Syst. Biol. 69(1):184–193, 2020

DOI:10.1093/sysbio/syz042
Advance Access publication June 10, 2019

The Multispecies Coalescent Over-Splits Species in the Case of Geographically Widespread Taxa

E. Anne Chambers* and David M. Hillis

Department of Integrative Biology and Biodiversity Center, The University of Texas at Austin, Austin, TX 78712, USA
**Correspondence to be sent to: Department of Integrative Biology and Biodiversity Center, The University of Texas at Austin, 2415 Speedway #C0930, Austin, TX 78712, USA; E-mail: eachambers@utexas.edu.*

Received 19 November 2018; reviews returned 22 May 2019; accepted 24 May 2019
Associate Editor: Richard Glor

Abstract.—Many recent species delimitation studies rely exclusively on limited analyses of genetic data analyzed under the multispecies coalescent (MSC) model, and results from these studies often are regarded as conclusive support for taxonomic changes. However, most MSC-based species delimitation methods have well-known and often unmet assumptions. Uncritical application of these genetic-based approaches (without due consideration of sampling design, the effects of *a priori* group designations, isolation by distance, cytoplasmic–nuclear mismatch, and population structure) can lead to over-splitting of species. Here, we argue that in many common biological scenarios, researchers must be particularly cautious regarding these limitations, especially in cases of well-studied, geographically variable, and parapatrically distributed species complexes. We consider these points with respect to a historically controversial species group, the American milksnakes (*Lampropeltis triangulum* complex), using genetic data from a recent analysis (Ruane et al. 2014). We show that over-reliance on the program Bayesian Phylogenetics and Phylogeography, without adequate consideration of its assumptions and of sampling limitations, resulted in over-splitting of species in this study. Several of the hypothesized species of milksnakes instead appear to represent arbitrary slices of continuous geographic clines. We conclude that the best available evidence supports three, rather than seven, species within this complex. More generally, we recommend that coalescent-based species delimitation studies incorporate thorough analyses of geographic variation and carefully examine putative contact zones among delimited species before making taxonomic changes. [Classification; speciation; species concepts; species delimitation; taxonomy.]

Systematists attempt to understand and organize the diversity of life using two fundamental concepts: species and trees of relationships among species. Under this framework, species are viewed as individual, independently evolving metapopulation lineages, within which organisms typically mate and exchange genes (Wiley 1978; Mayden 1997; de Queiroz 1998, 2007). Species lineages split and give rise to new independent lineages, forming phylogenetic trees of species in the process. Within those trees, monophyletic groups of species, or clades, represent historical groups that share a common evolutionary origin.

The boundary between species and clades is not arbitrary, as life is clearly not organized in a continuum. Instead, there are clear reproductive and genetic breaks that allow different lineages to evolve on independent evolutionary pathways. Within sexual species, gene flow typically maintains cohesion such that lineages evolve as units through time (Ghiselin 1974; Templeton 1989). Ecological circumstances (selection for particular ecological roles) may also play a role in maintaining species, even in the case of asexual organisms (Fontaneto et al. 2007; Hillis 2007; Fontaneto and Barraclough 2015).

Although the theoretical distinction between species and clades is clear, the origins of new species are necessarily fuzzy, as are the beginnings of all ontological individuals (Ghiselin 1974; Frost and Hillis 1990; de Queiroz 1998). Species rarely split instantaneously into descendant lineages, and different biologists may use different operational criteria to detect a splitting event (de Queiroz 1998, 2007). Widespread, geographically variable, but continuously distributed species and species complexes present a particularly difficult problem for systematists, as their members may exhibit considerable biological divergence at continental scales. In some cases, this variation can be clinal and essentially continuous, with gene flow across the entire species range (e.g., Slatkin and Maddison 1990; Slatkin 1991). In other cases, a species complex might consist of multiple geographically, genetically cohesive, parapatric taxa with little or no gene flow between species where they come into contact (e.g., Hillis 1988). Intermediate conditions are also possible, such that gene flow is restricted but not entirely lacking between particular regional lineages, and such groups present a particular challenge for species delimitation (*Ensatina* salamanders provide a textbook example of such complexity and controversy; Wake and Schneider 1998).

Here, we explore the limitations of a commonly used approach for species delimitation that relies on the multispecies coalescent model (hereafter, MSC-based methods). Despite the known assumptions and limitations of these methods (Leaché and Fujita 2010; Olave et al. 2014; Eberle et al. 2016; Luo et al. 2018; Barley et al. 2018), they are often used in isolation for species delimitation and taxonomic change. We illustrate, using a case study, problems that may arise from inadequate consideration of *a priori* group designations, limited sampling, and lack of attention to contact zones in the context of one MSC-based species delimitation method.

Limitations of the Multispecies Coalescent Model for Species Delimitation

Misapplication of the MSC Model

The MSC has become an important conceptual framework for inferring relationships among species (species trees) from relationships among different genes (gene trees), while taking into account incongruence among gene trees that results from incomplete lineage sorting (Maddison 1997). Because genes trees are not always monophyletic within species lineages, the MSC was introduced as a way to detect recently divergent lineages from collections of gene trees (Knowles and Carstens 2007). However, several biological processes other than incomplete lineage sorting (including hybridization and geographic structuring of populations) can also contribute to discordance among gene trees, and the extent to which the MSC is able to estimate a species tree depends in part on how much discordance is limited to the process of incomplete lineage sorting within species (Sukumaran and Knowles 2017; Barley et al. 2018; Leaché and Fujita 2010).

The MSC has been implemented in several methods for species delimitation (e.g., Yang and Rannala 2010; Ence and Carstens 2011; Camargo et al. 2012; Fujita et al. 2012; Leaché et al. 2014), and some authors have argued that these methods present a more objective approach for testing species hypotheses compared to traditional methods of species delimitation (Leaché and Fujita 2010; Fujita et al. 2012). One commonly used method is Bayesian Phylogenetics and Phylogeography (BPP; Yang and Rannala 2010), which we examine here. Recently, as the limitations of BPP have been explored (Sukumaran and Knowles 2017; Barley et al. 2018; Leaché et al. 2018), it has become evident that this method does not necessarily delimit species boundaries, but may also identify other kinds of genetic structure within species.

