Contemporary Philosophy and Methods for Studying Speciation and Delimiting Species

Authors: Burbrink, Frank T., and Ruane, Sara

Source: Ichthyology & Herpetology, 109(3) : 874-894

Published By: The American Society of Ichthyologists and Herpetologists

URL: https://doi.org/10.1643/h2020073
Species delimitation is a first step for realizing the extent of biodiversity and is relevant for all downstream applications in biology. The production of large genome-scale datasets for non-model organisms combined with the development of methodological tools have allowed researchers to examine fine-scale processes of speciation such as timing of origin, degree of migration, population-size changes, selection, drift, and recombination. Studies using reptiles and amphibians have, in part, paved the way for the development and use of such methods for exploring speciation and delimitation. While these methodological tools have improved our understanding of processes of diversification, researchers are far from agreeing upon a set of criteria to delimit species. In cases where genetic lineages are discovered that are unique to geographic areas, researchers usually agree that two entities exist. Disagreement about taxonomic status often centers on the degree of reproductive isolation between taxa and probability of remaining distinct. However, reproductive isolation is frequently inferred without examining gene flow, understanding the nature of hybrid zones, or determining the amount and type of introgression. Here, we review some of the vexing problems for delimiting reptiles and amphibians, which include isolation by distance, gene flow and differential allelic introgression, hybrid zone dynamics, and the nature of genomic islands of divergence. We also respond to recent literature criticizing model-based species delimitation in North American snakes in the context of these methodological advancements and address how scientists can move forward with studies on speciation.
The first step for identifying candidate species is to examine the structure of genetic or morphological discontinuities over a landscape. Most species initially formed due to some degree of extrinsic reproductive isolation, such as allopatry or parapatry (Kozak and Wiens, 2006), though the less common sympatric speciation is possible (not examined here; see Fitzpatrick et al., 2009; Edwards et al., 2020). Ironically, while researchers have not agreed on how to delimit species, even with whole-genome datasets, studying divergence has become more tractable (Galtier, 2019). Where two historical entities have been identified, processes of divergence can be studied along several predictable avenues of research.

Herein, we review methods used to dissect the gray zone of speciation, usually involving a complex process that includes the six major parameters in evolutionary biology: mutation, genetic drift, selection, migration, recombination, and population demography (Harvey et al., 2019; Peñalba and Wolf, 2020). Currently, no method has been developed to seamlessly provide a single framework to model speciation using all of these parameters simultaneously. However, we examine progress made to integrate many of these parameters, highlighting coalescent theory-based methods, which, for modern studies in herpetology, have dominated the field over the last two decades. Beyond the assumptions and limitations of the coalescent, we address the importance of properly detecting lineages, understanding selection, the formation of hybrid zones and maintenance of reproductive isolation, and the effects of recombination, genetic drift, and differential introgression of alleles on species boundaries. We then respond to some recent criticisms of model-based methods for inferring species boundaries in phylogenetic systems and provide suggestions for ways to ameliorate these problems.

**IDENTIFYING GEOGRAPHIC LINEAGES AND ISOLATION BY DISTANCE**

The first step for identifying candidate species is to examine the structure of genetic or morphological discontinuities over a landscape. Most species initially formed due to some degree of extrinsic reproductive isolation, such as allopatry or parapatry (Kozak and Wiens, 2006), though the less common sympatric speciation is possible (not examined here; see Fitzpatrick et al., 2009; Edwards et al., 2020). Functionally, detecting independently evolving lineages using genetic data should minimally show that putative species occupy geographically discrete areas. These putative species must have acquired more substitutions via selection or drift between them than those within their respective populations (Rannala, 2015). Since the advent of phylogeography, detecting geographic structure over a landscape has been accomplished by applying a measure of reproductive isolation that is not restrictive in the way that it requires strict reproductive isolation. This non-phylogenetic species concept in many cases will group unique, non-sister taxa because they fail to be reproductively isolated. Therefore, classification will not follow genealogy and the BSC will distort phylogenetic history (Rosen, 1978; Cracraft, 1983; Velasco, 2008). It may be just as useful to recognize the independent history of lineages as species prior to hybridization particularly since the existence of these lineages must be recognized before understanding the extent of hybridization (Nelson and Platnick, 1981).

Another concern voiced by some taxonomists is that when using genetic datasets, the potential to artificially subdivide groups that are not representative of unique species increases (Bauer et al., 2011; Galtier, 2019; Leaché et al., 2019). Methodologies should be able to differentiate species from artificially designated genetic clusters regardless of the upper boundary of loci analyzed by researchers. This then increases the number of species and 2) correctly determining if geographic lineages represent unique evolutionary trajectories. For the second part, applying a metric of strict reproductive isolation is unhelpful; two species can maintain unique evolutionary trajectories even when some level of gene flow is maintained heterogeneously throughout the genome (Mallet, 1995; Edwards et al., 2020). Ironically, while researchers have not agreed on how to delimit species, even with whole-genome datasets, studying divergence has become more tractable (Galtier, 2019). Where two historical entities have been identified, processes of divergence can be studied along several predictable avenues of research.

Herein, we review methods used to dissect the gray zone of speciation, usually involving a complex process that includes the six major parameters in evolutionary biology: mutation, genetic drift, selection, migration, recombination, and population demography (Harvey et al., 2019; Peñalba and Wolf, 2020). Currently, no method has been developed to seamlessly provide a single framework to model speciation using all of these parameters simultaneously. However, we examine progress made to integrate many of these parameters, highlighting coalescent theory-based methods, which, for modern studies in herpetology, have dominated the field over the last two decades. Beyond the assumptions and limitations of the coalescent, we address the importance of properly detecting lineages, understanding selection, the formation of hybrid zones and maintenance of reproductive isolation, and the effects of recombination, genetic drift, and differential introgression of alleles on species boundaries. We then respond to some recent criticisms of model-based methods for inferring species boundaries in phylogenetic systems and provide suggestions for ways to ameliorate these problems.

**IDENTIFYING GEOGRAPHIC LINEAGES AND ISOLATION BY DISTANCE**

The first step for identifying candidate species is to examine the structure of genetic or morphological discontinuities over a landscape. Most species initially formed due to some degree of extrinsic reproductive isolation, such as allopatry or parapatry (Kozak and Wiens, 2006), though the less common sympatric speciation is possible (not examined here; see Fitzpatrick et al., 2009; Edwards et al., 2020). Functionally, detecting independently evolving lineages using genetic data should minimally show that putative species occupy geographically discrete areas. These putative species must have acquired more substitutions via selection or drift between them than those within their respective populations (Rannala, 2015). Since the advent of phylogeography, detecting geographic structure over a landscape has been accomplished by applying a measure of reproductive isolation that is not restrictive in the way that it requires strict reproductive isolation. This non-phylogenetic species concept in many cases will group unique, non-sister taxa because they fail to be reproductively isolated. Therefore, classification will not follow genealogy and the BSC will distort phylogenetic history (Rosen, 1978; Cracraft, 1983; Velasco, 2008). It may be just as useful to recognize the independent history of
via tree-based or population clustering methods, plotting the locations of individuals from these groups on a map, and subsequently testing species boundaries and processes responsible for the origins of these groups using coalescent or integrated-data models (Avise, 2000a; Pritchard et al., 2000; Knowles and Carstens, 2007a; Knowles et al., 2007; Jombart et al., 2008; Fujita and Leaché, 2010; Leaché and Fujita, 2010).

