ECOLOGY

Host susceptibility to snake fungal disease is highly dispersed across phylogenetic and functional trait space

Frank T. Burbrink,[1]* Jeffrey M. Lorch,[2] Karen R. Lips[3]

Emerging infectious diseases (EIDs) reduce host population sizes, cause extinction, disassemble communities, and have indirect negative effects on human well-being. Fungal EIDs have reduced population abundances in amphibians and bats across many species over large areas. The recent emergence of snake fungal disease (SFD) may have caused declines in some snake populations in the Eastern United States (EUS), which is home to a phylogenetically and ecologically diverse assembly of 98 taxa. SFD has been documented in only 23 naturally occuring species, although this is likely an underestimate of the number of susceptible taxa. Using several novel methods, including artificial neural networks, we combine phylogenetic and trait-based community estimates from all taxa in this region to show that SFD hosts are both phylogenetically and ecologically randomly dispersed. This might indicate that other species of snakes in the EUS could be currently infected or susceptible to SFD. Our models also indicate that information about key traits that enhance susceptiblity is lacking. Surveillance should consider that all snake species and habitats likely harbor this pathogen.

INTRODUCTION

Emerging fungal diseases have devastating effects on wildlife abundance and species diversity (1). Among terrestrial vertebrates, these fungal diseases have commonly been recorded in amphibians (2, 3), bats (4), and now snakes [snake fungal disease (SFD) caused by *Ophidiomyces ophidiodiicola*] (5, 6). The origins of these diseases and their effect on physiology, population demography, species loss, rates of colonization, and community assembly are important for mitigating their impacts on wildlife. Although it is immediately important to generate a list of susceptible taxa, unfortunately, predictions about how widespread particular diseases are across the tree of life are generally not well understood, likely owing to the time and expertise to survey broad regions and sample taxa with low encounter rates (7–9). Within amphibians and bats, the host range of fungal diseases affecting these groups may appear ecologically and taxonomically random at broad geographic and phylogenetic scales; however, nonrandom ecological and phylogenetic patterns have been reported locally (10–14). The broad host range exhibited by fungal pathogens is believed to play a major role in the catastrophic effects they can have on host communities (1). Understanding host ecological and evolutionary diversity is also complicated by cases where the same taxon may respond uniquely across different areas, given the timing of the disease invasion and environmental heterogeneity (15).

Predicting the diversity of taxa susceptible to any disease requires understanding the degree of host specificity. For example, even at the population level, specificity and genotype, evolutionary dynamics, demographics, and ecology of the host and genetic and phenotypic diversity of the pathogen all play an important role in determining how rapidly and to what extent a quickly evolving pathogen will spread through a population or across species (16, 17). Previously, it was considered that pathogens rarely cause species extinction, but both theoretical models and empirical evidence indicate that extinction is possible given the host abundance size and the presence of

reservoirs (18, 19). Extirpation of taxa in local communities can produce negative downstream consequences such as changing predator and prey relationships, altering habitat quality, enhancing community disassembly, or even affecting the prevalence of disease relevant to human health (7, 20–22).

Understanding the taxonomic extent of host susceptibility when a disease emerges regionally would allow us to predict (i) the capacity of the pathogen to infect host species not yet recorded, (ii) potential interactions among host species and disease transmission given the ecology, and (iii) whether the taxonomic identity of hosts is related to the dispersion across the phylogeny. With information about phylogenetic dispersion of hosts from a small number of observations, we can better understand whether infected taxa are restricted to specific clades having particular traits. If infected species are part of taxonomic groups occurring outside of the known range of the pathogen, then the disease could occur in related species in other areas, particularly if the disease was not constrained to the specific ecology of a host. In this case, ecology of the hosts would also be randomly dispersed, and the disease would be expected to spread over a broader range of habitats and across the phylogeny.

The emerging pathogen, *Ophidiomyces ophiodiicola*, that causes SFD has been recorded in wild snakes inhabiting much of the Eastern United States (EUS), which mostly includes the coastal, forested, and adjacent prairie habitats and dry areas east of the Continental Divide, although recently, the pathogen was detected on three species of snakes commonly found in Europe (23). The disease primarily affects the skin of snakes. Although molting may help resolve mild cases of SFD, conditions experienced by wild snakes may result in development of life-threatening infections (6). Death in wild snakes infected with SFD is likely related to multiple factors including physical damage to tissues and behavioral modifications that lead to increased risk of predation, environmental exposure, and starvation (6). This disease has been previously reported in 23 species from the EUS, although it is unclear how widely distributed *O. ophiodiicola* is over the taxonomy and phylogeny of snakes in this region. Snakes in the EUS represent a diverse group of colubroid families, including viperids, colubrines, and natricines, all sharing a common ancestor as recently as 50 million years ago (24).

[1]Department of Herpetology, American Museum of Natural History, 79th Street at Central Park West, New York, NY 10024, USA. [2]U.S. Geological Survey, National Wildlife Health Center, 6006 Schroeder Road, Madison, WI 53711, USA. [3]Department of Biology, University of Maryland, College Park, MD 20742, USA.
*Corresponding author. Email: fburbrink@amnh.org

As expected from this diversity, snakes in this region have a wide range of ecologies in primarily fossorial, terrestrial, or aquatic habitats (25). It is unclear whether *Ophidiomyces* can infect taxa with particular ecologies important to snakes, such as habitat, body size, reproductive mode, clutch size, diet, or diel pattern.

Snakes are generally secretive, and securing information on the extent of species-specific susceptibility of the disease from all taxa in the EUS is difficult. Specifically, morbidity, mortality, and population declines suspected to be associated with SFD have been documented predominantly in snake populations that were extensively monitored (5, 6, 26); however, little is known about species vulnerability across eastern North America.

Here, we examine the distribution of wild hosts infected with *Ophidiomyces* in a phylogenetic and ecological context to estimate dispersion. We use phylogenetic estimates at various spatial-phylogenetic scales covering the EUS, total United States (US), and the world and apply community phylogenetic methods and machine-learning approaches to examine several competing scenarios to understand how dispersed this disease is over trait and phylogenetic space: (i) In cases where infected taxa are phylogenetically and ecologically closely related, relative to a random sample of taxa, underdispersion is predicted, and the disease should only be confined to those taxa. (ii) Alternatively, if infected taxa are overdispersed both phylogenetically and ecologically, then only unrelated taxa should be susceptible to the disease when compared to a random distribution of species' relationships and ecologies. Instances where this would happen are likely rare, but it is possible that infected species are targeted to a narrow range of ecologies but are phylogenetically unrelated; these taxa are ecologically convergent (and underdispersed) but phylogenetically overdispersed, which suggests that only taxa with particular traits are susceptible. (iii) Closely related infected taxa may have very different ecologies, particularly where competitive exclusion or evolution into distinct niches occurs; here, taxa will be phylogenetically underdispersed but ecologically over-dispersed. (iv) The infected taxa could be completely random phylogenetically and ecologically, which would suggest that the disease is not targeting any particular group of snakes that have any specific traits.