Reliance on A Priori Grouping and Problems with Limited Sampling

Many MSC-based methods (including BPP) use clustering algorithms for initial population-level assignment of individuals to groups which are subsequently validated using the MSC-based method (see Carstens et al. 2013 for a full review). The number of individuals and loci sampled play a significant role in ensuring programs such as Structure or Structurama (Pritchard et al. 2000; Huelsenbeck et al. 2011) infer appropriate groups for testing (Rittmeyer and Austin 2012; Olave et al. 2014; Hime et al. 2016). Limited geographic sampling can produce the appearance of distinct genetic clusters, even when samples are drawn from continuous clines or geographically structured populations (Hedin et al. 2015; Barley et al. 2018). Consider, for example, two extreme alternatives. In one case, distinct species lineages have a narrow contact zone with little to no gene flow or hybridization. In another case, a single species exhibits a geographic cline with gradual genetic change across geographic space. Distinguishing these two scenarios requires thorough sampling across the cline or contact zone. If sampling is limited and genetic information is obtained only from geographically distant populations, clustering methods may be incapable of distinguishing between these two scenarios (Irwin 2002; Schwartz and McKelvey 2008; Rittmeyer and Austin 2012; Puechmaille 2016; Bradburd et al. 2018).

There has been extensive discussion of the limitations of MSC-based methods. Overall, depending on taxonomic, geographic, and genetic sampling, BPP can yield variable results in delimitation (Setiadi et al. 2011; Olave et al. 2014; Reid et al. 2014; Zhang et al. 2014; Hime et al. 2016; Barley et al. 2018). Here, we extend this literature by providing a reanalysis of an existing data set from a published study (Ruane et al. 2014) to illustrate the impact of using limited data on BPP's ability to delimit species. Particularly, we focus on the ramifications of using nuclear genetic data sets with limited sampling and little phylogenetic signal, combined with a strong conflicting signal from interspecific introgression of mitochondrial DNA, on species delimitation. We emphasize that our analysis is not necessarily a criticism of the MSC-based method BPP itself, but rather its application to inappropriate data sets in species delimitation studies.

Case Study: the Lampropeltis triangulum Complex

Ruane et al. (2014) sought to clarify species boundaries and relationships in the *Lampropeltis triangulum* complex (American milksnakes) using an MSC-based approach and concluded that genetic evidence supported the recognition of seven species in what had traditionally been considered a single species (Williams 1988). Based primarily on results from BPP, Ruane et al. (2014) elevated seven groups in the *L. triangulum* complex to full species status.

Using the Ruane et al. (2014) data set, we show that sparse geographic sampling, combined with a conflicting signal from interspecific introgression of mitochondrial DNA, led to over-splitting of the American milksnake complex. We first detail inconsistencies observed in the *a priori* clustering analyses and consider the information that can be inferred from such analyses, and then examine the insights that can be gained from an examination of gene trees. We then propose reasons that species splits were recognized despite the lack of supporting evidence from the clustering analyses or evidence of any genetic or reproductive gaps between species. Finally, we perform additional tests on two of the newly recognized species that demonstrate the tendency of BPP to over-split species in the case of limited sampling across broad geographic ranges.

A Priori Grouping

As discussed above, individuals are often assigned to groups (or putative species) before input into MSC-based methods like BPP. This is usually accomplished using clustering methods that report the relative support for each individual's assignment into different clusters. Ruane et al. (2014) assigned individuals to clusters in two different ways. First, they used the program Structurama (Huelsenbeck et al. 2011), which searches for deviations from Hardy–Weinberg equilibrium expectations across sampled gene loci, and then assigns individuals to genetic groups that minimize these deviations. Ruane et al. (2014) also constructed a mitochondrial DNA gene tree, from which they identified groups for subsequent population assignment.

Ruane et al. (2014) found that Structurama did not distinguish between their *a priori* geographic groups *gentilis* (western milksnakes) and *triangulum* (eastern milksnakes). When we repeated the Ruane et al. (2014) Structurama analysis (see online Supplementary Appendix 1 available on Dryad at http://dx.doi.org/10.5061/dryad.7hs34mj), the highest support was given to different cluster numbers depending on the run, indicating the data were not informative enough to provide consistent and robust results across different runs (Supplementary Appendix 1 available on Dryad). However, there were a few consistencies. We observed that Structurama almost always grouped Ruane et al.'s (2014) nominal taxa *polyzona*, *abnorma*, and *micropholis* into a single cluster; that *gentilis* and *triangulum* were always assigned to the same cluster, and that *elapsoides* and *annulata* were generally shown as composites of multiple genetic clusters (Supplementary Fig. S1 available on Dryad). The fact that Structurama consistently showed no division between *gentilis* and *triangulum* is especially noteworthy, as this indicates that the samples of these putative taxa, collected thousands of kilometers apart from New York to Montana to Arizona, do not deviate significantly from Hardy–Weinberg equilibrium expectations (for the data examined) across the breadth of North America. Because of this observation, we will be focusing on these two purported taxa for our subsequent analysis.

Gene Tree Phylogenies and Introgression of Mitochondrial Genomes

Ruane et al. (2014) collected data on 11 nuclear genes and one mitochondrial gene. Using the same data set reported by Ruane et al. (2014), we constructed gene trees for all 11 nuclear genes (one representative nuclear gene is shown in Fig. 1, and the rest are shown in Supplementary Fig. S2 and Appendix 2 available on Dryad), as well as the single mitochondrial gene (Supplementary Fig. S3 available on Dryad). The overall amount of divergence between the 11 nuclear genes was low (0.017–0.076 substitutions per site between the most divergent samples of the *L. triangulum* complex), especially compared to the higher divergence of the single mitochondrial gene (0.225 substitutions per site for samples within the *L. triangulum* complex). Among recently divergent species, we would not expect congruence among all gene trees, and some gene trees would not be expected to be monophyletic within species lineages. However, there is no evidence of any consistent nuclear genetic divergence, nor evidence for any reproductive isolation, at the contact zones between some of the purported species recognized by Ruane et al. (2014). For example, geographically closest individuals of *gentilis* and *triangulum* are genetically indistinguishable at every nuclear locus (Supplementary Fig. S2 available on Dryad). The lack of any genetic break at the contact zone of these purported species suggests that their division is an arbitrary split in a population continuum, rather than a break between distinct species.

Closely related species are expected to retain some shared interspecific polymorphisms. Indeed, humans and chimpanzees are known to share genetic polymorphisms that are thought to have arisen in their common ancestor (e.g., Fan et al. 1989). Nonetheless, humans and chimpanzees are also estimated to be diagnostically distinct across 4% of their genomes (Varki and Altheide 2005). Interspecific differences between humans and chimpanzees (which total approximately 125 million nucleotides) far exceed all intraspecific polymorphisms, and only a small percentage of the latter are shared across these species (Varki and Altheide 2005). Georges et al. (2018) emphasized the importance of such diagnostic differences as evidence for species boundaries and lineage independence. In contrast, there are no diagnostic nucleotide differences among the nuclear genes sampled by Ruane et al. (2014) between *gentilis* and *triangulum*. Given that there is no evidence of deviations from Hardy–Weinberg equilibrium expectations (as shown in the Structurama analyses), and no evidence of even a single nuclear gene that consistently differs between these two purported species, then what is the basis for hypothesizing the existence of these species lineages? Is there any reason to expect any biological differences between two "species" that exhibit no known genetic differences?