Crucially, inferences on species boundaries rendered by clustering of allelic data alone or in the context of geography and coalescent estimators may be in error given isolation by distance (IBD) effects (Wright, 1943; Slatkin, 1993). Isolation by distance proposes that migration among populations decreases continuously with increasing spatial distance between populations. Where intermediate populations remain unsampled, distinct but spurious lineages appear to have formed near spatial endpoints. This is in contrast to when geographic groupings form due to adaptation to unique habitats (i.e., ecological speciation or isolation by ecology or adaptation), become isolated by physical barriers preventing gene flow (isolation by barrier), or a combination of both (Mayr, 1963; Avise, 2000a; Rundle and Nosil, 2005; Schluter, 2009; Pyron and Burbrink, 2010; Pyron et al., 2015). Multiple phylogeographic studies on reptiles and amphibians have demonstrated the importance of such environmental and physical barriers in structuring populations and promoting speciation, e.g., African frogs (Barratt et al., 2018), Anolis (Gray et al., 2019), Boa (Card et al., 2016), Hyla (Dufresnes et al., 2018, 2019), and Liopholis (Atkins et al., 2020). Poor sampling across the ranges of species will not account for population connectivity, and it is likely to exacerbate the probability of finding artificial geographic groups (Mason et al., 2020). It is imperative therefore to establish that the observed geographic clustering is not predicted by IBD to avoid subdividing potential lineages within a species unnecessarily (Yang, 2004; Jensen et al., 2005). The minimum divergence across the landscape must be in excess of what would be predicted by IBD alone to understand the effect of contemporary or historical climate and geographic barriers on genetic structure. However, knowing the minimum amount of sampling necessary to exclude the effects of IBD is complex given the interactions among several variables: the size of the range, dispersal distance, genetic variation, and variable migration rates among populations. This is likely to be an especially important consideration for herpetological studies. For many reptiles and amphibians, extensive range-wide sampling is difficult given that populations may occur at low frequencies and seasonally, and the total area occupied can be large and discontinuous and cover numerous political boundaries.

Several methods have been used to address the effects of IBD on population structures given genomic data and the location of individuals alone. For instance, effective estimates of migration surfaces (EEMS; Petkova et al., 2016) can visualize where genetic similarity by migration decays faster than expected by IBD, thus highlighting areas of reduced gene flow at geographic boundaries. This method has been used successfully in studies on reptiles and amphibians to identify areas of low migration (Homola et al., 2019; Myers et al., 2019b; Chan and Brown, 2020; Burbrink et al., 2021). Modeling population structure given continuous (geographic) and discrete (reproductive isolation) processes can be used to separate clines vs. geographic clusters using programs such as conStruct (Bradburd et al., 2018). Additionally, redundancy analyses or machine learning can be used to isolate non-spatial effects on genetic structure, fully supplementing reliance on incorrect Mantel tests (Diniz-Filho et al., 2013; Burbrink et al., 2021; Martin et al., 2021). Resistance or conductive surfaces representing spatial and environmental covariates on genetic dissimilarity between mapped individuals can also partial out the effects of space of population structure (Peterman and Pope, 2021), though these methods may only indicate that genetic changes are occurring over environmental gradients. Identifying candidates that may be distinct species occurs where geographically continuous groups are inferred from divergences in excess of those predicted by IBD, which may be due to adaptations to different environments or isolation at particular barriers. For many researchers, determining if these entities are species still comes down to the degree of gene flow, which can then be further examined via coalescent methods, diffusion techniques, or hybrid zone dynamics.

**COALESCENT SPECIES DELIMITATION**

Many contemporary methods for studying population history and speciation rely on some form of the coalescent model (Kingman, 1982; Griffiths and Tavaré, 1994; Wakeley, 2008). Within herpetology, coalescent-based methods have been frequently used to help identify candidates for additional investigation and/or delimit species across every continent inhabited by reptiles and amphibians (e.g., Hemidactylus, Leaché and Fujita, 2010; Ptychozoon, Brown et al., 2012; Euparkerella, Fusinatto et al., 2013; Leptopolis, Portillo et al., 2015; Plesiodon, Kuchta et al., 2016, 2018; Carlia, Afonso Silva et al., 2017; Eupusophus, Correa et al., 2017; Tropidurus, Domingos et al., 2017; Pseudechis, Maddock et al., 2017; Podarcis, Ponsis et al., 2017; Rana, Yang et al., 2017; Mesaspis, Solano-Zaleta and Nieto-Montes de Oca, 2018; Eurycea, Devitt et al., 2019; Philothamnus, Engelbrecht et al., 2019; Melanophryniscus, Pie et al., 2019; Pauropus, Slavenko et al., 2020; Leptotyphlops, Busschau et al., 2021). Coalescent theory models alleles within populations (or species) backwards through time until lineages merge (i.e., coalesce; for a more detailed description see Degnan and Rosenberg, 2009; Edwards, 2009). Using neutral loci, these methods allow us to predict changes in population size and address timing of divergences leading to testable theories about the origins of species (Knowles and Maddison, 2002; Drummond et al., 2005; Fujita et al., 2012; Burbrink et al., 2016; Myers et al., 2019a, 2019b). Additional parameters can be added to basic coalescent models to assess migration rates between lineages (Beeri and Felsenstein, 2001; Won and Hey, 2005), while other coalescent-based methods can be used to examine recombination and identify departures from neutrality to detect unique signatures of selection (McVean et al., 2002; Harris and Jensen, 2020). Such methods can be scaled up from the population level to examine relationships among species; this species-specific model is referred to as the multispecies coalescent (MSC; Rannala and Yang, 2003) and may include parameters to address migration, timing, and historical demography (Hey, 2010). Alternatively, these parameters can be estimated without with the MSC using diffusion approximation with the joint frequency spectrum (Gutenkunst et al., 2009). As with any robust statistic, increasing the size of the genetic datasets scales with...
increasing accuracy provided that the model is correct (Felsenstein, 2006; Liu et al., 2009; Ruane et al., 2015a). Coalescent methods provide an explanation of why individual gene trees often do not follow the same branching patterns of population or species trees, a concept known as incomplete lineage sorting (ILS; Maddison and Knowles, 2006; Degnan and Rosenberg, 2009). Looking backwards in time from contemporary lineages, ILS describes the common phenomenon where alleles fail to coalesce within the most immediate speciation event. Genetic polymorphisms scale with population size; therefore, larger ancestral populations and recent speciation events will exacerbate the problem of ILS. Using the MSC for genome-scale datasets incorporates ILS by modeling species/population splitting given both the timing of divergence and historical demography. Prior to considering ILS, many authors used the principle of reciprocal monophyly by gene to determine if candidate taxa represent real groups (“good” species). This prediction may be incorrect for young species with large population sizes, which are likely to suffer from ILS and therefore will not generate reciprocally monophyletic genes (Carstens and Knowles, 2007). Furthermore, when ancestral populations are large, the timing for reaching reciprocal monophyly for even half of all loci can be extremely long (Hudson and Coyne, 2002; Carstens and Knowles, 2007; Rannala and Yang, 2020). Thus applying the criterion of reciprocal monophyly may result in failure to recognize distinct species. Alternately, small and young founder populations could be incorrectly recognized as species given the criterion of reciprocal monophyly (Rannala and Yang, 2020).

