Systematics and Biodiversity (2020), **18(3)**: 269–280

Taylor & Francis
Taylor & Francis Group

Research Article

Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance

NICHOLAS A. MASON[1]*, NICHOLAS K. FLETCHER[1], BRIAN A. GILL[2,3], W. CHRIS FUNK[2,3] & KELLY R. ZAMUDIO[1]

[1]Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall, Ithaca, New York 14853, USA
[2]Department of Biology, Colorado State University, Fort Collins, Colorado 80523, USA
[3]Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado 80523, USA

(Received 17 September 2019; accepted 6 December 2019)

Species are a fundamental unit of biodiversity that are delimited via genetic data and coalescent-based methods with increasing frequency. Despite the widespread use of coalescent-based species delimitation, we do not fully understand the sensitivity of these methods to potential sources of bias and violations of their underlying assumptions. One implicit assumption of coalescent-based species delimitation is that geographic sampling is adequate and representative of genetic variation among populations within the lineage of interest. Yet exhaustive geographic sampling is logistically difficult, if not impossible, for many taxa that span large geographic expanses or occupy remote regions. Here, we examine the impact of geographic sampling on the output of Bayes-factor delimitation with SNAPP, a popular coalescent-based species delimitation pipeline. First, we demonstrate the problematic nature of sparse geographic sampling and isolation by distance for species delimitation using simulated data sets of populations connected by different levels of gene flow. We then examine whether similar trends are present in an empirical dataset of *Andesiops* mayflies (Ephemeroptera: Baetidae) from a high elevation transect in the Ecuadorian Andes. In both the simulated and empirical analyses, we systematically exclude geographically intermediate sites to quantify the impact of geographic sampling and isolation by distance on coalescent-based species delimitation. We find that removing intermediate sites with genetically admixed individuals incorrectly favors multi-species delimitation scenarios. Oversplitting is especially pronounced when isolation by distance is strong, but exists even when gene flow among neighboring populations is relatively high. These findings highlight the importance of adequate geographic sampling in species delimitation and urge caution in interpreting the output of such methods when species' distributions are sparsely sampled and in systems characterized by strong patterns of isolation by distance.

Key words: gene flow, isolation by distance, multispecies coalescent, speciation, species delimitation, taxonomy

Introduction

Species are a core unit of biological analyses and species delimitation is a fundamental goal of systematic biology. Accurate species-level classifications enable studies of phylogenetic relationships, lineage and phenotypic diversification, and biogeographic history, among other long-standing lines of inquiry. Despite the far-reaching importance of alpha taxonomy, species delimitation—or the process of determining whether one or two species exist in a given lineage—is complicated and often controversial in practice (Bauer et al., 2011; Carstens, Pelletier, Reid, & Satler, 2013; Fujita & Leaché, 2011; Heller, Frandsen, Lorenzen, & Siegismund, 2013). This ongoing controversy stems in large part from disagreements over species concepts, speciation criteria, and how morphological, ecological, behavioral, and genetic data should be evaluated and integrated into delimitation pipelines and decisions (Barrowclough, Cracraft, Klicka, & Zink, 2016; De Queiroz, 2007; Gill, 2014; Toews, 2014). Continued improvements in our ability to incorporate different types of data into species delimitation pipelines should lead to a more objective and stable classification of the tree of life. However, our understanding of how biases

Correspondence to: Nicholas A. Mason. E-mail: nicholas.albert.mason@gmail.com
*Current address: Nicholas A. Mason Museum of Vertebrate Zoology, 3101 Valley Life Sciences Building, University of California, Berkeley, CA 94720, USA.

ISSN 1477-2000 print / 1478-0933 online

https://dx.doi.org/10.1080/14772000.2020.1730475

Published online 14 Apr 2020

and violations of the assumptions underlying coalescent-based methods affect delimitation inferences have not kept pace with the implementation of these methods themselves.