In contrast to the low levels of nuclear divergence discussed above, upon reconstruction of the mitochondrial gene tree, Ruane et al. (2014) found clear evidence for multiple captures of *L. alterna* mitochondrial DNA within western North American populations of the *L. triangulum* complex (i.e., the populations referred to as the forms *gentilis* and *annulata*; Supplementary Fig. S3 available on Dryad). These western populations of *L. triangulum* have mitochondrial haplotypes that are deeply embedded within those of *L. alterna*, which in turn has a mitochondrial genome that is more closely related to species of the *L. getula* complex and *L. extenuata* than to the eastern North American populations of *L. triangulum* (Supplementary Fig. S3 available on Dryad). These introgression events appear to have happened several times and are still ongoing (note the nearly

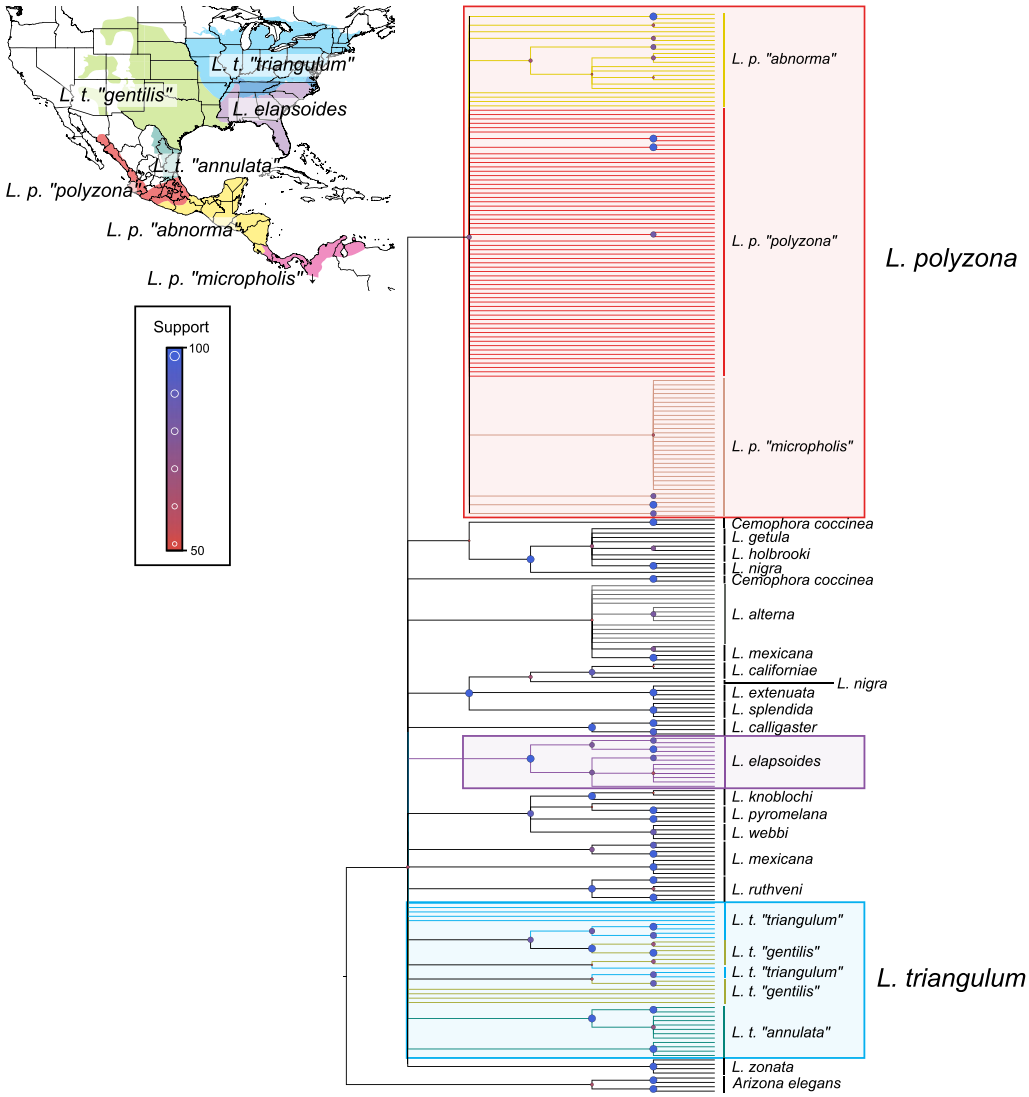

FIGURE 1. Majority-rule consensus gene tree constructed with nuclear gene 2CL8, used as a representative tree to illustrate consistencies observed across all 11 nuclear gene trees. From this tree, it is clear that Central and South American milksnake lineages (*polyzona, abnorma,* and *micropholis*) form a monophyletic cluster with little resolution. *L. elapsoides* is consistently recovered as a monophyletic lineage, while remaining U.S. lineages (*triangulum, gentilis,* and *annulata*) are rarely resolved and exhibit no diagnostic differences. Remaining gene trees are given in Supplementary Figure S2 available on Dryad.

identical mitochondrial DNA haplotypes of *L. alterna* and *L. triangulum* where the two coexist in Val Verde County, TX, USA; Supplementary Fig. S3 available on Dryad). Indeed, the only consistent genetic difference between *gentilis* and *triangulum* is that individuals assigned to *gentilis* have introgressed mitochondrial DNA from *L. alterna*, whereas individuals assigned to *triangulum* do not. No single nucleotide from any of the sampled nuclear genes follows this same pattern.

BPP Analysis

Ruane et al. (2014) first ran BPP using Structurama assignments as terminal lineages on guide trees, resulting in high support for six lineages within the *L. triangulum* complex (recall that *gentilis* and *triangulum* were initially treated as a single lineage by Ruane et al. 2014 based on their Structurama assignments). We found the same result when we performed the same analysis using unguided BPP (Yang and Rannala 2014; posterior probability = 99.2%; Supplementary Table S1 and Appendix 3 available on Dryad).