Coalescent methods have recently been at the heart of controversial species delimitations (Sukumaran and Knowles, 2017; Leaché et al., 2019). The MSC has been criticized for improperly identifying species under the assumption that they delimit population structure and not species structure (Sukumaran and Knowles, 2017), resulting in taxonomic inflation affecting downstream fields such as macroecology, macroevolution, conservation, and medicine. However, the MSC will likely delimit population structure and not species structure and will not generate reciprocally monophyletic genes (Carstens and Knowles, 2007). Furthermore, when ancestral populations are large, the timing for reaching reciprocal monophyly for even half of all loci can be extremely long (Hudson and Coyne, 2002; Carstens and Knowles, 2007; Rannala and Yang, 2020). Thus applying the criterion of reciprocal monophyly may result in failure to recognize distinct species. Alternately, small and young founder populations could be incorrectly recognized as species given the criterion of reciprocal monophyly (Rannala and Yang, 2020).

Coalescent methods have recently been at the heart of controversial species delimitations (Sukumaran and Knowles, 2017; Leaché et al., 2019). The MSC has been criticized for improperly identifying species under the assumption that they delimit population structure and not species structure (Sukumaran and Knowles, 2017), resulting in taxonomic inflation affecting downstream fields such as macroecology, macroevolution, conservation, and medicine. However, the MSC will likely delimit population structure and not species structure depending on the group and nature of the data (Leaché et al., 2019). For example, if we chose multiple deeply sampled individuals from four North American ratsnakes of the genus Pantherophis, the MSC would correctly delimit with a probability of 1.0 at each node: *P. guttatus* (*P. obsoletus*, *P. bairdi*). This topology and the support for each of these four species would be consistent with our knowledge until ~2001 (see Conant and Collins, 1998). Exploring population structure within each species across their geographic ranges using coalescent-based species delimitation methods would find additional structure consistent with what has subsequently been described as species within *Pantherophis* *obsoletus*, *P. guttatus*, and *P. vulpinus* (Burbrik et al., 2000, 2021; Burbrik, 2001, 2002; Crother et al., 2011; Myers et al., 2020). Whether these additional delimitations represent species or populations may be difficult to answer using the MSC alone, though morphology and ecology have aided delimitation in some of these.

Researchers have attempted to model a property referred to as the conversion factor in a protracted speciation model, which should infer when populations convert to species using neutral genetic data (Sukumaran and Knowles, 2017; Leaché et al., 2019). It is unclear what biological process the conversion factor represents, though it presumably functions as a surrogate for the origin of adaptive traits or endogenous selection related to intrinsic reproductive isolation. However, Leaché et al. (2019) pointed out that inferring if an entity is a species or a population using the protracted speciation model of Sukumaran and Knowles (2017) is arbitrary given that the input data (genomic) are uninformative for understanding this conversion process. Additionally, the methodology requires reference to a related group where the taxonomy has been established for estimating the conversion factor in the target group. Therefore, philosophical and practical knowledge about speciation is not inherent to the study system but rather borrowed from “good” species.

The MSC can be used to identify candidate species, such as when using the program Bayesian Phylogenetics and Phylogeography (BPP; Yang and Rannala, 2010; Rannala and Yang, 2020) or when using newer software that simultaneously addresses gene flow (Jackson et al., 2016; Flouris et al., 2019; Smith and Carstens, 2020). A disturbing trend suggests that populations may be artificially delimited as species with increasing numbers of loci using BPP (Leaché et al., 2019). However, it is unclear if other MSC-based methods used to delimit populations/species have this problem (Bryant et al., 2012). Rather than relying solely on model predictions to delimit species, it is possible to use estimates from other parameters derived from coalescent models and determine when migration times exceed $10^5$ generations, where migration rates ($N_m$) $<1/generation$, and where long-term migration rates are significantly lower than the frequency at which F1 hybrids are produced (Leaché et al., 2019). For taxa in the gray zone, only a single or a few combinations of these parameters might cross the particular numerical threshold to indicate species status, still leaving researchers with a difficult decision to make. Unfortunately, the underlying issue with using MSC models, even those that incorporate gene flow, is that researchers fail to provide a measurable criterion that identifies when a population differs from a species using neutral genetic data only. This may then render MSC metrics to delimit species alone problematic without qualification. Ultimately, if we cannot decide on a well-formulated definition of species given the intricacies of process discoverable in the genome, then we cannot expect models or metrics to do that job alone.

Applying only a single measure to delimit species is unlikely to account for the complexity of parameter space needed to model speciation and therefore has greater potential for failure (Carstens et al., 2013). Most herpetologists, however, do not use MSC or other genetic-based methods exclusively but rather address putative species delimitations explicitly with other independent data such as geographic spatial and environmental niche data, ecology, morphology, estimates of migration rates, reproductive isolation, mating calls, and even chemical communication systems. These multifaceted approaches are common in herpetology, such as in *Aegkistrodon* (Burbrik and Guiler, 2015); *Cuora* (Spinks et al., 2012); *Discoglossus*, *Pelodytes* (Dufresnes et al., 2020b); *Eurycea* (Devitt et al., 2019); *Eutropis* (Barley et al., 2013); *Gonatodes* (Pinto et al., 2019); *Heteronotia* (Zoaya et al., 2019); *Lampropeltis* (McKelvy and Burbrik, 2017); *Liolema* (Camargo et al., 2012); *Mogophrys* (Liu et al., 2018); *Mimophis* (Ruane et al., 2018); *Plestiodon* (Pavón-Vázquez et al., 2018); *Phyllocautalus* (Koch et al., 2016); *Pithecopus* (Ramos et al., 2019); *Phrynosoma* (Leaché et al., 2009); *Pristimantis* (Ortega-Andrade et al., 2015); *Spea* (Neal et
al., 2018); *Tropidurus* (Domingos et al., 2017); *Typhanoptychus* (Chaplin et al., 2019); *Uperoleia* (Clulow et al., 2016). At least in herpetological studies, most evolutionary lineages discovered are partially or completely spatially isolated, suggesting that frequently used coalescent-based methods do not find random geographic structure across a species’ range. Demystifying the link between species delimitation and the processes of speciation using coalescent techniques in part provides some rationale for the decision that authors make when choosing to formally name or not name a taxon. At a minimum, coalescent methods can provide the date of origin and degree of gene flow sustained through time for taxa of interest as part of species delimitations and descriptions. However, there are many other avenues of research important for extending our understanding of speciation and delimitation.

We continue our discussion by examining methodologies to address processes of speciation that are often not examined by typical coalescent techniques alone. These often represent exciting new and rapidly developing methods using genomic data to understand how lineages diverge and examine genetic variation in the context of hybrid zones and clines, genic speciation, differential introgression of loci, and recombination landscapes.