RESULTS

Phylogenetic species variability

We first show that there is no phylogenetic signal for SFD-infected taxa at any spatial scale, where tree transformations yielding no phylogenetic signal (Pagel's $\lambda = 0$) are always significantly better supported (ΔAIC > 20; $P < 1.3 \times 10^{-6}$) than untransformed trees. We used three phylogenetic trees with taxa from (i) the EUS, (ii) all US species, and (iii) a large sampling of global taxa to estimate phylogenetic species variability for all taxa infected with SFD in the wild and those infected both in the wild and in laboratory settings relative to a random shuffling of infected taxa. We determined at each spatial scale (global, US, and EUS) that our real values of phylogenetic species variability (iPSV) were not lower than expected (underdispersed), suggesting that the 23 known infected taxa were not phylogenetically more closely related than a random sampling of 23 taxa (Fig. 1). For our infected taxa from the EUS, we estimated a mean iPSV at 0.633, which was not significantly over- or underdispersed relative to 1000 randomly shuffled groups (nPSV) retaining the same number of species of infected taxa (95% nPSV = 0.574 to 0.689; Fig. 1). We found that iPSV at 0.427 is phylogenetically underdispersed relative to both the global sample of wild infected taxa [nPSC 95% confidence interval (CI), 0.555 to 0.678] and a combination of wild- and laboratory-infected taxa (iPSV = 0.625, nPSC 95% CI, 0.572 to 0.680). At the scale of US, iPSV (0.230) was random (95% nPSC = 0.245 to 0.326).

Trait variability

We also examined how infected taxa in the EUS were dispersed with respect to ecologically important traits, including habitat preference,
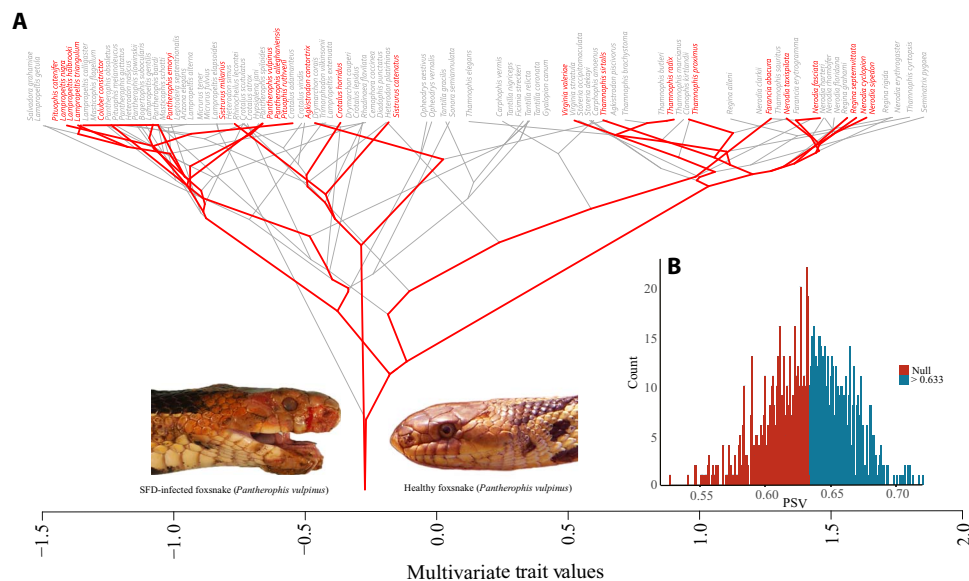

Fig. 1. (A) Traitgram of snakes found in the EUS, which projects the topology of the molecular phylogeny related to traits (on the *x* axis), showing SFD-infected taxa in red. **(B)** PSV density distribution of 1000 null communities (*n* = 23) assembled from the random samples taken from the 92 sampled taxa in EUS. The distribution colored blue indicates the proportion of the distribution above the actual PSV value (0.633) for the real SFD-infected community.

parity, average clutch size, diel pattern, diet, and maximum body size, relative to other species found in the EUS. Using all taxa and traits, we estimated three optimal multivariate clusters representing natural ecological/trait groupings using the average silhouette width criteria (Fig. 2). Similarly, the Duda-Hart test significantly rejected a single cluster in favor of >1 cluster. The clusters represented three groups of ecologically distinct taxa, which included the following types: (i) large-bodied terrestrial, (ii) aquatic, and (iii) small leaf-litter or fossorial snakes (see Fig. 2 and the Supplementary Materials). We found that infected taxa were not ecological outliers; individuals belong to groups 1, 2, and 3, and were not significantly different from their group medoids ($P > 0.05$) relative to other taxa in that category. Similarly, we found that ecological variance among the 23 infected taxa were not significantly smaller than 1000 random combinations of 23 taxa from the pool of uninfected taxa ($P = 0.11$).

Functional-phylogenetic metrics

Results examining the relative contribution of both traits and phylogeny to SFD susceptibility using functional-phylogenetic distance (FPDist) metrics showed a conserved pattern, where the trait distances increase with phylogenetic distances (Fig. 3). The effects of both traits and phylogeny, considered simultaneously, showed no pattern of significant under- or overdispersion relative to 1000 random simulations. This suggests that taxa with SFD were random with respect to the combined influence of traits and phylogeny.