Given the divergent mitochondrial DNA haplotypes in western populations of the combined *triangulum–gentilis* lineage (the introgressed haplotypes from *L. alterna*; Supplementary Fig. S3 available on Dryad), Ruane et al. (2014) next tested whether BPP would support a division between *triangulum* and *gentilis*, despite their Structurama results. To conduct this test, they ran BPP with a guide tree generated from their mitochondrial gene tree, assigning these two lineages to different groups. BPP strongly supported this split

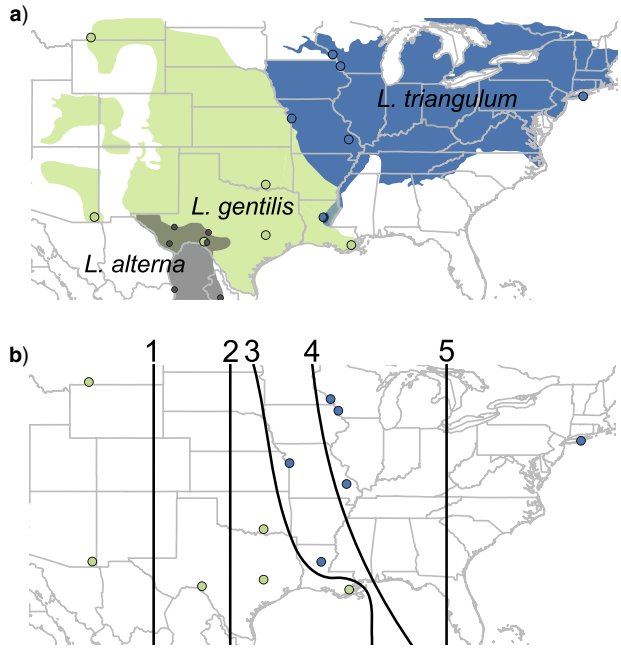

FIGURE 2. Results and group assignment from five runs of unguided BPP. a) Points represent samples with nuclear gene data from Ruane et al. (2014), with the same data used here, and ranges shaded according to the Ruane et al. (2014) final population assignment. b) Population assignment between each of five runs of BPP (Supplementary Appendix 1 available on Dryad).

as well, while still differentiating *L. alterna*, thus leading Ruane et al. (2014) to propose *L. gentilis* and *L. triangulum* as distinct species.

Given that no nuclear genes (Fig. 1 and Supplementary Fig. S2 available on Dryad) show evidence of genetic differentiation between *gentilis* and *triangulum*, and even Structurama fails to separate individuals of these taxa, the only basis for distinguishing these taxa appears to be the introgressed mitochondrial DNA. Similar cases of mitochondrial DNA capture have confounded species delimitation in other taxa (e.g., polar bears versus brown bears: Miller et al. 2012; freshwater mussels: Chong et al. 2016), and deep intraspecific polymorphisms of mitochondrial DNA have similarly affected species delimitation in other studies (e.g., Folt et al. 2019). Without the introgressed mitochondrial DNA, there would have been no basis for testing a split between *gentilis* and *triangulum*. Therefore, we tested if the support from BPP for the division of *gentilis* and *triangulum* was limited to the distribution of the introgressed *L. alterna* DNA (as examined by Ruane et al. 2014; see Fig. 2a), or if it was simply a reflection of the geographic proximity of samples taken from across a broad geographic distribution. In other words, does BPP support any east–west split of the *gentilis–triangulum* populations at any point in the combined continental distribution of these forms, or is the split tested by Ruane et al. (2014) at the break in introgressed mitochondrial DNA distinctive?

Using the nuclear data set from Ruane et al. (2014), we tested five east–west splits of the *gentilis–triangulum* populations using BPP, including the split tested by Ruane et al. (2014; Split 3 in Fig. 2b), as well as two splits farther west (Splits 1 and 2 in Fig. 2b), and two splits farther east (Splits 4 and 5 in Fig. 2b; Supplementary Appendix 3 available on Dryad). If the support from BPP reported by Ruane et al. (2014) for Split 3 reflects a real split between species, and is not simply a reflection of genetic similarity of geographically proximate populations on either side of an arbitrary line, then, we would expect much stronger support from BPP for Split 3 than for Splits 1, 2, 4, or 5. In contrast, if BPP is simply supporting any split that results in clustering of two groups of geographically proximate samples from a broad distribution of a single species, we would expect to see support for all five splits in Figure 2b. We found the latter result: regardless of the geographic split between populations, BPP indicated very high support (posterior probability = 100% for Splits 1–4, and posterior probability > 96% for Split 5; Supplementary Table S1 available on Dryad) for all five of the east–west splits of the *gentilis–triangulum* cline.

Our empirical results support the simulations of Barley et al. (2018), who demonstrated that if samples are taken from separated geographic localities from a single species that exhibits isolation by distance, BPP consistently supports the separated geographic clusters as distinct species. That result is in contrast to the simulations of Zhang et al. (2011), who simulated a stepping-stone model and found that only in cases of relatively high migration rates did BPP falsely recover low support for a single species. As noted by Barley et al. (2018), the results from theoretical studies depend largely on parameters used in the respective simulations. Our results suggest that the simulations conducted by Barley et al. (2018) better match the empirical system studied by Ruane et al. (2014) than do the simulations of Zhang et al. (2011). Note that even if the split between *gentilis* and *triangulum* reported by Ruane et al. (2014) represented an actual species split, BPP also supports all the other east–west geographic splits shown in Figure 2b.

We do not suggest that any of the alternative species splits in Figure 2b represent "better" species delimitation in the *L. triangulum* complex compared to those examined by Ruane et al. (2014). Rather, our analysis merely demonstrates that BPP supports virtually *any* geographic partition of samples in this potential continental cline as "species." But clearly, splits 1–5 in Figure 2b cannot all be true species splits, as they each are mutually inconsistent with one another. BPP does not provide stronger support for the *gentilis–triangulum* split than it does for other east–west splits of the samples.