**REPRODUCTIVE ISOLATION AND HYBRID ZONES**

The degree of reproductive isolation necessary for lineages to remain distinct is complex and may not rely on estimates of migration per generation. This is somewhat counterintuitive in that lineages must first be discovered to determine if and how much gene flow occurs between them but would not be discoverable if gene flow was extensive enough to erase the signal and presence of those lineages. Therefore, if distinct evolutionary lineages have been discovered, then gene flow has not been extensive enough to extinguish their presence. It is clear that species may form and remain distinct even in the presence of gene flow (Mayr, 1963; Van Valen, 1976; Coyne and Orr, 2004; Nosil, 2008, 2012; Feder et al., 2012; Harrison and Larson, 2014; Ravinet et al., 2017). This has been extensively documented across a diversity of reptiles and amphibians (e.g., *Bufo*, Wooten et al., 2019; *Ctenophorus*, *Litoria*, Melville et al., 2017; *Crocdyulus*, Hekkala et al., 2015; *Graptemys*, Godwin et al., 2014; *Liolaemus*, Olave et al., 2011; *Natrix*, Kindler et al., 2017; *Plethodon*, Chatfield et al., 2010; *Salamandra*, Canestrelli et al., 2015; *Terrapene*, Cureton et al., 2011; *Thamnophis*, Kapfer et al., 2013; *Triturus*, Arntzen, 2018; also see reviews on gene flow between species in amphibians and reptiles in Nadachowska, 2010 and Wollenberg Valero et al., 2019). Hybridization between well-defined species is even found across all major clades of lizards (Jančučová-Lásková et al., 2015). As might be expected given widespread hybridization, the variation in how genes introgress between taxa indicates that a hard criterion of complete reproductive isolation without qualification is too restrictive for setting species limits. Importantly, many species remain distinct despite secondary contact (Ackermann and Bishop, 2010; Paiseur and Rieseberg, 2016; Kumar et al., 2017; Dong et al., 2020). Many of these organisms may have reached an evolutionary endpoint referred to as partial reproductive isolation, where species remain distinct and yet are still in genetic contact long after forming (Servedio and Hermisson, 2020). For example, many hybrid zones likely formed during the Pleistocene, and for many squamates interaction between species would have continued for >30,000 generations (Hewitt, 2011).

Isolation with secondary contact can also produce species where Dobzhansky-Müller incompatibilities occur. Here genetic drift or selection in isolation forms numerous unique alleles between lineages, and, upon secondary contact, epistatic interaction among these different alleles in hybridizing individuals are selected against (Dobzhansky, 1937; Muller, 1942). This model of genomic divergence can happen rapidly, via the “snowball effect,” as negative interactions among loci in hybrids increases exponentially relative to the number of unique alleles evolving during isolation (Orr, 1995; Orr and Turelli, 2001). Processes of divergence producing unique areas in the genome adapted to particular environments or having negative epistatic interactions with other genes in hybridization point to the difficulty delimiting species when relying only on timing of lineage separation and whether gene flow occurs. The expectation is that eventually with stronger selection against hybrids due to the maintenance of genomic islands of divergence, increasing linkage disequilibrium, and drift, many markers should show divergence between species. A sober approach to species delimitation therefore embraces hybridization to better understand how species are interacting ecologically and how genomes respond to admixture.

In most stable hybrid zones, it is expected that selection against hybrids occurs thus reinforcing reproductive isolation (Howard, 1993; Garner et al., 2018). At the other extreme, new species may be formed as a result of hybridization (e.g., *Pelophylax*, Dubey et al., 2019; *Vipera*, Zinenko et al., 2016); these hybrid species may be unable to reproduce with their parental taxa (as seen in some teiid lizards; Reeder et al., 2002) and can significantly increase species diversity for a group (e.g., plethodontid salamanders; Patton et al., 2020). Additionally, adaptive introgression of beneficial alleles may help the receiver better adapt to local environments (Abbott et al., 2013; Bierne et al., 2013; Ma et al., 2019), reducing the possibility of extinction. It is clear that contemporary and ancient hybridization among many organisms is detectable among even commonly studied species (Edwards et al., 2016; Mallet et al., 2016; Figueiró et al., 2017; Kumar et al., 2017; Mason et al., 2019). For instance, using phylogenomic methods that account for reticulating hybridization, Burbink and Gehara (2018) inferred that the ancestor of the Mexican kingsnake species group (*Lampropeltis mexicana*, *L. webbi*, and *L. ruthveni*) formed via ancient hybridization between ancestors of a clade of milksnakes (*L. abnorma*, *L. micropsilus*, and *L. polyzona*) and a clade of other tri-colored *Lampropeltis* (including *L. pyromelana*, *L. zonata*, and *L. alterna*) in the late Miocene. The area of this ancient hybridization in the southwestern United States and northern Mexico is known for repeated secondary contact and hybridization among squamates in the Chihuahuan and Sonoran Deserts, including other snakes (Myers et al., 2019b) and lizards (*Crotaphytus*; McGuire et al., 2007). These studies highlight that diversification and origins of contemporary squamate communities likely include hybridization over long periods of time.

Studying hybrid zones is particularly useful for understanding degrees of reproductive isolation, but these investigations are not typically included in most delimitation studies likely due to difficulty obtaining adequate
samples and lack of *a priori* knowledge on the location of areas of contact. Hybrid zones between species may form when there is selection against hybrids (tension zones), when there is selection for hybrids (bounded superiority), or when there are ecological gradients resulting in reduced fitness for parental species in the environment of the other species (Barton and Hewitt, 1985; Barton and Gale, 1993). Selection can also be endogenous (genomic) or exogenous (environmental). Paradoxically, while hybrid zones indicate reproduction between species, their presence simultaneously demonstrates that some level of isolation had occurred. Introggression of loci into the parental taxa outside of the hybrid zone will vary given recombination and selection on the function of those genes (Feder et al., 2012; Wolf and Ellegren, 2017). In contrast, a neutral hybrid zone is defined by having no selection against hybrids, and free introgression of alleles among parental types. In a neutral hybrid zone, the two species should collapse into a single species due to hybrid swamping (Barton and Hewitt, 1981; Wolf et al., 2001). Making matters more complex, it is likely some of these different types of hybrid zones or areas with no hybridization are occurring in different parts of the ranges where species contact, or are changing through time.

Hybrid zones can be modeled to understand the timing of their formation, the spread of hybrids, the degree of selection on hybrids, and gene flow and reproductive isolation (Endler, 1977; Barton, 1979; Avise, 2000b; Nachman and Payeour, 2012; Gompert et al., 2017). For species delimitation, hybridization between lineages in the context of the gray zone should be qualified with respect to the width of hybrid zones relative to parental species ranges, degree of gene flow between the taxa, and which alleles are introgressing. The width of the hybrid zone (tension zone) can be predicted by dispersal rate and selection against hybrids or genetic distance (Barton and Gale, 1993); in reptiles, random dispersal and genetic distances are important predictors of the size of hybrid zones (McEntee et al., 2020). Still, the genomics of speciation and secondary contact are complex, and heterogeneous genomic divergence is not easily generalizable over the environmental landscape. Researchers are typically moving away from a strict interpretation of the biological species concept where reproductive isolation occurs at individual or coadapted genes, the so-called genic view of speciation (Wu, 2001).

