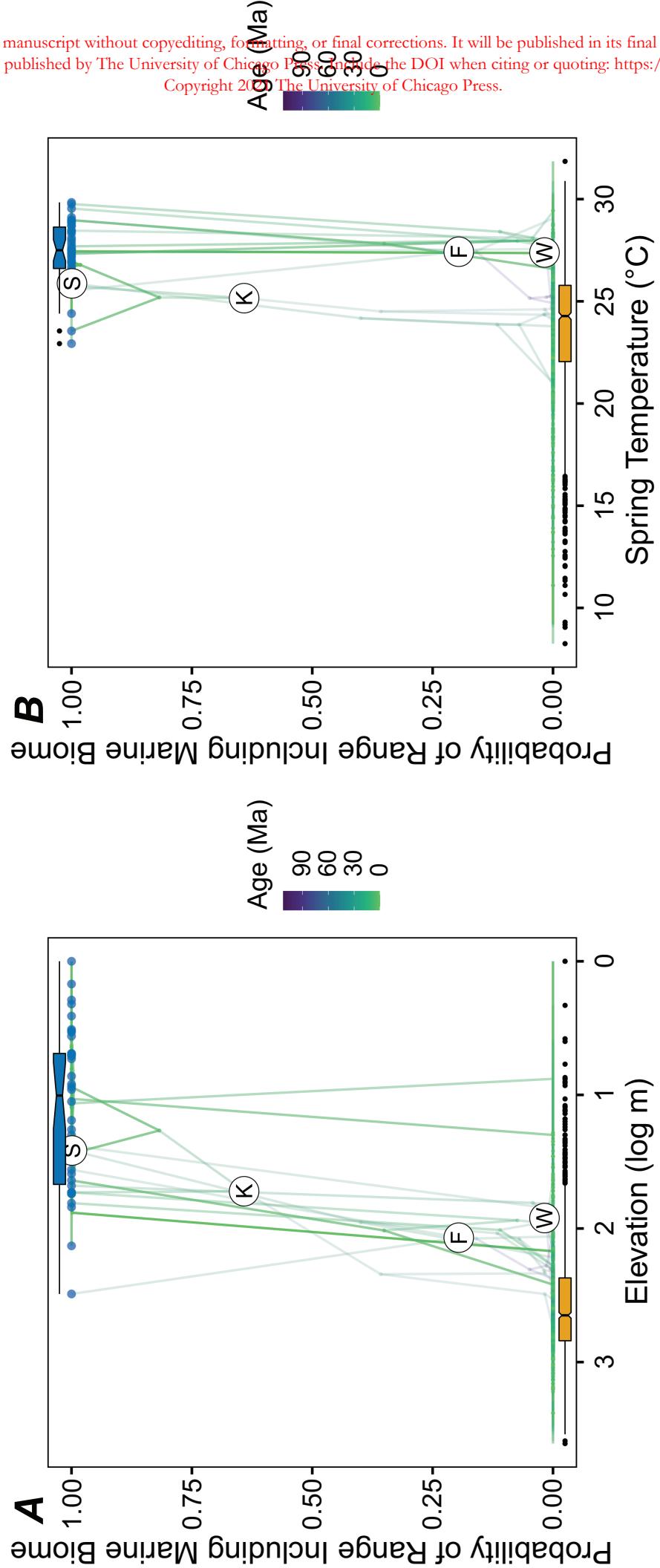