Artificial neural networks

To understand whether the combination of trait and phylogenetic data can accurately classify infected taxa, we developed an optimal artificial neural network (NN) with 11 input neurons containing scaled trait and phylogenetic information, a single layer of 3 hidden neurons, and 1 output neuron that identified whether taxa were infected or un-infected. We sampled 70% of our total trait and phylogenetic data to train the NN and 30% to test accuracy. This procedure was replicated 100 times with randomly selected training and test data sets. Whereas all randomly sampled training data sets converged over all replicate data sets with model size = 1 and decay = 0, accuracy predicting infected taxa was limited (Fig. 4). To understand whether our NN can identify infected species, we developed additional tests. Using simulated data over the phylogenetic tree used here and under optimal conditions where there is a perfect phylogenetic and trait association with the disease, we show that the NN performs with 100% accuracy, given these new training and test data; here, accuracy, ΔA, measures the difference between true NN accuracy and shuffled infected-identity accuracy (Fig. 5). Furthermore, we demonstrate that when most of these simulated trait data are completely randomized and the remaining minority of traits were randomly shuffled by 0 to 40% of the original trait values, that accuracy remains high. Using our real data, we show that ΔA was centered on 0 and the AUC value was at 0.63, suggesting that our real data have little ability to predict infected taxa. Generally, body size was the most important predictor variable among replicates, although we note that
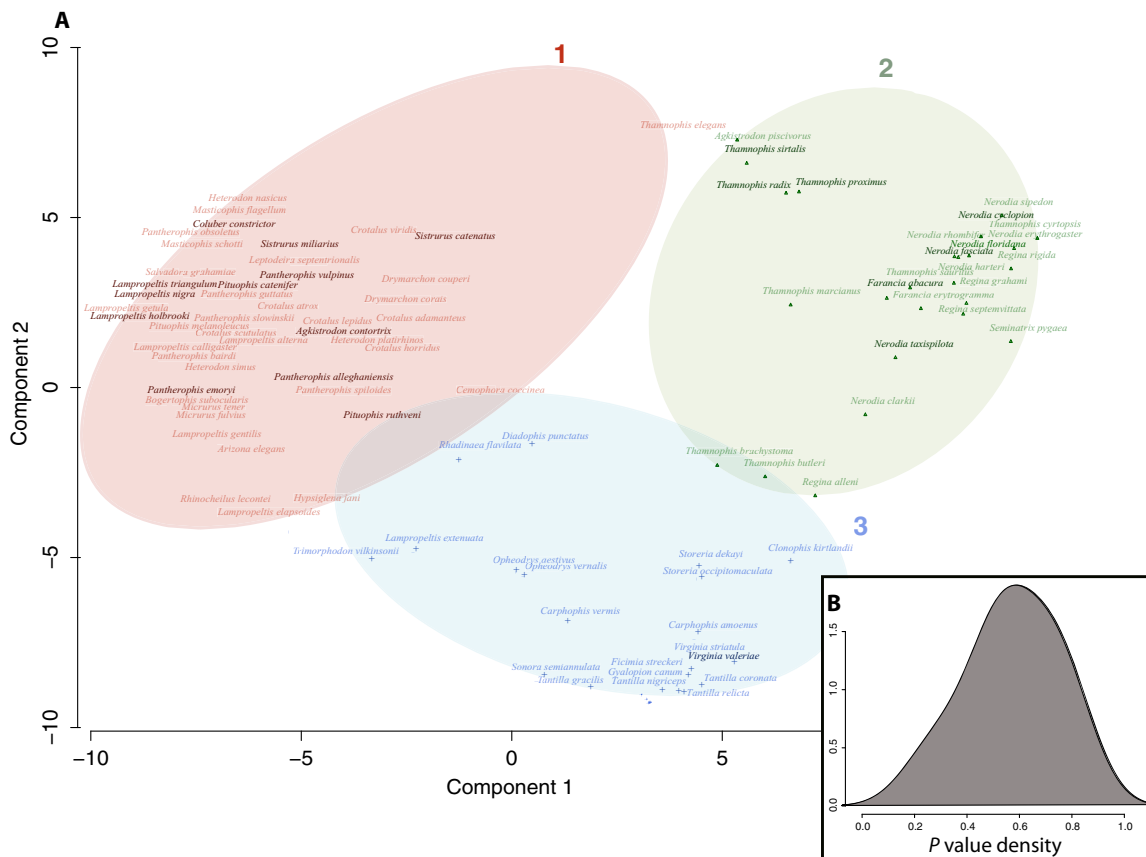
Fig. 2. (A) Multivariate clusters showing the three trait/ecological groups for snakes found in the EUS. Taxa infected with *O. ophidiodiicola* are bold within each group. **(B)** Distributions of *P* values showing the probability of having a Euclidian distance from each infected taxon to their group medoid are greater than those distances among all other uninfected taxa and medoid per group for all groups (see text). Infected taxa are generally not ecological outliers in any particular cluster.
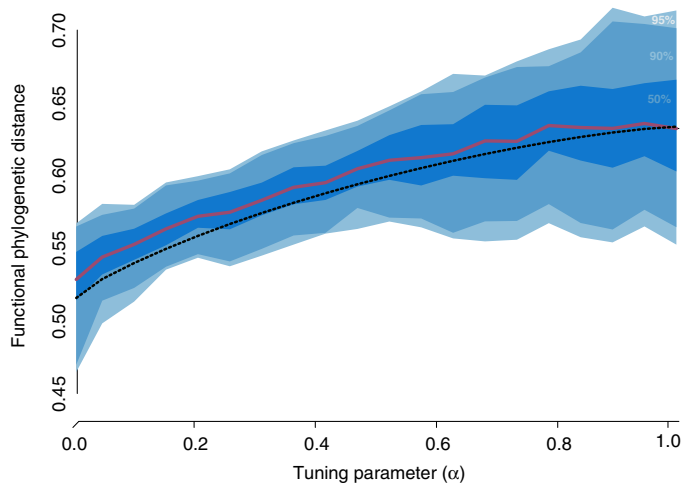
Fig. 3. FPDist for SFD-infected taxa, over all 20 values of the tuning parameter α, showing a conserved pattern where trait differences increase with phylogenetic differences. Here, α values closer to 0 indicate a strong contribution of traits, whereas values closer to 1.0 indicate a strong contribution of phylogeny, with intermediate values showing combined trait/phylogeny contributions. The shaded area is formed from 1000 null simulations with the average represented by the dark red solid line and the real SFD community represented by the black dashed line.

this may reflect the bias of larger infected taxa being more commonly observed. The populations of infected taxa were generally larger than those uninfected (*t* test; $t = -3.20$; $P = 0.002$).

Given that our conclusions suggest the possibility that all taxa may be infected, it then follows that the possibility of detecting SFD is biased by those taxa most commonly encountered by researchers. Summarizing the presence of each taxon across surveyed sites across the EUS (*25, 27*), logistic regression shows a significant association between commonality and binary presence of SFD ($P < 0.035$), although for every unit increase in commonality, log odds increased by only 0.019 (95% CI, 0.005 to 0.035). Furthermore, using the same phylogenetic relationships and trait data in our real study, but now with positive SFD paired with only the 23 most common taxa across sites in the EUS, we show that the ΔA metric is centered on 0.006, which was similar to results using real data.

DISCUSSION

Using a unique combination of phylogenetic methods that integrate ecology to understand the taxonomic range of an emerging infectious disease (EID), our results suggest that SFD is both phylogenetically and ecologically randomly dispersed in the EUS (Figs. 1 to 3). As currently known, taxa susceptible to infection by *Ophidiomyces* are likely a random sampling of the entire snake fauna of the EUS. Although it is not possible to predict which taxa are likely to be infected next, using phylogenetic relationships alone, it is possible that all 98 taxa in this region are susceptible to SFD. Although some wildlife diseases infect closely related species, in others, such as several fungal EIDs, the phylogenetic diversity of host taxa is quite large (*1, 28*).

Phylogenetic dispersion alone does not capture how the ecology of snakes relates to vulnerability to *Ophidiomyces*. Therefore, we demonstrate that taxa with SFD represent most of the basic ecological types of snakes found in the US (Figs. 2 and 3), which include snakes that are aquatic, terrestrial, fossorial, viviparous, oviparous, diurnal, nocturnal,
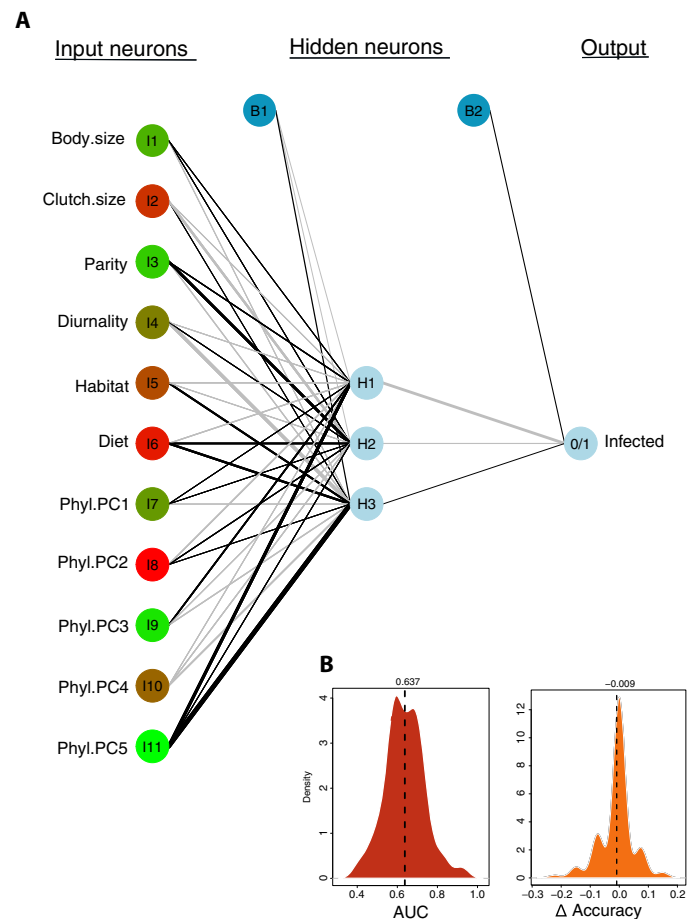