The Importance of Contact Zones

When splits are hypothesized within an otherwise continuous distribution, contact zone analyses have

traditionally been used to assess the degree of genetic isolation and gene flow between the putative taxa (Barton and Hewitt 1985; Derryberry et al. 2014). Systematists need to distinguish between widespread, clinal geographic variation within a species on one hand, versus distinct genetic and reproductive breaks between species on the other. This is especially important when species are thought to be distributed parapatrically, such that the species contact one another along narrow zones of potential gene flow. In such cases, the study of contact zones can reveal if (a) hybridization between the putative species is absent or rare; (b) the contact zones act as "genetic sinks" (thus restricting gene flow between the putative species); or (c) there is broad gene flow and integration between the putative species at the contact zone. Case (a) is uncontroversial, as it is consistent with virtually any concept of species (i.e., there is clear evidence that the taxa are reproductively isolated, evolutionary distinct, and independent lineages). In recent decades, many biologists have argued that case (b), or evidence of a narrow hybrid zone that acts as a "genetic sink" that strongly restricts gene flow between species, is also consistent with the hypothesis of distinct species (e.g., Sage and Selander 1979; Hafner et al. 1983; Yanchukov et al. 2006). In contrast, case (c) refutes the hypothesis that the lineages are evolving independently from one another (as there are no reproductive or genetic breaks between the lineages to support their independent evolution).

Examining the population genetic structure at contact zones can also determine selective forces that may be playing roles in driving, or maintaining, divergence (Sobel and Streisfeld 2015; Bertrand et al. 2016). Many approaches, genetic and otherwise, have been developed for examining contact zone interactions (e.g., Gompert and Buerkle 2010; Derryberry et al. 2014), although many species delimitation studies may require additional sampling for such an analysis.

Species Delimitation and Taxonomic Recommendations

Ruane et al. (2014) stated that they followed the "general lineage species concept" of de Queiroz (1998, 2007). In these two papers, de Queiroz argued that virtually all species concepts treat species as "separately evolving metapopulation lineages" that simply use different lines of evidence to assess the independence and isolation of lineages. In other words, virtually all "species concepts" are conceptualizing the same entities—namely, the individual, independent, evolving lineages of life, within which organisms typically mate and exchange genes (as we described in the opening of this article).

Species delimitation is typically a two-step process (see Hillis 2019). Taxonomists first group organisms into putative taxa using one of several criteria. These include 1) correlated diagnostic characters (including morphological, genetic, or behavioral attributes), which are often assessed in a hierarchical phylogenetic analysis; 2) deviations from Hardy–Weinberg equilibrium (as conducted, for example, by the programs Structure and Structurama); and 3) multivariate analyses that assess overall divergence, such as principal components analysis. These tests are all ways of identifying groups of individuals that appear to be different from one another. However, differences arise within species as well as between them, so a second step is needed to assess if the observed differences are evidence of independently evolving lineages, or if the observed variation simply represents geographic or population variation within species. If the groups in question come into geographic contact, then taxonomists typically assess lineage independence by looking for direct or indirect evidence of reproductive barriers between the groups at contact zones. Indirect evidence may include sharp geographic breaks in suites of morphological, genetic, and/or behavioral characters at the contact zones; direct evidence may include behavioral assessments of reproductive interactions between the putative species.

MSC-based approaches have also been used to assess the evolutionary independence of lineages (as in Ruane et al. 2014), but as shown in Barley et al. (2018), BPP does not appear to discriminate adequately between geographic clinal structure versus species boundaries. Although our results appear to be an empirical example of this scenario (Fig. 2b), without adequate sampling at purported contact zones there is no way to distinguish these two possibilities using BPP alone.

Despite the inability of BPP to distinguish between clinal variation versus speciation, we can use the data collected by Ruane et al. (2014) to ask if there is any evidence for sharp genetic or reproductive breaks among the various groups that they examined. Ruane et al. (2014) also presented an analysis to summarize their data, in the form of a SplitsTree analysis (Fig. 3; Huson and Bryant 2006). This tree does not represent any single gene tree, but is instead a summary of support and counter-support for various clusters of individuals examined by Ruane et al. (2014) across all examined loci. Individuals that are nearly identical across all loci are located adjacent to one another (separated by small branch lengths) on this tree; in contrast, individuals that differ across many loci are well separated. Therefore, we can look at the contact zones between each purported taxon, and ask if geographically adjacent individuals in different purported taxa exhibit any evidence for the genetic or reproductive breaks that are expected from separately evolving lineages. If there are none, then there is no reason to hypothesize a break between distinct species rather than a continuous geographic cline.

Figure 3 shows three hypotheses for species delimitation in American milksnakes projected on the Ruane et al. (2014) SplitsTree analysis, with a depiction of the distribution of the purported taxa. The one-species hypothesis of Williams (1988; shown in Fig. 3a) is refuted by two lines of evidence: first, there are far larger genetic gaps among subgroups of his *L. triangulum* than there are between those subgroups