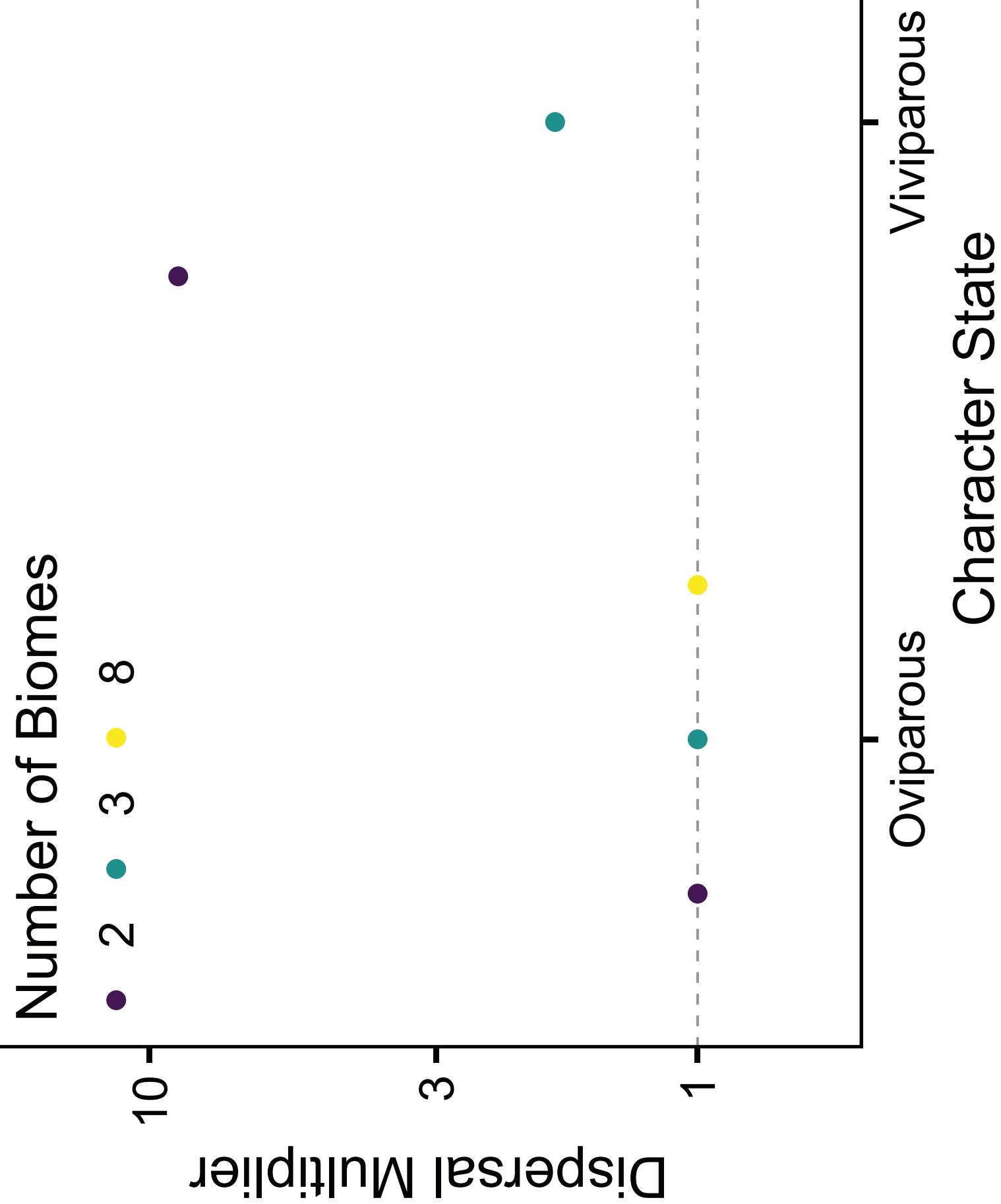