Variation in how alleles introgress is extensive given selection, pleiotropy, and the coupling of adaptive genes (Nosil et al., 2009; Nosil and Schluter, 2011; Payeour and Rieseberg, 2016; Gompert et al., 2017; Schmicki et al., 2017; Campbell et al., 2018; Jiggins, 2019). These processes of divergence and gene flow may not occur at a constant rate during speciation. Loci involved in adaptation and reproductive isolation experience little gene flow between species, which may be in contrast to the remainder of the genome where gene flow may remain high (Coyne and Orr, 2004). In these cases, the selected loci plus linked genomic regions that are not broken up by recombination will form islands of divergence between young species (Nadeau et al., 2012; Wolf and Ellegren, 2017). The presence of these genomic islands suggests that populations may be in the “act of speciation” and therefore provide opportunities to understand how lineages diverge and maintain independent identities. In populations that are diverging with ongoing gene flow, genomic islands will be generated at loci related to divergent ecologies or genomic incompatibilities (Hejase et al., 2020).

Understanding how genomic islands originate, their function, and even if they exist in particular study organisms is still an area of continuing exploration even among well-studied species. Genomic islands of divergence may have arisen rapidly as a result of selective sweeps or background selection unrelated to reproductive isolation (Noor and Bennett, 2009; Cruzickshank and Hahn, 2014; Duranton et al., 2018). Endogenous selection against hybrids at particular genomic island loci may also have formed in allopatry due to Dobzhansky-Muller incompatibilities and not be adapted to particular environments (Bierne et al., 2011). Identification of loci involved in reproductive isolation may be further complicated by variation in recombination rates across genomes. This variation can lead to spurious identification of genomic islands in low recombination regions even when under neutrality. Typical genomic scans do not account for variation in recombination rate, which, even under neutrality, can generate heterogeneous estimates of selection using typical measures such as *Fst* (Booker et al., 2020). Therefore, researchers are urged to consider the recombination landscape when looking for islands of genomic divergence containing loci that influence reproductive isolation.

Where genomic islands have been credibly identified, these loci then can be associated with pre- and postzygotic isolation and/or ecological and morphological differences via ecological speciation (Nosil et al., 2009; Marques et al., 2016; Mérot et al., 2017; Wolf and Ellegren, 2017). Additionally, these alleles or co-adapted genes may change rapidly along clines and lead to complete reproductive isolation (Mallet, 2008; Stankowski et al., 2017, 2019; Souissi et al., 2018). For many organisms, the presence of genomic islands of divergence where alleles are not traversing clines might suggest that reproductive isolation is not likely to be reversed. This could serve as a minimum criterion for defining a species and would be consistent with the genetic cluster view of speciation, where reproductive isolation was a cause but not a criterion for delimiting species. Under this particular view, species should be considered distinct when reduced recombinants and heterozygotes in loci are found where parental forms overlap (Mallet, 1995, 2020). However, we know of no studies in herpetology where such diverging loci were found first and then species delimited second. But, we underscore that species divergence can occur due to a small number of select genes, as in the case of hooded and carrion crows, where just a few loci control prezygotic isolation and also permit the phenotypes of each species to remain distinct despite gene flow among all other loci (Knief et al., 2019).

Genomic islands of divergence also hold promise for identifying key loci and associated traits that function to reproductively isolate species. Traditional species delimitation methods typically measure or score external anatomy or reproductive call data to often indirectly infer the effects of isolation via drift or selection. These methods have, of course, been useful for centuries but may fail to account for all of the morphological or physiological differences between
lineages. It is possible now to determine which traits are involved in reproductive isolation by examining significant differences between taxa across their genomes and then relate those differing alleles to morphology (Han et al., 2017). For example, an intriguing recent study (Rautsaw et al., 2020) looked at genomic divergence between the Florida water-snakes Nerodia clarkii, occurring in high-salinity environments, and N. fasciata, in freshwater environments. They demonstrated that despite recurring historical and contemporary gene flow between these taxa, several key loci were responsible for ecological divergence in these environments. Scanning the genome for selection, the authors identified 31 loci of interest using the annotated genome from the related Thamnophis elegans. With gene ontology (GO) enrichment, a bioinformatics initiative to classify genes and gene products across taxa, the authors identified a network of related terms underscoring ten genes involved with osmoregulation. These functional dissimilarities correlate to the primary differences in habitat salinity between N. clarkii and N. fasciata. It would be difficult to determine that these particular differences are occurring between these taxa using external morphological data alone. One of the most pressing needs now is having more fully annotated genomes representative of most families and subfamilies of reptiles and amphibians. This would lead to a better understanding of the primary axis of trait diversification in the gray zone.

Once taxa are established, dissecting how and when contemporary and historical landscapes isolated populations and facilitated speciation through time is achievable. Biologically relevant information, including ecology, physical barriers, gene flow, historical climate, landscape resistance, and geographic space can be used to predict the causes of species formation when using techniques such as redundancy analyses, conductance surface techniques, generalized dissimilarity modeling, and machine learning approaches (Ferrier et al., 2007; Zhang, 2010; Diniz-Filho et al., 2013; Burbrink and Gehara, 2018; Burbrink et al., 2021; Peterman and Pope, 2021). Such methods allow us to understand how predictor variables (e.g., environment, physical barriers) influence resulting population and species structure. For example, Myers et al. (2017a, 2019b) demonstrated divergence in a community of 13 species of snakes at the division between the Chihuahuan and Sonoran Deserts (the Cochise Filter Barrier) using subgenomic data combined with geographic distances and environmental heterogeneity. This research found that no single unifying process generated lineages among species of snakes within the genera Crotalus, Hypsiglena, Lampropeltis, Mastiocophus, Pituophis, Rhinocellus, Salvadora, Sonora, Thamnophis, and Trimorphodon. Myers et al. (2017a, 2019b) concluded that both isolation by distance and environment throughout the Pleistocene and Pliocene produced these deep lineages. In another recent example, Burbrink et al. (2021) examined the origins of the Eastern Ratsnakes (Pantherophis obsoletus complex) by integrating genomic data with models of diversification to include current and historical environmental data and analyses of hybrid zones. Paralleling the conclusions from Burbrink et al. (2000) but using genome-scale data with both coalescent and neural network methods, Burbrink et al. (2021) again found a strong role for the Mississippi River as a genetic barrier for these ratsnakes, along with evidence that sharp ecological transitions helped structure the four taxa throughout the Pleistocene.