Systematists have recently developed an array of methods to evaluate empirical support for alternative species delimitation scenarios using genetic data. Methods such as Bayesian Phylogenetics and Phylogeography (BPP; Yang, 2015), spedeSTEM (Ence & Carstens, 2011) and Bayes factor Delimitation using whole loci (BFD; Grummer, Bryson, & Reeder, 2014) or SNPs (BFD*; Leaché, Fujita, Minin, & Bouckaert et al., 2014) leverage the multispecies coalescent model (MSC; Rannala & Yang, 2003) to quantify support for alternative species delimitation scenarios. In brief, the multispecies coalescent model is a statistical framework that incorporates multi-locus data to evaluate alternative hypotheses of divergence among lineages while allowing for gene tree discordance under a neutral model of genetic drift (Fujita, Leaché, Burbrink, McGuire, & Moritz, 2012; Knowles & Carstens, 2007; Yang & Rannala, 2010). Coalescent theory incorporates demographic parameters, such as effective population sizes ($\theta = 4N_e\mu$) and time since lineage splitting (τ), to model the retrospective process of identity by descent via coalescence among alleles sampled from a given lineage. Users can then compare statistical support for alternative coalescent models with varying species delimitation hypotheses given the data at hand. Often, subspecific or historical taxonomic treatments or single-locus data sets guide putative delimitation scenarios, but species delimitation approaches based on genetic data and the multispecies coalescent model have also proven especially useful for cryptic or understudied taxa for which no data are available to establish *a priori* alternative delimitation scenarios based on phenotypic variation or infraspecific and interspecific taxonomy (Frankham, 2010). Thus, multispecies coalescent methods have been widely adopted as an empirical criterion to be evaluated alongside other data in integrative species delimitation pipelines (Fujita & Leaché, 2011; Fujita et al., 2012).

Despite their widespread use, we do not fully understand the sensitivity of coalescent-based species delimitation methods to violations of their underlying models and biases that may arise from different forms of missing data. For example, most MSC-based species pipelines methods assume that gene flow does not occur between putative species (Rannala & Yang, 2003), yet gene flow and hybridization are pervasive in nature (Taylor & Larson, 2019) and can impact topology and parameter estimation during species tree inference (Leaché, Harris, Rannala, & Yang, 2014; Luo, Ling, Ho, & Zhu, 2018). Fortunately, recent methodological advances explicitly model and incorporate gene flow into species delimitation pipelines (Jackson, Carstens, Morales, & O'Meara, 2016), but such methods have not been widely adopted yet. Researchers also disagree on whether MSC methods delimit lineages that have truly undergone speciation or merely exhibit population structure, which depends in part on which speciation model is adopted (extended vs. instantaneous speciation; Leaché, Zhu, Rannala, & Yang, 2019; Sukumaran & Knowles, 2017). Furthermore, the number and type of loci used may also affect species delimitation pipelines (Leaché, McElroy, & Trinh, 2018). For example, including more loci at the expense of fewer individuals leads to better resolved species trees and higher confidence metrics in species delimitation scenarios compared to restricting analyses to loci that have no missing data (Gottscho et al., 2017; O'Connell & Smith, 2018).

Another assumption implicit in spatial studies of genetic variation is that geographic sampling is comprehensive, such that most or all populations, biogeographic regions, and contact zones are represented. However, exhaustive geographic sampling is difficult—if not impossible—in many systems. Species distributions may be extremely large or undefined, span multiple countries, include remote or politically unstable regions, or present other logistical difficulties. The ramifications of inadequate geographic (and taxonomic) sampling have been considered in the context of DNA barcoding (Dupuis, Roe, & Sperling, 2012; Hebert, Cywinska, Ball, & deWaard, 2003; Moritz & Cicero, 2004), in which the omission of admixed or intermediate populations generally leads to an overestimation of species richness. Inadequate or uneven geographic sampling is particularly problematic in the face of isolation by distance, a non-random mating process by which geographically proximate individuals are more closely related to each other than more distant individuals due to limited dispersal (Rousset, 1997; Wright, 1943). Indeed, various studies have documented overclustering, or an inflated number of inferred populations, in the presence of isolation by distance (Frantz, Cellina, Krier, Schley, & Burke, 2009; Perez et al., 2018).

The question of whether individuals represent 'clusters' or 'clines' amid uneven sampling has been explored in the broader context of population genetics (Bradburd, Coop, & Ralph, 2018; Rosenberg et al., 2005). Sampling regimes and program settings can have a large impact on the number of population clusters inferred (Janes et al., 2017; Wang, 2017), although this problem is somewhat mitigated by including large panels of loci (Rosenberg et al., 2005). Moreover, new methods have been developed that simultaneously estimate population clusters and clines by examining signals of isolation by distance alongside discrete population structure (Bradburd et al., 2018). While these studies represent substantial progress in our understanding of the impacts that sampling regimes and isolation by distance have on inferring population clusters and clines, the effects of isolation by distance and geographic sampling have received less attention in the context of coalescent-based species delimitation (but see Barley, Brown, & Thomson, 2018). Understanding how modern, widely used methods behave amid violations of their underlying assumptions is important, given the widespread implications of species delimitation for basic and applied science (Frankham et al., 2012; Hedin, 2015; Isaac, 2004).