Figure 7

This is the author's accepted manuscript without copyediting, formatting, or final corrections. It will be published in its final form in an upcoming issue of The American Naturalist, published by The University of Chicago Press. Include the DOI when citing or quoting: <https://doi.org/10.1086/716015>

Copyright 2021 The University of Chicago Press.



Supplement to

Ecological filtering and exaptation in the evolution of marine snakes

The American Naturalist

William Gearty^{1,2,*}, Elsie Carrillo^{1,3}, and Jonathan L. Payne¹

¹Department of Geological Sciences, Stanford University, Stanford, CA 94305, USA

²School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

³Department of Ecology and Evolutionary Biology, University of California-Santa Cruz, Santa Cruz, CA 95064, USA

*Corresponding author (email: wgearty@unl.edu)

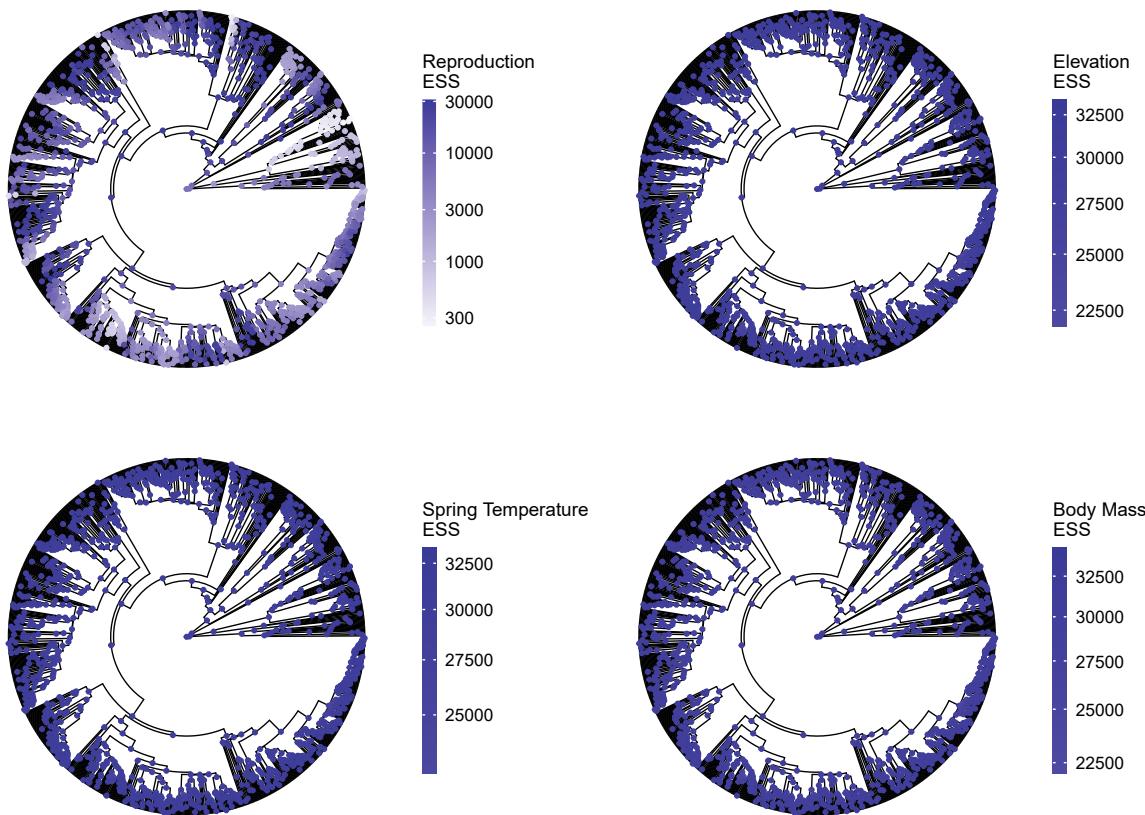


Figure S1: Effective sample sizes (ESS) for node ancestral state estimation with *threshml*.

Nodes are colored by ESS for each of the four co-estimated variables. An ESS value greater than 200 is generally accepted as indicating convergence. Note that roughly the same color scale is used for all four plots (but different limits are used for the legends).