Fig. 4. (**A**) The preferred artifical NN model showing the following groups of input neurons (trait variables), the first layer of hidden neurons (H1 to H3), and the output variables predicting presence or absence of the diease (0/1). In addition, this NN diagram shows both intercept terms (bias), B1 and B2, which increases NN effeciency, and synaptic weights indicated by the thickness and bolding of lines connecting the neurons. (**B**) Density distributions and means (dashed lines) for the two NN metrics taken from 100 randomly chosen replicate data sets: (i) area under the receiver operator characteristic curve (AUC), where values >0.70 are considered reliable predictors; and (ii) difference between real and randomly shuffled accuracy (ΔA; see text), where values centered on zero indicate poor NN-predictive function.

communally denning during brumation, and solitary brumation, as well as those with diets composed of vertebrates, other snakes, and mostly invertebrates (*24*). Note that the known infected taxa have larger body sizes than the uninfected, although this estimate may represent observational bias in the field based on commonality of the species across sites in the EUS. Our methods therefore indicate that researchers should consider that SFD has the potential to infect a wide range of snake species in different ecologies and should not be biased by taxonomy and habitats. Because too little is understood about the epidemiology of SFD, it is unclear which traits, if any, influence susceptibility to *Ophidiomyces* infection. We show that few broad-scaled traits characterizing the major ecological groups of snakes in the EUS influence which species will be infected; our analyses suggest that many species not currently documented as being vulnerable are likely susceptible or already infected with *Ophidiomyces* (Figs. 1 and 2). Therefore, surveys of SFD should not be biased by taxon or ecology. In addition, given the spatial dispersion
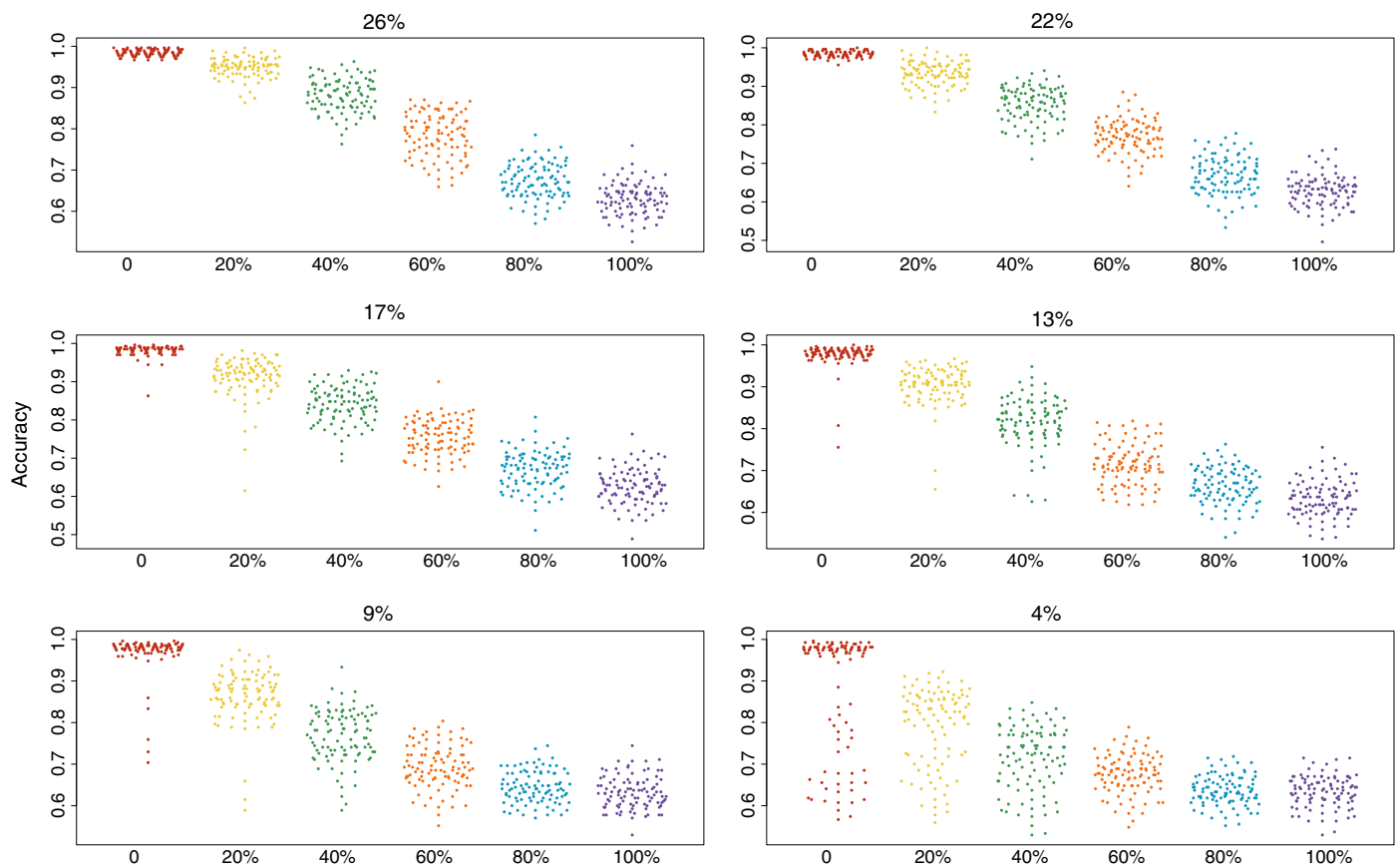
Fig. 5. Beeswarm plots showing the accuracy of the NN for 100 data simulations per category, where the percentage of trait variables were randomized by 0, 20, 40, 60, 80, and 100%, and the remaining traits were randomized by 100% (random noise). For example, in the upper left hand graph, 26% of variables were randomized by 0 to 100%, whereas the remaining 74% of variables were randomized completely by 100%.

of the disease (6), outbreaks could likely occur throughout this entire region. Unfortunately, the basic geographic information on the occurrence of SFD is derived from a few hosts; we therefore underscore that more knowledge about the basic biology and ecology of *Ophidiomyces* is necessary to predict the spatial range of this fungus. Without knowing how this disease is transmitted in the wild, it is imperative that care is taken when handling or transporting any species of snake found in this region. This region also contains several federally listed threatened or endangered taxa including *Drymarchon couperi*, *Nerodia clarkia taeniata*, *Nerodia erythrogaster neglecta*, and *Pituophis melanoleucus lodingi* as well as locally listed species throughout every state (see the Environmental Conservation Online System).

Over large geographic areas and deep evolutionary time, the mechanisms that influence the presence of EIDs include rates of dispersal and extinction, potential transmission by humans, and, in this study, the phylogenetic diversity of local hosts (9). At broader spatial scales, which here included the US or the world, we find these taxa to also be phylogenetically randomly dispersed or underdispersed, respectively. In the US, the most recent common ancestor among the infected taxa ranges from 50 to 75 million years ago (24). This does not suggest that the origins of SFD infections in snakes are this old, but rather, the disease can affect taxa from all groups that share a most recent common ancestor in the Tertiary or Late Cretaceous. The evolutionary history of the pathogen itself has not been examined, although a recent phylogeny of *Ophidiomyces* suggests that three main groups exist, two sister clades

in wild snakes in Europe and North America and one deeper clade represented by an isolate from a captive snake in the United Kingdom (23).