CONTROVERSIAL SPECIES DELIMITATIONS IN HERPETOLOGY

Recently, Hillis (2019, 2020) and Chambers and Hillis (2020) have written three papers discussing species delimitation that fall short of considering modern theory and application. These papers primarily target delimitations in North American snakes, though as demonstrated here, many other researchers use the same methods in herpetology for taxa found in Central and South America, Asia, Europe, Australia, and Africa. These authors contend that species delimitations from multiple herpetological studies were unfounded because they represent arbitrary slices of clines or conjecture that hybrid zones are too wide for species to be valid. However, they ignore that hybridization across the tree of life is common, these species diverged during the Pleistocene or even earlier and have maintained independent evolutionary trajectories, and incomplete gene (lineage) sorting is one expectation of the speciation process (Wiley, 1978; Knowles and Carstens, 2007b; de Queiroz, 2007; Payne and Riceberg, 2016; Gompert et al., 2017; Barth et al., 2020; Moran et al., 2021).

Describing his proposal of sensible nomenclatural practice, Hillis (2019) provides a decision tree to determine when to change taxonomy (Hillis, 2019: fig. 1). Oddly, this proposal uses criteria that may be applied to species delimitation but simultaneously attempts to specify when to revise names above the species level (see text for example regarding the genus Rana). We point out that this decision tree makes a number of hidden and false assumptions and cannot accurately reflect the way species should be delimited. The first stage on hypothesis formation suggests grouping individuals into potential lineages; unless known ahead of time, this stage tests the existence of lineages and subsequently the assignment of individuals into lineages. The second stage that guides the remaining taxonomic treatments rests entirely on the unqualified criterion that species should be sufficiently reproductively isolated to be considered independently evolving. The crux of this proposal relies only on the application of the biological species concept. The word “sufficiently” also suggests that determining reproductive isolation is subjective, and, as we have reviewed here, reproductive isolation without some qualification is not useful. Furthermore, whether reproductive isolation occurs has never been tested on most described and named reptile and amphibian species.

Passing the second stage, the reader is directed to the right if the lineages fail to show sufficient reproductive isolation and rather represent geographic variation. It is unclear how geographic variation (no reproductive isolation) and distinct species (reproductive isolation) are mutually exclusive given that distinct species also can be geographically variable. Proceeding down the right-hand path, Hillis then asks if the geographically variable lineages have an available name(s). This is more complex than it appears as it must mean that these lineages represent subspecies, but there is no mention of the word or concept in the paper. Without considering that these are subspecies, then of course the organism that is composed of multiple lineages that are not reproductively isolated has an already applied name (the organism that you started the process with). Unfortunately, even if they are to be considered subspecies, it would not be clear if they represent entities with distinct evolutionary histories that
happen to also be incompletely reproductively isolated or just geographic variants (“handles of convenience”). The latter is not a unit of evolution and therefore should not receive a name. If the geographically variable lineage does not have a name, then this leads to a box that says to fix the taxonomic problem with the least disruption to the current taxonomy. This is unhelpful because the user has to fix the taxonomy according to the proper nomenclatural rules regardless of the level of disruption, or change their results (or the interpretation of those results). Turning to the left side of the decision tree where we have reproductive isolation, we are correctly instructed to determine valid names and follow the pertinent nomenclatural rules. If there are no available names, we are again instructed to not be disruptive to the current taxonomy, though how we should do that outside of the specific nomenclatural rules or altering the perception of our results is not explained. If there are available names, then we are questioned if the groups are monophyletic or if the species are valid. It is unclear at this taxonomic level how monophyly is being approached, perhaps reciprocal monophyly, and a positive conclusion has already established that we have valid species. If yes, then we are instructed to “then leave the nomenclature the #&%& alone!” but there was no step for applying the valid names. It is uncertain from this decision tree just who has the final authority to decide when “taxonomy is fixed” and if multiple iterations of the decision tree are permitted given new data or analyses. If not, this approach hampers species discovery and provides no help for understanding just what data, methods, or philosophy aid with contemporary species delimitations. Moreover, it is difficult to comprehend what philosophy and methods are approved given that Hillis recently co-authored a paper on the phylogeography of *Eurycea* using the MSC to delimit species, but without directly assessing degree of reproductive isolation or gene flow (Devitt et al., 2019).

Recently, Chambers and Hillis (2020) posited that species delimited in the American milksnakes (*Lampropeltis triangulum* complex; Ruane et al., 2014) “represent arbitrary slices of continuous geographic clines.” Although Chambers and Hillis (2020) initially acknowledged that loci may not be monophyletic for recently diverged species, they subsequently used the expectation that nuclear loci should be monophyletic as an argument to refute the Ruane et al. (2014) delimitations that re- Elevated previously described milksnake species. Their point is made despite widespread research showing that reciprocal monophyly among gene trees will not be the case for most taxa due to incomplete gene sorting, particularly given recent speciation times and large population sizes (Maddison and Knowles, 2006; Knowles and Carstens, 2007b). Rather than test whether the probability that the 11 nuclear gene trees for the milksnakes were consistent with the expectations of monophyly and species limits given coalescent predictions that account for sorting and migration, Chambers and Hillis (2020) concluded several proposed species delimitations were unfounded. However, neither monophyletic gene trees nor the ambiguous criterion of “consistent nuclear divergence” are true for their own proposed milksnake taxon of “*L. triangulum,*” which is also paraphyletic in the supplemental gene trees (fig. S2 of Chambers and Hillis, 2020). Chambers and Hillis (2020) also attempted to use BPP to demonstrate that the support for two milksnake taxa (*L. triangulum* and *L. gentilis*) from Ruane et al. (2014) was not indicative of species divergence. To do so, they randomly sliced the ranges of the North American milksnakes, combined samples into groups within the slices (see fig. 2b in Chambers and Hillis, 2020) and found these new groupings could be delimited with high support, thus concluding that the delimitations in Ruane et al. (2014) were faulty. Based on the sampling scheme used for the random slices, IBD could result in similarly well-supported delimitations (Chambers and Hillis, 2020: fig. 2b), but Chambers and Hillis (2020) used no other data or methods to examine IBD, ILS, or migration to better understand actual processes (e.g., see Burbarkin et al., 2021) nor were the marginal likelihoods of the arbitrary-slice models compared to the original model of Ruane et al. (2014) to further validate their own results. Despite stating that “new species designations, especially of well-studied groups, are best made after careful consideration of all sources of relevant evidence,” Chambers and Hillis (2020) ignored other available works and datasets, which include morphological as well as additional genetic data (e.g., the original descriptions of the species; Ruane, 2015; Ruane et al., 2015b). When conducting species delimitation or revising taxonomy based on previously collected data, we would advocate researchers gather and assess a preponderance of evidence. At least for the milksnakes, this would include obtaining new genetic (preferably genomic) data, integrating spatial and environmental information, and using appropriate methods to delimit species reflective of processes of divergence.