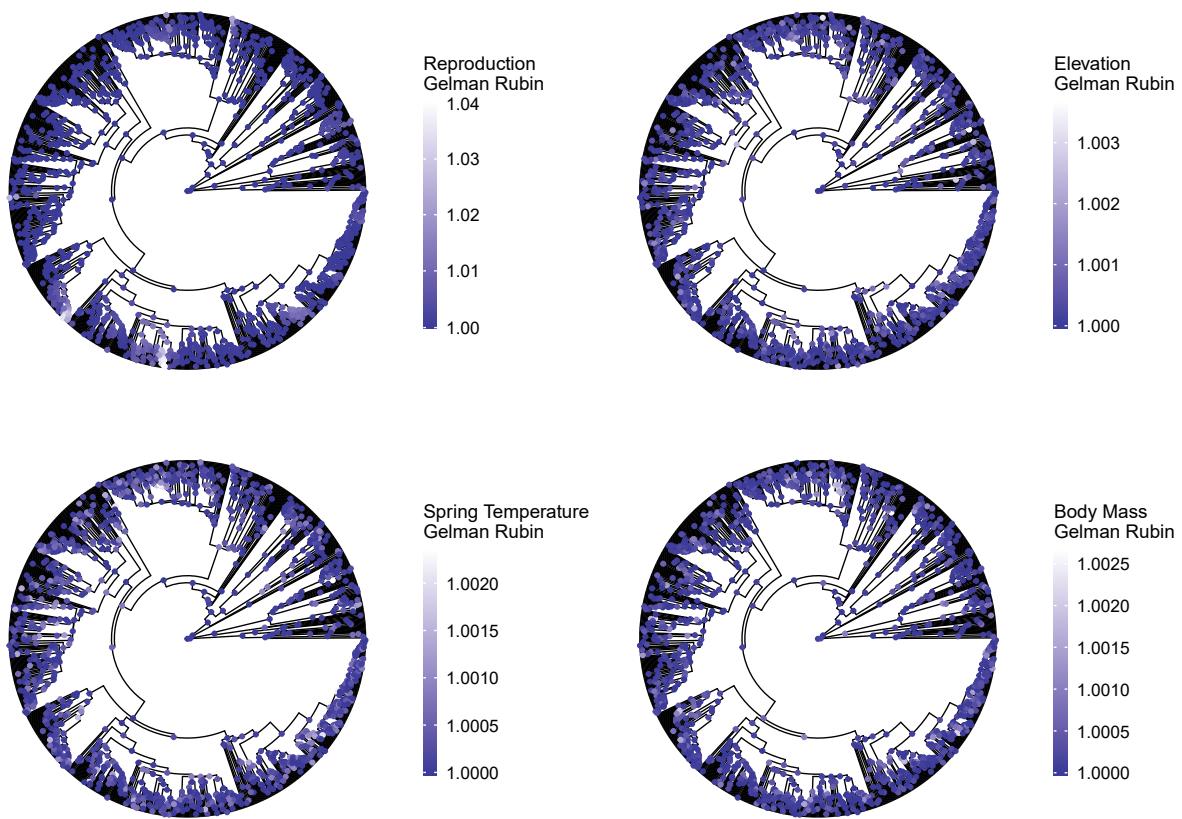


Figure S2: Gelman Rubin values for node ancestral state estimation with *threshml*. Nodes are colored by Gelman Rubin values for each of the four co-estimated variables. A Gelman Rubin value less than or equal to 1.1 is generally accepted as indicating convergence. Note that all scales have the same minimum value, but each has a different maximum value.