Nearly all known groups in the EUS have tested positive for SFD, which include the rattlesnakes (Viperidae), ratsnakes, milksnakes, kingsnakes and racers (two deep lineages of Colubrinae), and watersnakes and gartersnakes (Natricinae) (25). Two species of grasssnakes (*Natrix*; Natricinae) and one adder species (*Vipera*; Viperidae) recently tested positive for the pathogen in the United Kingdom and the Czech Republic (23). All of these European and North American families and subfamilies are part of Colubroidea, which excluding the remainder of Alethinophidia (for example, boas, pythons, sunbeam, and pipe and shieldtail snakes) and typhlopoids (blindsnakes) represent ~83% of all snakes (29). Although infections caused by *Ophidiomyces* have not been reported in the venomous coral snakes (Elapidae) found in the EUS, SFD has been associated with disease in a captive Australian elapid (30), suggesting species from the US in this family may also be vulnerable. Species from all of these infected subfamilies and families are found on every continent in the world (31), although the disease is only known in these groups from wild populations in EUS and Europe (5, 6, 23). Given the broad diversity of snakes infected, the range of biomes occupied, and that SFD can live outside of the host, it is plausible that *Ophidiomyces* is present globally or at least occurs throughout temperate regions of the world.

Our study mainly focused on snake species found in the EUS. However, *Ophidiomyces* may pose a disease risk to snake species on a global

level. From a phylogenetic perspective, if ecology of the host and geographic area are negligible to *Ophidiomyces*, then the disease should occur in colubroid species throughout the New World. Snakes from other groups host this disease in a captive setting (for example, Acrochordidae, Boidae, Elapidae, and Pythonidae) (*30*, *32*), although infected individuals from these groups have not, to our knowledge, been found under natural conditions. If snakes from these divergent families are also capable of developing SFD in the wild, then hosts for *Ophidiomyces* are much more dispersed at nearly the deepest phylogenetic scales, which would include 88% of all extant snakes in Alethinophidea (*29*). In addition, if the fungus (or particular strains of the fungus) is introduced outside of its native range and into naïve snake populations in other parts of the world, then the consequences could be as devastating, as evidenced by other vertebrate fungal diseases (*1*), particularly given the possibility that the disease could spread via the snake trade through sources originating among species found in the EUS.

Greater surveillance of SFD globally could enhance our understanding of the phylogenetic and ecological extent of the disease, as well as the possible areas of origin. A better understanding of the timing and area of origin and modes of transmission of SFD is needed to predict whether SFD will spread beyond known areas. Even in the US, it is unknown whether SFD was introduced in the last few decades from the pet trade, as the relatively recent confirmation of SFD in wild snakes would indicate, or was present for many generations, as field reports and large distances among outbreaks might suggest (*6*). Therefore, considering the intersection among phylogenetic, temporal, and spatial variables of the disease, natural history collections should be considered as a valuable resource for monitoring and modeling recent historical biogeography of SFD, provided accurate identification of *Ophidiomyces* can be rendered from these sources (*19*).

Although the NN shows low accuracy when predicting which particular species is susceptible to SFD, this may be because not all infected taxa are known to be infected, and key predicting variables are unknown. Alternatively, our NN simulations (Fig. 5) and functional-phylogenetic tests suggest that all snakes are likely susceptible to SFD regardless of phylogenetic relationship, associated traits, or primary ecology. To better understand which taxa will become infected and the probability of the disease emerging in other biomes and regions, more information is needed about the transmission of *Ophidiomyces* in the wild, pathogen population dynamics relative to changes in the host's ecology, phenotype, and genome, and the basic biological processes associated with exposure to new hosts, as well as potential factors related to host resistance (*33*).

Emerging wildlife diseases may have large effects on nontarget taxa. In some instances, the loss of wildlife may affect the community composition of other taxa or, with the loss of some groups, reduce essential wildlife services (*7*, *20*, *21*). Snakes are likely no different. They consume a massive variety of prey in the EUS and, for example, provide essential services like reducing populations of pests or acting as ecosystem monitors (*34*). In the EUS, snakes represent the dominant group of squamates (up to 100% of all species in some communities), and loss of these taxa would likely affect the structure of communities across multiple trophic levels.

Knowledge of how populations are affected by EIDs for each species and across communities and biomes is one of the primary goals for understanding the severity and long-term projections for disease management (*15*). Our predictions suggest that *Ophidiomyces* could likely infect more snake species in the EUS than is documented and that our knowledge of which species are infected is likely dependent on how

common they are in the community. Therefore, to advance our understanding of SFD, it is necessary to (i) screen a broad sampling of wild populations for the presence of *Ophidiomyces*; (ii) monitor susceptibility among taxa; (iii) estimate variance in detectability, intensity, and lethality among species and populations; (iv) understand the timing of SFD introduction (if applicable) and changes in lethality; (v) associate changes in infection intensity across seasons, climate, and biomes; and (vi) understand the extent to which the environment acts as a reservoir for the pathogen and how this interacts with snake hosts (*35*, *36*). Our analysis focused simply on susceptibility to SFD but did not account for potential differences in disease prevalence, severity, or population affects among species. Although SFD can be fatal and cause declines in some populations, individual recovery from the disease is also common (*6*); it is unknown how SFD affects survivability and recruitment across snake species and populations at a broader scale. Therefore, the effect of SFD on snake population demographics by host species is needed to understand whether *Ophidiomyces* reduces populations uniformly across taxa with different ecologies, physiologies, demographies, and migration rates. Furthermore, basic knowledge of transmission of this disease across taxa and biomes is needed to understand whether *Ophidiomyces* is likely contained within EUS and Europe or naturally occurs (or could be spread) throughout other biomes and continents.

MATERIALS AND METHODS
Identity of susceptible wild snakes
We used the list of susceptible species described in the study of Lorch *et al.* (*6*) for our analyses restricted to the EUS and the US. Because it is unclear whether snakes that developed SFD in captivity were exposed to conditions suitable for disease development in the wild, host taxa for which the disease has been documented only in a captive setting were excluded (table S1). We also added three additional species to the list of infected taxa based on recent submissions of snakes with SFD to the U.S. Geological Survey, National Wildlife Health Center (NWHC). These included *Lampropeltis holbrooki* (collected on 16 April 2016 from Elk County, KS, USA; NWHC case numbers 27242-004 and 27242-005), *Nerodia cyclopion* (collected on 03 March 2016 from Iberville Parish, LA, USA; NWHC case number 44736-093), and *Pantherophis emoryi* (collected on 16 April 2016 from Elk County, KS, USA; NWHC case numbers 27242-001, 27242-002, and 27242-003). For our global analysis, we included two species, *Natrix natrix* and *Natrix tesselatus*, recently discovered to be infected with *Ophidiomyces* in Europe (*23*). A third, *Vipera berus*, tested positive for SFD but remained uninfected. We also conducted another test where we included all infected taxa found both in the wild and in the laboratory, which included additional species *Acrochordus* sp., *Eunectes murinus*, *Boiga irregularis*, *Nerodia clarkia*, *Pantherophis guttatus*, *Hoplocephalus bungaroides*, *Python regius*, *Python sebae*, *Agkistrodon piscivorus*, and *Crotalus adamanteus*.