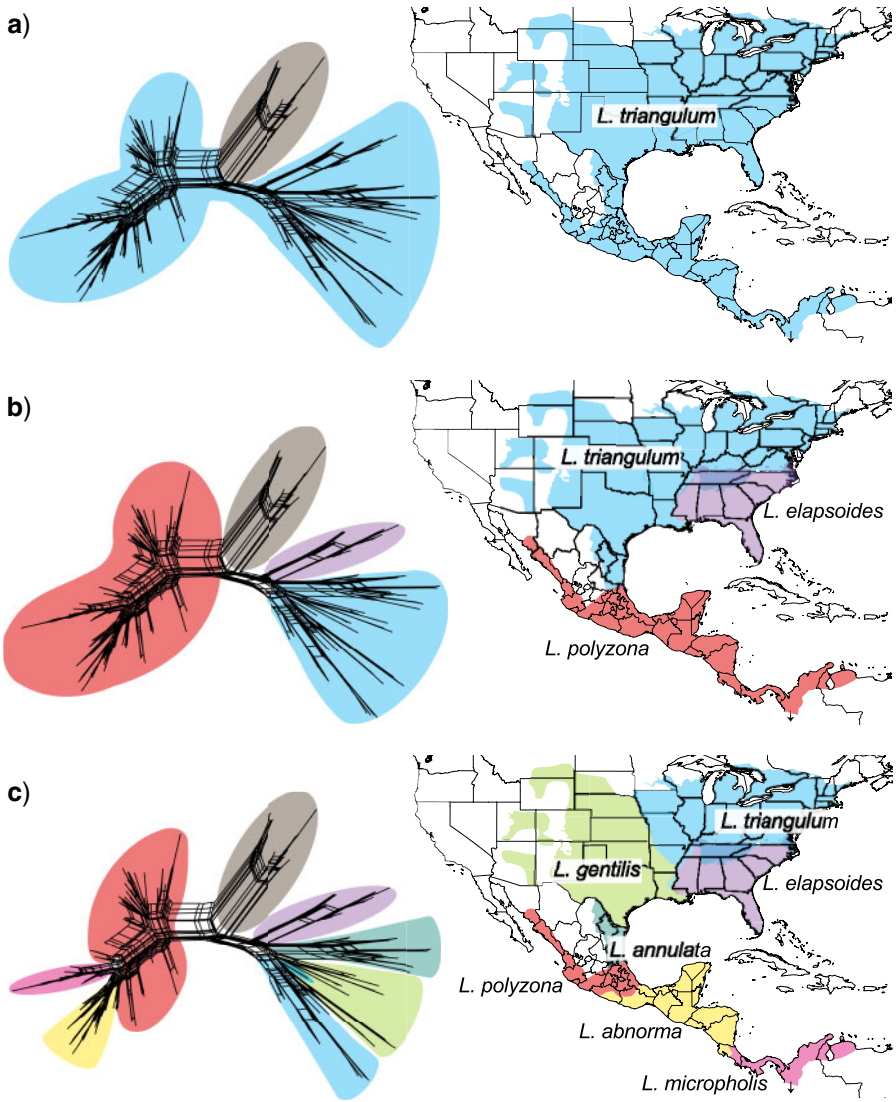

FIGURE 3. Three hypotheses for species delimitation in milksnakes (*Lampropeltis triangulum* complex) with SplitsTree networks (Huson and Bryant 2006) and ranges colored based on the proposed species given in each hypothesis (adapted from Ruane et al. 2014). The *Lampropeltis alterna* lineage shown in gray (not indicated on range map) is included because of its relevance to mitochondrial introgression events. a) Hypothesis 1: American milksnakes represent a single, polytypic species across their entire range (Williams 1988); b) Hypothesis 2: three species, *L. triangulum*, *L. elapsoides*, and *L. polyzona* (similar to Blanchard 1921); c) Hypothesis 3: the seven species proposed by Ruane et al. (2014).

and other well differentiated, sympatric species (i.e., *L. alterna*). Second, where these subgroups of Williams' *L. triangulum* come into contact, they are sympatric, and yet maintain large genetic gaps between individuals. Thus, we agree with Ruane et al. (2014) in rejecting the one-species hypothesis of Williams (1988).

Figure 3b presents an alternative taxonomic hypothesis that addresses the problems noted above, and divides *L. triangulum* of Williams (1988) into three distinct species: *L. triangulum*, *L. elapsoides*, and *L. polyzona*. This hypothesis is almost identical to the arrangement proposed by Blanchard (1921), although he noted that collections at the time were not sufficient in lower Central America to firmly establish the relationship between the nominal forms *L. polyzona* and *L. micropholis*, and he tentatively treated those two species as distinct as well, pending further collection of intermediate populations. There are substantial, consistent genetic breaks across multiple loci among all three of the species recognized in this hypothesis. In addition, where any two of these three species come into geographic contact, there are areas of known sympatry, accompanied by genetic divergence across multiple loci. *L. triangulum* and *L. elapsoides* are known to occur sympatrically, with little or no hybridization, across a broad area of parts of Kentucky, Tennessee, Alabama, Georgia, North Carolina, and Virginia in the United States (indeed, this region of sympatry was discussed by Williams 1988). The known area of sympatry between *L. triangulum* and *L. polyzona* in northern Veracruz, Mexico is much smaller, with the two species reported together from just a single locality (also reported by Williams

1988). Thus, all the genetic and geographic data appear to support the recognition of these three species.

In contrast, the remaining taxa recognized by Ruane et al. (2014; Fig. 3c) exhibit no known areas of sympatry or any evidence of sharp genetic breaks at or near purported contact zones. Instead, individuals on either side of purported contact zones (other than the ones noted in Fig. 3b) are genetically much more similar to one another than they are to other geographically distant individuals in their own taxon. This is not consistent with the expectation for independently evolving lineages. As with the human/chimpanzee example discussed earlier, we might expect similarities in occasional genes through independent lineage sorting, but we would still expect large genetic gaps across most loci in comparisons of individuals drawn from different species. No such genetic gaps exist between geographically adjacent samples of *micropholis–abnorma–polyzona*, or between geographically adjacent samples of *annulata–gentilis–triangulum* (Figs. 2b and 3c and Supplementary Figure S2 available on Dryad). These findings are also largely consistent with the Structurama results (Supplementary Fig. S1 available on Dryad), except that *annulata* does appear to show significant Hardy–Weinberg equilibrium deviations from the *gentilis–triangulum* grouping. However, such deviations are not surprising given the large geographic distance between the samples examined by Ruane et al. (2014) of *annulata* versus *gentilis–triangulum* (e.g., the distance between the closest samples examined for nuclear genes of *annulata* and *gentilis–triangulum* is ~485 km). Better sampling is needed to determine if the relatively small genetic differences and Hardy–Weinberg deviations between these groups are indicative of geographic clines or species breaks.

Despite frequent discussion of integrative approaches for species delimitation (Dayrat 2005; Leaché et al. 2009; Padial et al. 2010; Schlick-Steiner et al. 2010; Fujita et al. 2012; Derkarabetian and Hedin 2014; Huang and Knowles 2016; Renner 2016), researchers sometimes use limited data and rely on results generated by a single analysis to delimit species. Consideration of morphological, behavioral, and ecological data, particularly including analyses at contact zones, is a critical part of testing species hypotheses (Zhang et al. 2011; Edwards and Knowles 2014; Pante et al. 2015; Solís-Lemus et al. 2015). We recommend against taxonomic changes on the basis of analyses of limited geographic samples, and demonstrate that in such cases BPP can support many groupings that are inconsistent with species.

Although we present an alternative hypothesis to that presented by Ruane et al. (2014) in Figure 3b, we emphasize that the data presented by Ruane et al. (2014) are inadequate to fully examine the species boundaries in this group. The existing data do appear to support the species delimited in Figure 3b, but it is certainly possible that additional genetic and geographic sampling will demonstrate the existence of additional species boundaries in this group. However, we see no convincing evidence from the data presented by Ruane et al. (2014) to support the additional species recognized in Figure 3c.

Species delimitation is not a simple process. In well-studied, widely distributed taxa, species designations should incorporate multiple sources of evidence regarding geographic variation (of genes, morphology, and behavior), reproductive isolation, and gene flow. New species designations, especially of well-studied groups, are best made after careful consideration of all sources of relevant evidence. Nomenclatural changes to well-studied groups should be made only after due consideration of all available data (e.g., Setiadi et al. 2011; Barley et al. 2013; Hedin et al. 2015; Pante et al. 2015; Pyron et al. 2016; Folt et al. 2019). Although a conservative approach to taxonomic change can risk underestimating diversity (Padial et al. 2010), this is preferable to making poorly supported taxonomic changes with each new data set and analysis (Hillis 2019).