In the case of the species delimitations for North American copperheads (*Agkistrodon contortrix* complex) from Burbarkin and Guiler (2015), Hillis (2020) took issue with the presence of a hybrid zone between proposed species. His criticisms were based on the assumption that the hybrid zone was neutral, which would indicate hybrid swamping of separate lineages and result in the eventual collapse of detectable species. While stating that there were historical lineages with hybrid zones, he also believes contrary to this that species delimitation is based on arbitrary slices of a continuous range. The lineages exist and are discoverable regardless of whether a hybrid zone exists. Hillis fails to provide any evidence that there is no selection against hybrids or that introgression was occurring. Burbarkin and Guiler (2015) demonstrated that *A. contortrix* and *A. laticinctus* diverged in the late Pliocene or early Pleistocene, were consistent with previous morphological treatment of the two groups (Gloyd and Conant, 1990), and thus considered them two species despite documenting areas of hybridization (Burbarkin and Guiler, 2015). These snakes hybridize at the intersection of two primary ecoregions in the United States, the eastern temperate forests and the Great Plains (Omernik, 1987; Bailey, 1995). This division between the Great Plains and forested regions of the Eastern Nearctic has been recognized as an ecological barrier for other reptile and amphibian taxa (Costa et al., 2008) and serves as a full or partial boundary for ~80 species occurring in the Eastern Nearctic and ~40 species primarily occurring throughout the Great Plains (see distributions from Powell et al., 2016). If these hybrid zones are neutral and ancient, then the expectation here is that these distinct and old lineages should be swamped (and thus eliminated) via hybridization given rates of life-time dispersal. We can conservatively predict the spread of the copperhead hybrid zone under neutrality by using the equation: \( w = 2.51\sigma\sqrt{T} \) (Barton and Gale, 1993; Bailey et al., 2015).
Here, given the constant (2.51), root mean squared (RMS) of lifetime dispersal from natal site (σ) and generations since the formation of the hybrid zone (T), we can estimate width (w) of the hybrid zone under neutrality. This equation can be rearranged to solve for T to predict how many generations it would take given lifetime dispersal distances for copperheads to form a growing neutral hybrid zone that would cover the entire width of both parental species ranges, here ~2,300 km (Gloyd and Conant, 1990; Burbbrink and Guither, 2015). For copperheads, σ = 4 km (though dispersal distances in the southern parts of their range where they are not constrained to hibernacula may be greater over a lifetime, Fitch, 1966; Smith et al., 2009; G. Schuett and C. Smith, pers. comm.) and a generation time of three years (Gloyd and Conant, 1990), the current hybrid zone in copperheads at ~400 km can be achieved in 1,588 generations (4,764 years). These hybrid zones are likely older than this given stability and placement of these ecoregions during the Holocene. Moreover, if these taxa have been in contact since formation (0.96–2.5 My, Burbbrink and Guither, 2015), and new analyses suggest they have been (Gehara et al., 2020, bioRxiv: 2020.12.04.410670), then hybrid swarming would overtake the range of both taxa in ~2,480 generations (15,7440 years), thus eliminating our ability to detect these as distinct lineages. At longer dispersal distances (8 km over a lifetime), extinction of the independent species via hybridization can occur in as little as 6,900 years. It is likely that these taxa and others with ancient hybrid zones have remained in a state of partial reproductive isolation as a stable evolutionary endpoint (Servedio and Hermisson, 2020); these species occur in different geographic areas with unique ecologies and have remained distinct in the face of hybridization since their time of origin. In this case, the contemporary hybrid zone would not suggest that two lineages are in the act of fusing. Of course, we recommend that future studies test all of these hypotheses and examine the width and nature of this hybrid zone with genomic data and adequate samples throughout the connecting ranges.

For the copperheads, Hillis (2020) suggested that these lineages are distinct and detectable, yet because they form a hybrid zone, they should be treated as subspecies. We argue that there is no need to use subspecies here or elsewhere given that it is unclear what this rank represents biologically. While arbitrary thresholds have been applied in some organisms, specific criteria that separate subspecies from species based on degree of gene flow or dynamics of hybrid zones have never been established for this taxonomic rank. Furthermore, as we have discussed above, it is difficult enough to determine when a population becomes a species, sliding another rank in between these terms would result in less clarity. It is also uncertain how downstream applications should use this rank relative to species. For example, should subspecies, which could represent diverging lineages (protracted speciation) be enumerated along with species to investigate processes of diversification, ecological evolution, or community assembly? Finally, the term is subjective; it has not historically been used consistently or meaningfully with respect to taxonomy (Wilson and Brown, 1953; Burbbrink et al., 2000). In some cases, they are considered evolutionary lineages (or incipient species) still connected reproductively with other such lineages. In that case, they are only different from species by degree of reproductive isolation but not by the primary formation of discoverable historical lineages (see de Queiroz, 2020). In other cases, they are considered “handles of convenience” or arbitrary range slices from continuously variable morphological data and should not warrant a name.

Moreover, there are many taxa with the applied species rank that also have hybrid zones, e.g., Hyperolius (Bell and Irian, 2019), Pseudacris (Lemmon et al., 2007), Podarcis (Pinho et al., 2009), Sternotherus (Scott et al., 2019), Tympanocryptis (Melville et al., 2019), though the nature of reproductive isolation is likely specific to each species group. Rather, it is clearer to describe two species, explain when they originated, define where the location of a hybrid zone is, quantify how much gene flow occurs, and investigate the nature of the introgressing loci. We do not need special classes for those taxa that have no or arbitrarily small hybrid zones (“species”) and those that have larger hybrid zones (“subspecies”). The size of a hybrid zone will vary between taxa and is likely related to lifetime dispersal, ecology, degree of selection against hybrids, and the genomic architecture of the species studied.

Many herpetological systematists are now using genome-scale data and computational modeling to address questions about species divergence (Morando et al., 2020; Reyes-Velasco et al., 2020; Streicher et al., 2020) but also understand that data sources and methods progress over time. So, it is pointless to criticize authors for using what data and methods were available to them at the time, whether those be long-deceased authors such as Blanchard, Cochran, Cope, Dickerson, and Dunn, or more recent authors. For example, 30 years ago Hillis described a new species of Synephis (Hillis, 1990) from only two specimens, one heavily damaged. While one could criticize this initial taxonomic decision, given that Hillis did not account for morphological variation within the genus, incorrectly scored material from the type specimens, and could not reliably define the distribution of the new species, it would be pointless to do so until additional new data were analyzed to refute or support the hypothesis (see Torres-Carvajal et al., 2015, for a more modern treatment of this genus).

Ignoring that there is both a historical and financial framework for science conducted at any point in time, Hillis (2020) suggested that numerous studies that have delimited species using mtDNA are likely incorrect, e.g., the Lampropeltis getula complex (Pyron and Burbbrink, 2009a). Of course, we recognize that some delimitations may be incorrect and that using a single locus to infer history can be problematic, especially for studying detailed processes of divergence (Toews and Brelsford, 2012; Spinks et al., 2014). However, given their small effective population sizes, low recombination, and high rates of substitution, mtDNA genes have been extremely useful for identifying likely candidates that represent cryptic species (Moritz et al., 1987; Funk and Omland, 2003; Piganneau et al., 2004; Rubinoff and Holland, 2005; Zink and Barrowclough, 2008; Edwards and Bensch, 2009; Allio et al., 2017). However, many of these studies also examine other sources of data as well to delimit species. For example, the studies on kingsnakes did not solely use mtDNA to determine species, but included ecological niche information and geography, and considered previous works on morphology (Pyron and Burbbrink, 2009a, 2009b). Hillis (2020) trivializes these studies by ignoring both the depth of mtDNA divergence among these taxa (the Late Miocene) and any corroborating evidence from ecology, morphology.
Ambystoma macrodactylum confirmed the identity of mtDNA lineages of Bynoe's Gecko biogeographic structure within a group exists and warrants used this marker successfully to initially determine that gene flow alone. However, there are many studies that have course result in error and is impossible to adequately assess distant related taxa. Ignoring such ancient divergences at known and shared biogeographic barriers discounts the historical magnitude of landscape effects on structuring deep lineages among distantly related taxa.