Phylogenetic trees
We determined whether infected snake taxa deviated from null phylogenetic models that include the pool of snakes from three main areas: (i) global, (ii) the US, and (iii) the EUS. For global distributions, we used the largest molecular phylogeny of snakes containing 1583 taxa sampled from all continents where snakes occur (*37*), and from the US, we used a phylogeny containing 144 resident species, which was then pruned to the 92 colubroid taxa that were sampled in the EUS

(text S1). Methods for the assembly of taxa and phylogenetic estimation using all 144 species were taken from the study of Burbrink and Myers (25). Note that this tree, dates, and support were similar to previous estimates where all of the main families, subfamilies, and tribes (Crotalinae, Natricinae, Dipsadinae, Lampropeltini, and Colubrinae) were well supported and genera are monophyletic (31).

Phylogenetic dispersion of SFD

We first determined whether there was a phylogenetic signal among taxa in the EUS, US, and global scales by transforming the phylogenies using Pagel's λ set to zero (no phylogenetic signal) and then fitting the infected data to these λ-transformed and untransformed trees using the R package GEIGER (38). Further, to examine phylogenetic diversity among the infected taxa at global, US, and EUS scales, we calculated the richness-independent measure of PSV (39) using the R package (40) Picante and the function phylostruct (41). To determine whether the infected taxa (iPSV) were phylogenetically over, under, or randomly dispersed relative to an infected sample of the same size, we generated 1000 random groups of taxa from the EUS, total US, and global pools of taxa from these trees while preserving species richness and phylogenetic relationships. We then calculated a null distribution of PSVs from these 1000 groups. We determined whether the infected taxa were overdispersed (iPSV > 95% of null groups) or underdispersed (iPSV < 95% of null groups) from each of the three regions.

Trait diversity data

To characterize the basic ecology of snakes in the US, we previously recorded habitat preference, parity, average clutch size, diel pattern, diet, and maximum body size from the study of Burbrink and Myers (25) (table S2). Body and clutch size were log-transformed, and presence or absence was scored for viviparity or oviparity, nocturnal or diurnal, and diel patterns. Dummy variables using 0/1 were coded for the habitat categories (terrestrial, aquatic, and fossorial) and diet (mammals, birds, lizards, snakes, frogs, salamanders, arthropods, worms, and gastropods). Note that these metrics provide a standard estimate for basic ecologically relevant traits in snakes (25).

Trait dimensionality and SFD

We determined whether the known infected taxa were ecologically different from the taxa not confirmed to have SFD. To reduce the dimensionality of this trait data set, which includes both continuous and categorical traits, we calculated all pairwise dissimilarities using Gower distances (42) in the R package Cluster (43). With these distances, we estimated the number of distinct clusters around medoids using the optimum average silhouette width (44) with the pamk function in Cluster. This test also determined the optimal number of clusters (k = 2 to 15), and with the Duda-Hart test (45) we estimated whether a single cluster for these data is likely.

With these clusters, we determined whether the distance of each infected taxon to their particular cluster medoid was significantly greater than the distances from all other taxa in that cluster. This provides another level of evidence about how closely fit the infected taxa are to their cluster.

In addition, we also determined whether the total trait variance for the infected taxa was smaller than a random sample of the uninfected taxa while preserving the same richness as the infected taxa (n = 23). We converted Gower distances to the principal coordinate scores for each taxon using principal component decomposition with the pcoa function in ape (46). We reconstructed 1000 data sets of 23 randomly selected uninfected taxa and calculated variance to determine how often this value was larger than the variance from the infected taxa; the expectation was that if infected taxa were ecologically more closely related, then their overall trait variance should be significantly lower than the variance of the uninfected taxa.

Relative influence of trait and phylogenetic diversity on occurrence of SFD

Patterns of dispersion by phylogeny or ecologically important traits alone may not be helpful in determining which of the two factors are relatively driving susceptibility to SFD. Therefore, we examined the relative contribution of traits and phylogeny within the known SFD-infected taxa. We first generated a traitgram (47), which is a phylogeny of all snakes in the EUS with known infected taxa highlighted but scales branch lengths and position against trait values.

We determined how significant the relative influence of traits and phylogeny on the SFD-infected taxa was by estimating FPDist using both phylogenetic distance (PDist) and trait functional distance (FDist), calculated as the Euclidian distance among taxa across all traits (47). We related these two factors using the equation FPDist = $(\alpha \text{PDist}^p + (1 - \alpha) \text{FDist}^P)^{1/p}$ to determine the relative contribution of the phylogenetic (PDist) and the functional (FDist) component via a tuning parameter α, which ranges from 0, where FPDists are dominated by trait distances, to 1, where FPDists are dominated by phylogenetic distance. We determined, over 20 incremental increases of α from 0 to 1, whether functional or phylogenetic under- or overdispersion exists. We then compared these to 1000 randomized simulations that, given significance, reveal the relative contribution of traits and phylogeny regarding SFD susceptibility. For example, α values between 0 and 1 indicate that both functional and phylogenetic contributions were important for SFD susceptibility. Here, FPDist will decrease with an increase in α for phylogenetically diverging taxa showing a necessary decreasing contribution of traits and increasing contribution of phylogeny. The metric increases with α where SFD occurs in convergently evolving taxa due to an increasing contribution of phylogeny [see Fig. 1 in Cadotte et al. (47)]. This may indicate a situation where none of the scored characters reflect real functional differences among the SFD-infected taxa, but that phylogeny is suggestive that some undetermined traits may predict susceptibility. For functional underdispersion, traits should be more similar than expected, suggesting that SFD was infecting taxa with a particular trait related to a particular ecology. By contrast, it was expected that in a situation where phylogenetic underdispersion with a lack of functional underdispersion occurs, taxa with SFD should be phylogenetically more closely related compared to a random assemblage of taxa. For functional overdispersion in this community, traits reflect greater differences than expected, whereas for phylogenetic overdispersion with a lack of functional-trait overdispersion, it was expected that taxa with SFD were phylogenetically less closely related than expected. A mixture of phylogenetic and functional over- and underdispersion or null patterns was possible within a community. Where nonsignificance occurs at any value of α, we predicted that the infected taxa were random with respect to both phylogeny and traits.

Predicting occurrence of SFD using artificial NNs

To determine whether the combination trait and phylogenetic data can accurately classify infected taxa, we used artificial NNs, a type of machine-learning method that does not require prespecified relationships between covariate predictor and response variables and does not assume that the data were distributed in any particularly way (48).

The NN method will allow us to maximize classification accuracy of this disease in snakes (0, uninfected; 1, infected) by taking ecological and phylogenetic information as input neurons connected via synapses to hidden neurons and then output a classification neuron. The NN will determine the relationship between trait and phylogenetic variables and presence of disease by taxon. For this data set where the identity of all infected taxa in the wild likely remains unknown, NN will not be negatively affected by unknown interactions among variables and collinearity among data arising from sampling bias (48).