Supplementary Material

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.7hs34mj.

Acknowledgments

We thank Sara Ruane and Frank Burbrink for guidance on the milksnake reanalysis. We also thank Carole Baldwin, Anthony Barley, Peter Beerli, David Cannatella, Kevin de Queiroz, Harry Greene, Tracy Heath, Aleta Quinn, Jordan Satler, Robert Thomson, and April Wright for comments and conversation related to the manuscript. Finally, we would like to thank Richard Glor, Adam Leaché, and 13 anonymous reviewers for helpful suggestions and critique that improved this article.

References

Barley A.J., Brown J.M., Thomson R.C. 2018. Impact of model violations on the inference of species boundaries under the multispecies coalescent. Syst. Biol. 67:269–284.

Barley A.J., White J., Diesmos A.C., Brown R.M. 2013. The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skins. Evolution. 67:3556–3572.

Barton N.H., Hewitt G.M. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst. 16:113–148.

Bertrand J.A.M., Delahaie B., Bourgeois Y.X.C., Duval T., García-Jiménez R., Cornuault J., Pujol B., Thébaud C., Milá B. 2016. The role of selection and historical factors in driving population differentiation along an elevational gradient in an island bird. J. Evol. Biol. 29:824–836.

Blanchard F.N. 1921. A revision of the king snakes: genus Lampropeltis. Bull. U.S. Natl. Mus. 114:1–260.

Bradburd G.S., Coop G.M., Ralph P.L. 2018. Inferring continuous and discrete population genetic structure across space. Genetics. 210:33–52.

Camargo A., Morando M., Avila L.J., Sites J.W. Jr. 2012. Species delimitation with ABC and other coalescent-based methods: a test of accuracy with simulations and an empirical examples with lizards of the *Liolaemus darwinii* complex (Squamata: Lacertidae). Evolution. 66:2834–2849.

Carstens B.C., Pelletier T.A., Reid N.M., Satler J.D. 2013. How to fail at species delimitation. Mol. Ecol. 22:4369–4383.
Chong J.P., Harris J.L., Roe K.J. 2016. Incongruence between mtDNA and nuclear data in the freshwater mussel genus Cyprogenia (Bivalvia: Unionidae) and its impact on species delineation. Ecol. Evol. 6:2439–2452.
Dayrat B. 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. Lond. 85:407–415.
de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard D.J., Berlocher S.H., editors. Endless forms: species and speciation. Oxford: Oxford University Press. p. 57–75.
de Queiroz K. 2007. Species concepts and species delimitation. Syst. Biol. 56:879–886.
Derkarabetian S., Hedin M. 2014. Integrative taxonomy and species delimitation in harvestmen: a revision of the western North American genus Sclerobunus (Opiliones: Laniatores: Travunioidea). PLoS One. 9:e104982.
Derryberry E.P., Derryberry G.E., Maley J.M., Brumfield R.T. 2014. HZAR: hybrid zone analysis using an R software package. Mol. Ecol. Res. 14:652–663.
Durand E., Jay F., Gaggiotti O.E., François O. 2009. Spatial inference of admixture proportions and secondary contact zones. Mol. Biol. Evol. 26:1963–1973.
Eberle J., Warnock R.C.M., Ahrens D. 2016. Bayesian species delimitation in Pleophylla chafers (Coleoptera)—the importance of prior choice and morphology. BMC Evol. Biol. 16:94.
Edwards D.L., Knowles L.L. 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proc. R. Soc. Lond. B Biol. Sci. 281:20132765.
Ence D.D., Carstens B.C. 2011. SpedeSTEM: a rapid and accurate method for species delimitation. Mol. Ecol. Res. 11:473–480.
Fan W.M., Kasahara M., Gutknecht J., Klein D., Mayer W.E., Jonker M., Klein J. 1989. Shared class II MHC polymorphisms between humans and chimpanzees. Hum. Immunol. 26:107–121.
Folt B., Bauder J., Spear S., Stevenson D., Hoffman M., Oaks J.R., Wood P.L. Jr., Jenkins C., Steen D.A., Guyer C. 2019. Taxonomic and conservation implications of population genetic admixture, mito-nuclear discordance, and male-biased dispersal of a large endangered snake, Drymarchon couperi. PLoS One. 14:e0214439.
Fontaneto D., Herniou E.A., Boschetti C., Caprioli M., Melone G., Ricci C., Barraclough T.G. 2007. Independently evolving species in asexual bdelloid rotifers. PLoS Biol. 5:914–921.
Fontaneto D., Barraclough T.G. 2015. Do species exist in asexuals? Theory and evidence from bdelloid rotifers. Integr. Comp. Biol. 55:253–263.
Frost D.R., Hillis D.M. 1990. Species in concept and practice: herpetological applications. Herpetologica. 46:87–104.
Fujita M.K., Leaché A.D., Burbrink F.T., McGuire J.A., and Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27:480–488.
Georges A., Gruber B., Pauly G.B., White D., Adams M., Young M.J., Kilian A., Zhang X., Shaffer H.B., Unmack P.J. 2018. Genomewide SNP markers breathe new life into phylogeography and species delimitation for the problematic short-necked turtles (Chelidae: *Emydura*) of eastern Australia. Mol. Ecol. 27:5195–5213.
Ghiselin M.T. 1974. A radical solution to the species problem. Syst. Zool. 23:536–544.
Gompert Z., Buerkle C.A. 2010. Introgress: a software package for mapping components of isolation in hybrids. Mol. Ecol. Res. 10:378–384.
Hafner J.C., Hafner D.J., Patton J.L., Smith M.F. 1983. Contact zones and the genetics of differentiation in the pocket gopher *Thomomys bottae* (Rodentia: Geomyidae). Syst. Zool. 32:1–20.
Hedin M., Carlson D., Coyle F. 2015. Sky island diversification meets the multispecies coalescent—divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. Mol. Ecol. 24:3467–3484.
Hillis D.M. 1988. Systematics of the *Rana pipiens* complex: puzzle and paradigm. Annu. Rev. Ecol. Syst. 19:39–63.
Hillis D.M. 2007. Asexual evolution: can species exist without sex? Curr. Biol. 17:R543–R544.
Hillis D.M. 2019. Species delimitation in herpetology. J. Herpetol. 53:3–12.
Hime P.M., Hotaling S., Grewelle R.E., O'Neill E.M., Voss S.R., Shaffer H.B., Weisrock D.W. 2016. The influence of locus number and information content on species delimitation: an empirical test case in an endangered Mexican salamander. Mol. Ecol. 25:5959–5974.
Huang J., Knowles L.L. 2016. The species versus subspecies conundrum: quantitative delimitation from integrating multiple data types within a single Bayesian approach in Hercules beetles. Syst. Biol. 65:685–699.
Huelsenbeck J.P., Andolfatto P., Huelsenbeck E.T. 2011. Structurama: Bayesian inference of population structure. Evol. Bioinform. 7:55–59.
Huson D.H., Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254–267.
Irwin D.E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution. 56:2383–2394.
Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. Syst. Biol. 56:887–895.
Leaché A.D., Fujita M.K. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. Lond. B Biol. Sci. 277:3071–3077.
Leaché A.D., Fujita M.K., Minin V.N., Bouckaert R.R. 2014. Species delimitation using genome-wide SNP data. Syst. Biol. 63:534–542.
Leaché A.D., Koo M.S., Spencer C.L., Papenfuss T.J., Fisher R.N., McGuire J.A. 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proc. Natl. Acad. Sci. U.S.A. 106:12418–12423.
Leaché A.D., Zhu T., Rannala B., Yang Z. 2018. The spectre of too many species. Syst. Biol. 68:168–181.
Luo A., Ling C., Ho S.Y.W., Zhu C. 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst. Biol. 10.1093/sysbio/syy011.
Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.
Mayden R.L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge M.F., Dawah H.A., Wilson M.R., editors. Species: The Units of Biodiversity. London: Chapman and Hall. p. 381–423.
Miller W., Schuster S.C., Welch A.J., Ratan A., Bedoya-Reina O.C., Zhao F., Kim H.L., Burhans R.C., Drautz D.I., Wittekindt N.E., Tomsho L.P., Ibarra-Laclette E., Herrera-Estrella L., Peacock E., Farley S., Sage G.K., Rode K., Obbard M., Montiel R., Bachmann L., Ingólfsson Ó., Aars J., Mailund T., Wiig Ø., Talbot S.L., Lindqvist C. 2012. Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. Proc. Natl. Acad. Sci. U.S.A. 109:E2382–E2390.
Olave M., Sola E., Knowles L.L. 2014. Upstream analyses create problems with DNA-based species delimitation. Syst. Biol. 63:263–271.
Padial J.M., Miralles A., De la Riva I., Vences M. 2010. The integrative future of taxonomy. Front. Zool. 7:16.
Pante E., Puillandre N., Viricel A., Arnaud-Haond S., Aurelle D., Castelin M., Chenuil A., Destombe C., Forcioli D., Valero M., Viard F., Samad S. 2015. Species are hypotheses: avoid connectivity assessments based on pillars of sand. Mol. Ecol. 24:525–544.
Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155:945–959.
Puechmaille S.J. 2016. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol. Ecol. Res. 16:608–627.
Pyron R.A., Hsieh F.W., Lemmon A.R., Lemmon E.M., Hendry C.R. 2016. Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (*Storeria*). Zool. J. Linn. Soc. Lond. 177:937–949.
Reid N.M., Hird S.M., Brown J.M., Pelletier T.A., McVay J.D., Satler J.D., Carstens B.C. 2014. Poor fit to the multi-species coalescent model is widely detectable in empirical data. *Syst. Biol.* 63:322–333.
Renner S.S. 2016. A return to Linnaeus's focus on diagnosis, not description: the use of DNA characters in the formal naming of species. Syst. Biol. 65:1085–1095.