Using mtDNA as the sole marker to delimit species can of course result in error and is impossible to adequately assess gene flow alone. However, there are many studies that have used this marker successfully to initially determine that biogeographic structure within a group exists and warrants further investigation. For example, Zozaya et al. (2019) confirmed the identity of mtDNA lineages of Bynoe's Gecko (Heteronotia binoei complex) using pheromone data, which corresponded with earlier phylogeographic studies on these lizards (Fujita et al., 2010; Moritz et al., 2016). In this more recent study, the authors also examined the morphology and pheromonal signatures for each of the mtDNA lineages. The geckos from the ten mtDNA lineages were phenotypically similar yet variable within a lineage; however, the pheromone composition of each of the mtDNA lineages was substantially different. This suggests that chemical communication in these geckos is an important prezygotic isolating mechanism. The work of Zozaya et al. (2019) indicates that hard-to-observe traits, such as pheromones, may be the key features that differentiate seemingly cryptic taxa and that mtDNA remains a useful tool for screening candidate taxa.

There is also increasing evidence that inferred species boundaries from deep mtDNA divergences may be incorrect in many cases, given mitochondrial-nuclear incompatibilities. The 13 mtDNA protein-gene subunits interact with ~70 nuclear subunits to generate functional proteins in the electron transport chain (Lloyd and McGeehand, 2013). Rapid evolution of mtDNA within lineages therefore requires compensatory changes within the nuclear counterparts for these genes which produces tightly coupled gene complexes. Hybridization between deep lineages, therefore, disassembles these complexes and can reduce fitness (Dowling et al., 2008; Burton and Barreto, 2012; Ballard and Pichaud, 2014). These mtDNA divergences between species may produce cytonuclear (nuclear-mitochondrial) incompatibilities in hybrids with respect to the parental species, and between species in general, thus reinforcing reproductive isolation (Ma et al., 2016; Jhuang et al., 2017; Lamelza and Allion, 2017; Telschow et al., 2019), which has been recently explored in amphibians and reptiles (Gibeaux et al., 2018; Haenel and Del Gaizo Moore, 2018; Prokić et al., 2018; Firneno et al., 2020). For example, cytonuclear discordance in hybrids between lineages of Ambystoma macrodactylum result in poorer feeding performance (Lee-Yaw et al., 2014). On the other hand, introgression across species boundaries may substitute resident mitochondria suffering from a heavy mutational load (Sloan et al., 2017). There are multiple examples in reptiles and amphibians where mitochondria have introgressed between clearly distinct species such as in Ambystoma (Denton et al., 2014), Crotaphytus (McGuire et al., 2007), Lampropeltis (Bryson et al., 2007; Ruane et al., 2014; Burbrink and Gehara, 2018), Rana (Plotner et al., 2008), and sea turtles (Vilaça et al., 2012). Here mtDNA alone would fail to detect unique species or would provide distorted limits of ranges.

For some of these previous studies that relied on mtDNA, along with morphological or ecological data, many of the same delimitations or geographic lineages are supported by contemporary approaches. For example, using genomic data and accounting for IBD, Myers et al. (2019b) recently found the same two distinct lineages within the Lampropeltis getula complex previously delimited using mtDNA (Pyron and Burbrink, 2009a, 2009b). Similarly, using ~4,300 loci, Myers et al. (2020) found the same three cornsnakes (Pantherophis emoryi, P. guttatus, and P. slowinskii) previously delimited using only mtDNA (Burbrink, 2002) and with an additional deep divergence within P. guttatus. Barrow et al. (2018) found when using AHE markers that spatial clusters and some species trees were similar to mtDNA geographic structure within the species of frogs Hyla cinerea, H. squirella, and Lithobates sphenocephalus in the Southeastern US. In general, mtDNA delimitations are more likely to be correct when genetic distances are large and divergence times are old. Large mtDNA differences between groups cannot be simply disregarded given the importance for those mitochondrial proteins to interact with complementary nuclear subunits to efficiently generate ATP (adenosine triphosphate). We do not advocate that researchers rely entirely on single-locus data given that the possibility of producing many thousands of independent loci is available. However, rather than suggesting that most previous taxonomic research using mtDNA or even limited numbers of nuclear loci are incorrect, we prefer to remain neutral until new studies suggest otherwise. Ideally, researchers will include genomic data in revisionary work and integrate other data sources to help guide taxonomic revisions.

CONCLUSIONS

Species delimitations using any method under any concept are hypotheses (de Queiroz, 2005). While other researchers working in downstream fields that use species as a unit of study prefer stable taxonomies and properly identified species, we cannot lose sight that all delimitations exist in the realm of testable science, and therefore complete stability is unrealistic (Crother, 2009). Development of methods to better address IBD and reduce incorrect species delimitations will be useful if they can be seamlessly integrated with methods that identify geographic groupings and examine the nature of genetic introgression. Species delimitation does not rely on a cookbook approach where one follows a set of directions to make a conclusion about species status that will remain immutable. A rigid formula for delimiting taxa may immediately be on the lagging edge of science in the face of new data types and analyses. Instead, we agree with other authors (Rannala and Yang, 2020) that species delimitation should comprise an evidence-based approach that properly
examines genetic structure integrated with additional data types to better understand timing of origin, gene flow, and degrees of introgression. Of course, not all data types are required to delimit every species, and the current cost of generating genomic data is unfortunately too expensive for many researchers. Ideally though, scientists should use all available evidence to determine whether the species in question are best represented by current taxonomy. New methodologies and technologies are showing us that old ideas of speciation without allowing for gene flow are often in error. Speciation with gene flow provides exciting avenues of research for understanding a more complex view of the evolution of species not constrained by unrealistic notions of what species should be. By embracing these contemporary data and methods, researchers can fully recognize the complexity of species boundaries, realize a more accurate tree of life, and properly aid other fields and educational sources about biologically meaningful taxonomic units.

ACKNOWLEDGMENTS

We thank Justin Bernstein, Brian Crother, Kevin de Queiroz, Scott Edwards, Luke Harmon, Sean Harrington, Adam Leaché, Marcelo Gehara, Edward Myers, Alex Pyron, Brian T. Smith, and Jeff Streicher for either reviewing a draft of this manuscript or contributing to discussions about species limits. We thank Gordon Schuett and Chuck Smith for contributing to useful discussions on copperhead dispersal. FTB acknowledges support from the National Science Foundation (NSF-DEB; Dimensions USBIOTA1831241).

LITERATURE CITED


Burbrink, F. T., Y. L. Chan, E. A. Myers, S. Ruane, B. T. Smith, and M. J. Hickerson. 2016. Asynchronous demo-


Mimophis (Serpentes: Lamprophiidae). Systematics and Biodiversity 16:229–244.