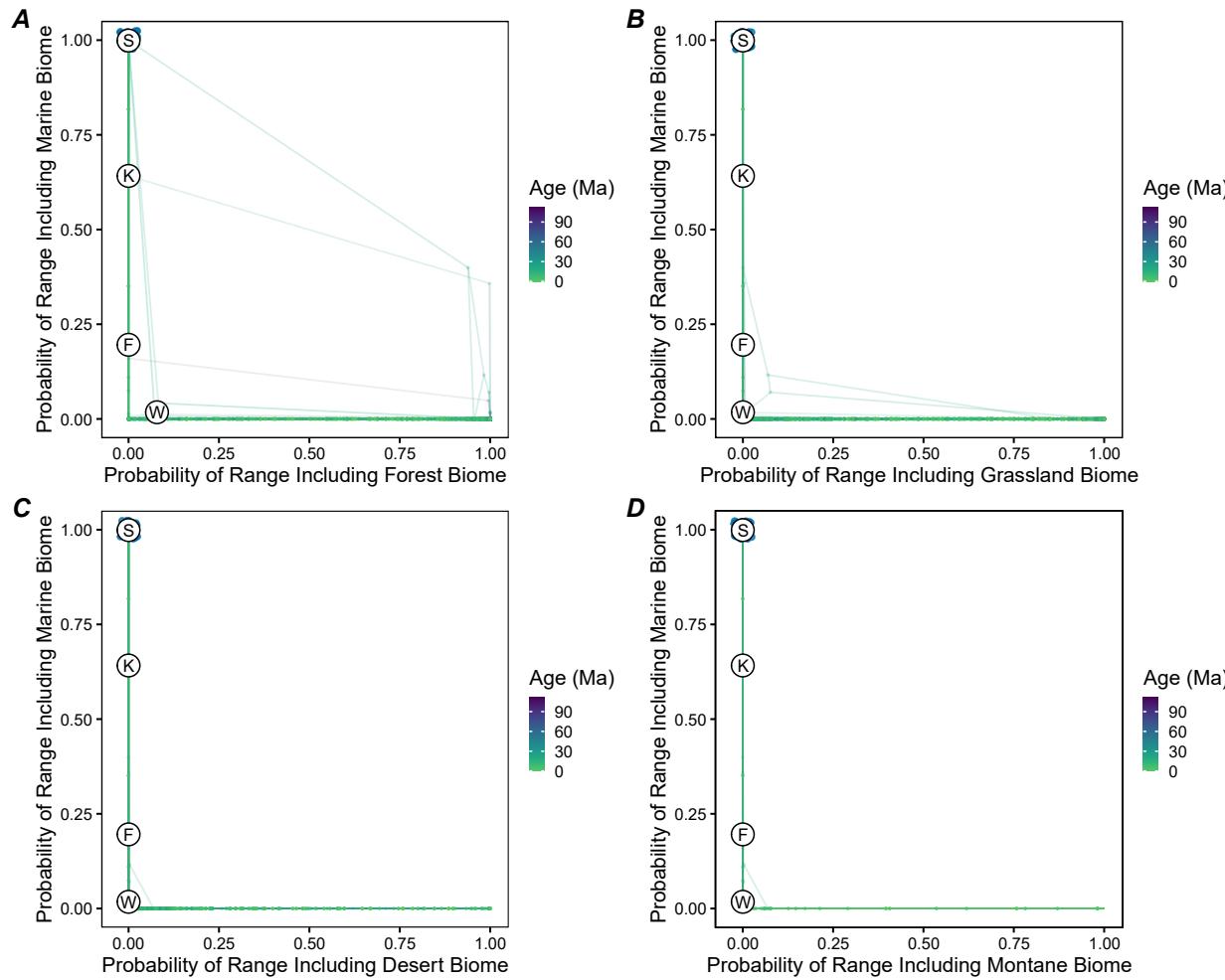


Figure S3: Phyloecospace of snakes for marine biome occupancy by forest (A), grassland (B), desert (C), and montane (D) biome occupancy. Lines and vertices represent the branches and nodes, respectively, of the phylogeny through time (color represents age) plotted by the probability of their estimated biome ranges including the marine biome and the probability of their estimated biome ranges including another biome. All probabilities were jointly estimated in the DEC+J model in *BioGeoBEARS*. Blue dots represent modern marine snakes with jitter to show point density. Node annotations as in Fig. 2.

								
Marine	Brackish	Freshwater	Semi-aquatic	Forest	Grassland	Desert	Montane	
# species (% of species)	38 (3.1%)	17 (1.4%)	92 (7.4%)	267 (21.5%)	921 (74.1%)	691 (55.6%)	221 (17.8%)	15 (1.2%)

Table S1: **Snake species biome occupancy.** Counts (and percentages) of snake species that occupy each of the defined biomes. A given species must occupy at least one biome but may occupy multiple biomes.

Model	LnL	N	d	e	j	AIC	AICc	AICc weight
DEC	-4308.112	2	0.007	0.000	0.000	8620.223	8620.233	0.009
DEC+J	-4302.432	3	0.007	0.000	0.002	8610.864	8610.884	0.991

Table S2: **Model comparison for BioGeoBEARS biome occupancy analysis results.** Best model based on AICc weight is highlighted in bold. Also shown are select model parameters: dispersal rate (*d*), local extirpation or range contraction (*e*), and jump-dispersal or founding event rate (*j*).

	Elevation (log m)	Temperature (°C)	Body Mass (log g)	Reproduction (liability)
Elevation (log m)	0.02962	-0.05971	0.00006	-0.00355
Temperature (°C)	-0.48869	0.73395	-0.00005	-0.19626
Body Mass (log g)	<i>0.01332</i>	-0.02848	0.03395	-0.01844
Reproduction (liability)	-0.09523	-0.17559	-0.19448	1.04640

Table S3: **Covariance matrix of traits as output by *threshml*.** Reproduction liability is used to calculate covariance (see Methods and Materials). Values in the lower diagonal (italicized) represent pairwise correlations for reference (calculated independently using only the measured species trait values).

Trait	Net Change			Rate of Change (/myr)		
	Average During Invasions	Average Outside of Invasions	p	Average During Invasions	Average Outside of Invasions	p
<i>Body Mass (log₁₀g)</i>	0.009	-0.004	0.46	0.006	-0.001	0.24
<i>Elevation (log₁₀m)</i>	-0.014	0.005	<0.001	-0.003	0.0003	0.008
<i>Temperature (°C)</i>	0.054	-0.021	0.45	0.017	-0.005	0.11
<i>Reproduction (P[viv])</i>	0.013	0.003	<0.001	0.002	0.001	<0.001

Table S4: **Estimated change along phylogenetic branches during marine invasions vs. outside of marine invasions.** Includes average net changes and average rates of change of various traits along individual phylogenetic branches during marine invasions (net change of marine biome occupancy probability > 0) versus outside of marine invasions. “p” columns represent p-values for Mann-Whitney U tests comparing the values for these two groups of phylogenetic branches. Significant P-values less than 0.05 are in bold.