Rittmeyer E.N., Austin C.C. 2012. The effects of sampling on delimiting species from multi-locus sequence data. Mol. Phylogenet. Evol. 65:451–463.
Ruane S., Bryson R.W. Jr., Pyron R.A., Burbrink F.T. 2014. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. Syst. Biol. 63:231–250.
Sage R.D., Selander R.K. 1979. Hybridization between species of the *Rana pipiens* complex in central Texas. Evolution. 33:1069–1088.
Schlick-Steiner B.C., Steiner F.M., Seifert B., Stauffer C., Christian E., Crozier R.H. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. Annu. Rev. Entomol. 55:421–438.
Schwartz M., McKelvey K. 2008. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conserv. Genet. 10:441–452.
Setiadi M.I., McGuire J.A., Brown R.M., Zubairi M., Iskander D.T., Andayani N., Supriatna J., Evans B.J. 2011. Adaptive radiation and ecological opportunity in Sulawesi and Philippine fanged frog (*Limnonectes*) communities. Am. Nat. 178:221–240.
Slatkin M. 1991. Inbreeding coefficients and coalescence times. Genet. Res. 58:167–175.
Slatkin M. and Maddison W.P. 1990. Detecting isolation by distance using phylogenies of genes. Genetics. 126:249–260.
Sobel J.M., Streisfeld M.A. 2015. Strong premating reproductive isolation drives incipient speciation in *Mimulus aurantiacus*. Evolution. 69:447–461.
Solís-Lemus C., Knowles L.L., Ané C. 2015. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution. 69:492–507.
Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. Proc. Natl. Acad. Sci. U.S.A. 114:1607–1612.
Templeton A. 1989. The meaning of species and speciation: a population genetics approach. In: Otte D., Endler J.A., editors. Speciation and its Consequences. Sunderland: Sinauer Associates. p. 3–27.
Varki A., Altheide T.K. 2005. Comparing the human and chimpanzee genomes: searching for needles in a haystack. Genome Res. 15:1746–1758.
Wake D.B., Schneider C.J. 1998. Taxonomy of the plethodontid salamander genus *Ensatina*. Herpetologica. 54:279–298.
Wiley E.O. 1978. The evolutionary species concept reconsidered. Syst. Zool. 27:17–26.
Williams K.L. 1988. Systematics and natural history of the American milk snake, *Lampropeltis triangulum*. 2nd ed. Milwaukee (WI): Milwaukee Public Museum.
Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. U.S.A. 107:9264–9269.
Yang Z., Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 31: 3125–3135.
Yanchukov A., Hofman S., Szymura J.M., Mezhzherin S.M., Morozov-Leonov S.Y., Barton N.H., Nürnberger B.N. 2006. Hybridization of *Bombina bombina* and *B. variegata* (Anura, Discoglossidae) at a sharp ecotone in western Ukraine: Comparisons across transects and over time. Evolution. 60:583–600.
Zhang C., Rannala B., Yang Z. 2014. Bayesian species delimitation can be robust to guide-tree inference errors. Syst. Biol. 63: 993–1004.
Zhang C., Zhang D.X., Zhu T., Yang Z. 2011. Evaluation of a Bayesian coalescent method of species delimitation. Syst. Biol. 60:747–761.