We first converted the predicted binary habitat, trait, and diet data into Gower distances and transformed these using pcoa, and scores from each primary axis were associated with each taxon. Similarly, phylogenetic distances were transformed using pcoa, and scores from the first five axes were associated with each taxon (see previous section). The predictor data set—composed of phylogenetic, body and clutch size, habitat, and dietary variables—were scaled from 0 to 1 for each variable. From this, 70% were randomly chosen for training the NN and 30% for testing. With the R package caret, we generated an NN model using the train function with 25 bootstrapped replicates to assess decay and accuracy for model convergence over one, two, and three layers of hidden neurons (49). We tested accuracy using the predict and postResample functions for our best model with the 30% retained samples. To determine the robustness of our model, we resampled the testing and training procedure 100 times, estimated the accuracy of predicting infected taxa for each sample, and subtracted this from a random shuffling of infected taxon identity (ΔA). A distribution of ΔA centered on 0 indicates that the model has little predictive power given the input neurons. Similarly, we predicted AUC to determine power of this model, where values lower than 0.7 indicate little ability to predict infected taxa.

To understand whether this NN could detect taxa with the disease over various optimal conditions, we simulated the disease containing the same number of taxa infected ($n = 23$) over the phylogeny used previously. Here, we chose the infected taxa to be the most inclusive species of a clade, therefore making a perfect association between phylogenetic relationship and disease. We then simulated continuous data and placed the same number of binary traits used in our real study on this tree to also have a perfect phylogenetic association with the disease. The continuous traits on this tree were simulated using fastBM in phytools (50). Using these variables, we tested the accuracy of the NN under conditions where 1 (4%), 2 (9%), 3 (13%), 4 (17%), 5 (22%), or 6 (26%) trait variables were randomly chosen, and data from each variable were shuffled by 0, 20, 40, 60, 80, and 100%; the remaining trait variables (74 to 96%) were randomized by 100% (random noise). All data sets and tests were replicated 100 times for each set of variables (1 to 6) and randomizing percentages.

If all taxa were infected and if detecting SFD was biased by those taxa most commonly encountered by researchers, then we expect to find a significant association bias between the disease and commonality. We summarized the presence of each taxon across surveyed sites across the EUS using data taken from two previous studies (25, 27) and performed a logistic regression using R where the binary presence of the disease was predicted by commonality across site. In addition, to determine whether unknown commonality yields similar NN predictions as with our real data, we replaced the identity of the infected and uninfected taxa using our real trait and phylogenetic data matrix; we then rescored the 23 most common taxa as being infected. We ran these NN 100 times and estimated ΔA.

REFERENCES AND NOTES

1. M. C. Fisher, D. A. Henk, C. J. Briggs, J. S. Brownstein, L. C. Madoff, S. L. McCraw, S. J. Gurr, Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**, 186–194 (2012).
2. L. Berger, R. Speare, P. Daszak, D. E. Green, A. A. Cunningham, C. L. Goggin, R. Slocombe, M. A. Ragan, A. D. Hyatt, K. R. McDonald, H. B. Hines, K. R. Lips, G. Marantelli, H. Parkes, Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9031–9036 (1998).
3. A. Martel, A. Spitzen-van der Sluijs, M. Blooi, W. Bert, R. Ducatelle, M. C. Fisher, A. Woeltjes, W. Bosman, K. Chiers, F. Bossuyt, F. Pasmans, *Batrachochytrium salamandrivorans* sp. nov. causes lethal chytridiomycosis in amphibians. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15325–15329 (2013).
4. D. S. Blehert, A. C. Hicks, M. Behr, C. U. Meteyer, B. M. Berlowski-Zier, E. L. Buckles, J. T. H. Coleman, S. R. Darling, A. Gargas, R. Niver, J. C. Okoniewski, R. J. Rudd, W. B. Stone, Bat white-nose syndrome: An emerging fungal pathogen? *Science* **323**, 227 (2009).
5. M. C. Allender, D. B. Raudabaugh, F. H. Gleason, A. N. Miller, The natural history, ecology, and epidemiology of *Ophidiomyces ophiodiicola* and its potential impact on free-ranging snake populations. *Fungal Ecol.* **17**, 187–196 (2015).
6. J. M. Lorch, S. Knowles, J. S. Lankton, K. Michell, J. L. Edwards, J. M. Kapfer, R. A. Staffen, E. R. Wild, K. Z. Schmidt, A. E. Ballmann, D. Blodgett, T. M. Farrell, B. M. Glorioso, L. A. Last, S. J. Price, K. L. Schuler, C. E. Smith, J. F. Wellehan Jr., D. S. Blehert, Snake fungal disease: An emerging threat to wild snakes. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150457 (2016).
7. P. Daszak, A. A. Cunningham, A. D. Hyatt, Emerging infectious diseases of wildlife—Threats to biodiversity and human health. *Science* **287**, 443–449 (2000).
8. F. Keesing, R. D. Holt, R. S. Ostfeld, Effects of species diversity on disease risk. *Ecol. Lett.* **9**, 485–498 (2006).
9. G. Suzán, G. E. García-Peña, I. Castro-Arellano, O. Rico, A. V. Rubio, M. J. Tolsá, B. Roche, P. R. Hosseini, A. Rizzoli, K. A. Murray, C. Zambrana-Torrelio, M. Vittecoq, X. Bailly, A. A. Aguirre, P. Daszak, A.-H. Prieur-Richard, J. N. Mills, J. F. Guégan, Metacommunity and phylogenetic structure determine wildlife and zoonotic infectious disease patterns in time and space. *Ecol. Evol.* **5**, 865–873 (2015).
10. A. J. Crawford, K. R. Lips, E. Bermingham, Epidemic disease decimates amphibian abundance, species diversity, and evolutionary history in the highlands of central Panama. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13777–13782 (2010).
11. S. J. Corey, T. A. Waite, Phylogenetic autocorrelation of extinction threat in globally imperilled amphibians. *Divers. Distrib.* **14**, 614–629 (2007).
12. V. Baláž, J. Vörös, P. Civiš, J. Vojar, A. Hettyey, E. Sós, R. Dankovics, R. Jehle, D. G. Christiansen, F. Clare, M. C. Fisher, T. W. J. Garner, J. Bielby, Assessing risk and guidance on monitoring of *Batrachochytrium dendrobatidis* in Europe through identification of taxonomic selectivity of infection. *Conserv. Biol.* **28**, 213–223 (2014).
13. K. G. Smith, K. R. Lips, J. M. Chase, Selecting for extinction: Nonrandom disease-associated extinction homogenizes amphibian biotas. *Ecol. Lett.* **12**, 1069–1078 (2009).
14. J. Zukal, H. Bandouchova, T. Bartonicka, H. Berkova, V. Brack, J. Brichta, M. Dolinay, K. S. Jaron, V. Kovacova, M. Kovarik, N. Martínková, K. Ondracek, Z. Rehak, G. G. Turner, J. Pikula, White-nose syndrome fungus: A generalist pathogen of hibernating bats. *PLOS ONE* **9**, e97224 (2014).
15. K. Lips, Overview of chytrid emergence and impacts on amphibians. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150465 (2016).
16. A. E. Savage, K. R. Zamudio, Adaptive tolerance to a pathogenic fungus drives major histocompatibility complex evolution in natural amphibian populations. *Proc. R. Soc. B Biol. Sci.* **283**, 20153115 (2016).
17. E. B. Rosenblum, T. Y. James, K. R. Zamudio, T. J. Poorten, D. Ilut, D. Rodriguez, J. M. Eastman, K. Richards-Hrdlicka, S. Joneson, T. S. Jenkinson, J. E. Longcore, G. Parra Olea, L. F. Toledo, M. L. Arellano, E. M. Medina, S. Restrepo, S. V. Flechas, L. Berger, C. J. Briggs, J. E. Stajich, Complex history of the amphibian-killing chytrid fungus revealed with genome resequencing data. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9385–9390 (2013).