Model	LnL	N	d	e	t12	t21	j	m1	m2	AIC	AICc	AICc weight
DEC+t12+t21	-325.531	4	0.001	0.000	0.002	0.006	0.000	1.000	1.000	659.062	659.094	0.000
DEC+J+t12+t21	-324.686	5	0.001	0.000	0.002	0.006	0.001	1.000	1.000	659.372	659.421	0.000
DEC+t12+t21+m2	-313.284	5	0.000	0.000	0.002	0.005	0.000	1.000	7.751	636.568	636.617	0.423
DEC+J+t12+t21+m2	-311.964	6	0.000	0.000	0.002	0.005	0.000	1.000	9.690	635.928	635.996	0.577

Table S5: Model comparison for BioGeoBEARS trait-dependent biome occupancy analysis results for reproductive mode with two biomes. Best model based on AICc weight is highlighted in bold. Also shown are select model parameters: dispersal rate (*d*), local extirpation or range contraction (*e*), jump-dispersal or founding event rate (*j*), transition rate from oviparity to viviparity (*t12*), transition rate from viviparity to oviparity (*t21*), multiplier representing effect of oviparity on dispersal (*m1*, fixed to 1), and multiplier representing effect of viviparity on dispersal (*m2*).

Model	LnL	N	d	e	t12	t21	j	m1	m2	AIC	AICc	AICc weight
DEC+t12+t21	-1374.393	4	0.008	0.000	0.002	0.006	0.000	1.000	1.000	2756.787	2756.819	0.000
DEC+J+t12+t21	-1374.474	5	0.008	0.000	0.002	0.006	0.000	1.000	1.000	2758.948	2758.997	0.000
DEC+t12+t21+m2	-1363.592	5	0.007	0.000	0.002	0.005	0.000	1.000	1.820	2737.183	2737.232	0.750
DEC+J+t12+t21+m2	-1363.681	6	0.007	0.000	0.002	0.005	0.000	1.000	1.821	2739.363	2739.431	0.250

Table S6: Model comparison for BioGeoBEARS trait-dependent biome occupancy analysis results for reproductive mode with three biomes. Best model based on AICc weight is highlighted in bold. Also shown are select model parameters: dispersal rate (*d*), local extirpation or range contraction (*e*), jump-dispersal or founding event rate (*j*), transition rate from oviparity to viviparity (*t12*), transition rate from viviparity to oviparity (*t21*), multiplier representing effect of oviparity on dispersal (*m1*, fixed to 1), and multiplier representing effect of viviparity on dispersal (*m2*).

Model	LnL	N	d	e	t12	t21	j	m1	m2	AIC	AICc	AICc weight
DEC+t12+t21	-4502.543	4	0.007	0.000	0.002	0.006	0.000	1.000	1.000	9013.087	9013.119	0.003
DEC+J+t12+t21	-4496.864	5	0.007	0.000	0.002	0.006	0.002	1.000	1.000	9003.728	9003.777	0.300
DEC+t12+t21+m2	-4500.573	5	0.006	0.000	0.002	0.005	0.000	1.000	1.173	9011.145	9011.194	0.007
DEC+J+t12+t21+m2	-4495.022	6	0.006	0.000	0.002	0.005	0.002	1.000	1.167	9002.043	9002.111	0.690

Table S7: Model comparison for BioGeoBEARS trait-dependent biome occupancy analysis results for reproductive mode with eight biomes. Best model based on AICc weight is highlighted in bold. Also shown are select model parameters: dispersal rate (*d*), local extirpation or range contraction (*e*), jump-dispersal or founding event rate (*j*), transition rate from oviparity to viviparity (*t12*), transition rate from viviparity to oviparity (*t21*), multiplier representing effect of oviparity on dispersal (*m1*, fixed to 1), and multiplier representing effect of viviparity on dispersal (*m2*).