18. H. McCallum, Disease and the dynamics of extinction. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2828–2839 (2012).

19. K. R. Lips, F. Brem, R. Brenes, J. D. Reeve, R. A. Alford, J. Voyles, C. Carey, L. Livo, A. P. Pessier, J. P. Collins, Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3165–3170 (2006).

20. P. T. J. Johnson, D. L. Preston, J. T. Hoverman, K. L. D. Richgels, Biodiversity decreases disease through predictable changes in host community competence. *Nature* **494**, 230–233 (2013).

21. R. S. Ostfeld, K. LoGiudice, Community disassembly, biodiversity loss, and the erosion of an ecosystem service. *Ecology* **84**, 1421–1427 (2003).

22. F. Keesing, L. K. Belden, P. Daszak, A. Dobson, C. D. Harvell, R. D. Holt, P. Hudson, A. Jolles, K. E. Jones, C. E. Mitchell, S. S. Myers, T. Bogich, R. S. Ostfeld, Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**, 647–652 (2010).

23. L. H. V. Franklinos, J. M. Lorch, E. Bohuski, J. Rodriguez-Ramos Fernandez, O. N. Wright, L. Fitzpatrick, S. Petrovan, C. Durrant, C. Linton, V. Baláž, A. A. Cunningham, B. Lawson, Emerging fungal pathogen *Ophidiomyces ophiodiicola* in wild European snakes. *Sci. Rep.* **7**, 3844 (2017).

24. R. A. Pyron, F. T. Burbrink, Extinction, ecological opportunity, and the origins of global snake diversity. *Evolution* **66**, 163–178 (2012).

25. F. T. Burbrink, E. A. Myers, Both traits and phylogenetic history influence community structure in snakes over steep environmental gradients. *Ecography* **38**, 1036–1048 (2015).

26. R. W. Clark, M. N. Marchand, B. J. Clifford, R. Stechert, S. Stephens, Decline of an isolated timber rattlesnake (*Crotalus horridus*) population: Interactions between climate change, disease, and loss of genetic diversity. *Biol. Conserv.* **144**, 886–891 (2011).

27. F. T. Burbrink, A. D. McKelvy, R. A. Pyron, E. A. Myers, Predicting community structure in snakes on Eastern Nearctic islands using ecological neutral theory and phylogenetic methods. *Proc. R. Soc. B Biol. Sci.* **282**, 20151700 (2015).

28. K. LoGiudice, R. S. Ostfeld, K. A. Schmidt, F. Keesing, The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 567–571 (2003).

29. P. Uetz, The Reptile Database (2009); available at www.reptile-database.org.

30. L. Sigler, S. Hambleton, J. A. Paré, Molecular characterization of reptile pathogens currently known as members of the chrysosporium anamorph of Nannizziopsis vriesii complex and relationship with some human-associated isolates. *J. Clin. Microbiol.* **51**, 3338–3357 (2013).

31. F. T. Burbrink, B. I. Crother, Evolution and taxonomy of snakes, in *Reproductive Biology and Phylogeny of Snakes*, D. M. Sever, R. D. Aldridge, Eds. (Science Publishers, 2011), pp. 19–54.

32. J. A. Paré, L. Sigler, K. L. Rypien, C. F. C. Gibas, Cutaneous mycobiota of captive squamate reptiles with notes on the scarcity of Chrysosporium anamorph of Nannizziopsis vriesii. *J. Hepetol. Med. Surg.* **13**, 10–15 (2003).

33. M. E. J. Woolhouse, D. T. Haydon, R. Antia, Emerging pathogens: The epidemiology and evolution of species jumps. *Trends Ecol. Evol.* **20**, 238–244 (2005).

34. S. J. Mullin, R. A. Seigel, *Snakes: Ecology and Conservation* (Cornell Univ. Press, ed. 1, 2009).

35. F. De Castro, B. Bolker, Mechanisms of disease-induced extinction. *Ecol. Lett.* **8**, 117–126 (2004).

36. E. H. Campbell Grant, E. Muths, R. A. Katz, S. Canessa, M. J. Adams, J. R. Ballard, L. Berger, C. J. Briggs, J. T. H. Coleman, M. J. Gray, M. C. Harris, R. N. Harris, B. Hossack, K. P. Huyvaert, J. Kolby, K. R. Lips, R. E. Lovich, H. I. McCallum, J. R. Mendelson III, P. Nanjappa, D. H. Olson, J. G. Powers, K. L. D. Richgels, R. E. Russell, B. R. Schmidt, A. Spitzen-van der Sluijs, M. K. Watry, D. C. Woodhams, C. L. White, Using decision analysis to support proactive management of emerging infectious wildlife diseases. *Front. Ecol. Environ.* **15**, 214–221 (2017).

37. J. F. R. Tonini, K. H. Beard, R. B. Ferreira, W. Jetz, R. A. Pyron, Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biol. Conserv.* **204**, 23–31 (2016).

38. L. J. Harmon, J. T. Weir, C. D. Brock, R. E. Glor, W. Challenger, GEIGER: Investigating evolutionary radiations. *Bioinformatics* **24**, 129–131 (2008).

39. M. R. Helmus, T. J. Bland, C. K. Williams, A. R. Ives, Phylogenetic measures of biodiversity. *Am. Nat.* **169**, E68–E83 (2007).

40. R. Core Team, R: A Language and Environment for Statistical Computing (2015); available at www.r-project.org/.

41. S. W. Kembel, P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, C. O. Webb, Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).

42. J. C. Gower, A general coefficient of similarity and some of its properties. *Biometrics* **27**, 857–871 (1971).

43. M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, Cluster: Cluster analysis basics and extensions, R package version 2.0.6 (2016).

44. C. Hennig, T. F. Liao, How to find an appropriate clustering for mixed-type variables with application to socio-economic stratification. *J. R. Stat. Soc. Series C Appl. Stat.* **62**, 309–369 (2013).

45. R. O. Duda, P. E. Hart, *Pattern Classification and Scene Analysis* (Wiley, 1973).

46. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).

47. M. Cadotte, C. H. Albert, S. C. Walker, The ecology of differences: Assessing community assembly with trait and evolutionary distances. *Ecol. Lett.* **16**, 1234–1244 (2013).

48. W. Zhang, *Computational Ecology: Artificial Neural Networks and Their Applications* (World Scientific, 2010).

49. M. Kuhn, Caret:Classification and Regression Training Package, R package version 6.0-77, **28** (2008).

50. L. J. Revell, Phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

Citation: F. T. Burbrink, J. M. Lorch, K. R. Lips, Host susceptibility to snake fungal disease is highly dispersed across phylogenetic and functional trait space. *Sci. Adv.* **3**, e1701387 (2017).

Science Advances

Host susceptibility to snake fungal disease is highly dispersed across phylogenetic and functional trait space

Frank T. Burbrink, Jeffrey M. Lorch and Karen R. Lips

ARTICLE TOOLS	http://advances.sciencemag.org/content/3/12/e1701387
SUPPLEMENTARY MATERIALS	http://advances.sciencemag.org/content/suppl/2017/12/18/3.12.e1701387.DC1
REFERENCES	This article cites 42 articles, 9 of which you can access for free http://advances.sciencemag.org/content/3/12/e1701387#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service