In this study, we considered the impacts of geographic sampling on species delimitation under the multispecies coalescent model. We focused on a widely used delimitation pipeline that uses the program SNAPP (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) to compare species delimitation scenarios with Bayes factors (BFD*; Leaché, Fujita, et al. 2014). We chose this particular pipeline because it leverages single nucleotide polymorphisms (SNPs), which are increasingly used in species delimitation studies as high-throughput sequencing pipelines become more affordable and accessible (Edwards, Shultz, & Campbell-Staton, 2015; Leaché & Oaks, 2017). We first examined the effects of isolation by distance and geographic sampling by simulated data sets with different, yet biologically realistic, levels of gene flow that represent various degrees of isolation by distance across a hypothetical landscape. Incorporating simulated data allowed us to explore a wide array of demographic scenarios beyond what would have been possible with empirical data alone to consider the impact of geographic sampling on heuristic delimitation models in taxa with different dispersal abilities. To compare our findings with the simulated data set to an empirical system, we then examined the effect of geographic sampling and isolation by distance on species delimitation among populations of mayflies in the genus *Andesiops* (Lugo-Ortiz & McCafferty, 1999) collected along an elevational transect of a single river drainage in the Ecuadorian Andes (Polato et al., 2017). The focal taxon (*Andesiops peruvianus*; Ulmer, 1920) is currently classified as single species, but is likely part of a species complex with multiple cryptic species (Polato et al., 2018). For the purposes of our study, we subsample populations from a single drainage as a simplified empirical system to examine the impacts of isolation by distance and geographic sampling 'adequacy'. In light of our findings, we provide general guidelines to empiricists that seek to use coalescent-based species delimitation pipelines in their own taxonomic research.

Materials and methods

Simulated data

We first generated simulated data sets of populations following a stepping stone-model with differing levels of isolation by distance. These simulated data sets are intended to represent single species with variable amounts of population structure across geographic space, ranging from panmixia to highly structured populations. In cases where our simulated data sets are generated with very little gene flow (i.e., $4\,\mathrm{N_e}m = 0.5$ or $4\,\mathrm{N_e}m = 1$), one could argue that the extremes of the transect represent different species, as is invoked in 'ring species' complexes (Alcaide, Scordato, Price, & Irwin, 2014; Irwin, 2005). For the purposes of our analysis and discussion, however, we consider each of these simulated data sets to represent a single species. We simulated data using the program msLandscape (House & Hahn, 2018), which is a wrapper for the program ms (Hudson, 2002) that allows users to generate simulated genetic data sets according to a specified landscape of population connectivity. To simulate a sampling regime comparable to a linear, simplified sampling transect (such as a montane river), we generated a data set with nine populations designated as 'columns' situated across one 'row' with four individuals per population. Each simulated data set contained 100 polymorphic SNPs. We altered the strength of gene flow ($4\,\mathrm{N_e}m$) in our msLandcape simulations by changing the '–m' flag in ms to one of six values: 0.5, 1, 5, 10, 50, or 100. These gene flow values were chosen to reflect variation observed among organisms with diverse life histories such as cichlids, fruit flies, chimpanzees, marine fish, and sunflowers (Hey, 2006; Hey & Nielsen, 2004; Kane et al., 2009; Won, Sivasundar, Wang, & Hey, 2005). We also used the flag '-ej' to generate data for an outgroup that is required for SNAPP. The remaining settings were left at default values. We visualized these simulated data using DAPC and Structure with $K = 2$ (Fig. 1). To further contextualize the degree of genetic differentiation for each simulated set of populations, we calculated F_{ST} between the extreme populations on either side of the imaginary transect.

We then established four geographic sampling scenarios to examine the effect of progressively excluding geographically intermediate populations with admixed individuals along the simulated transect (Fig. S1). The first scenario included all nine populations, while the second scenario omitted the centermost locality of the

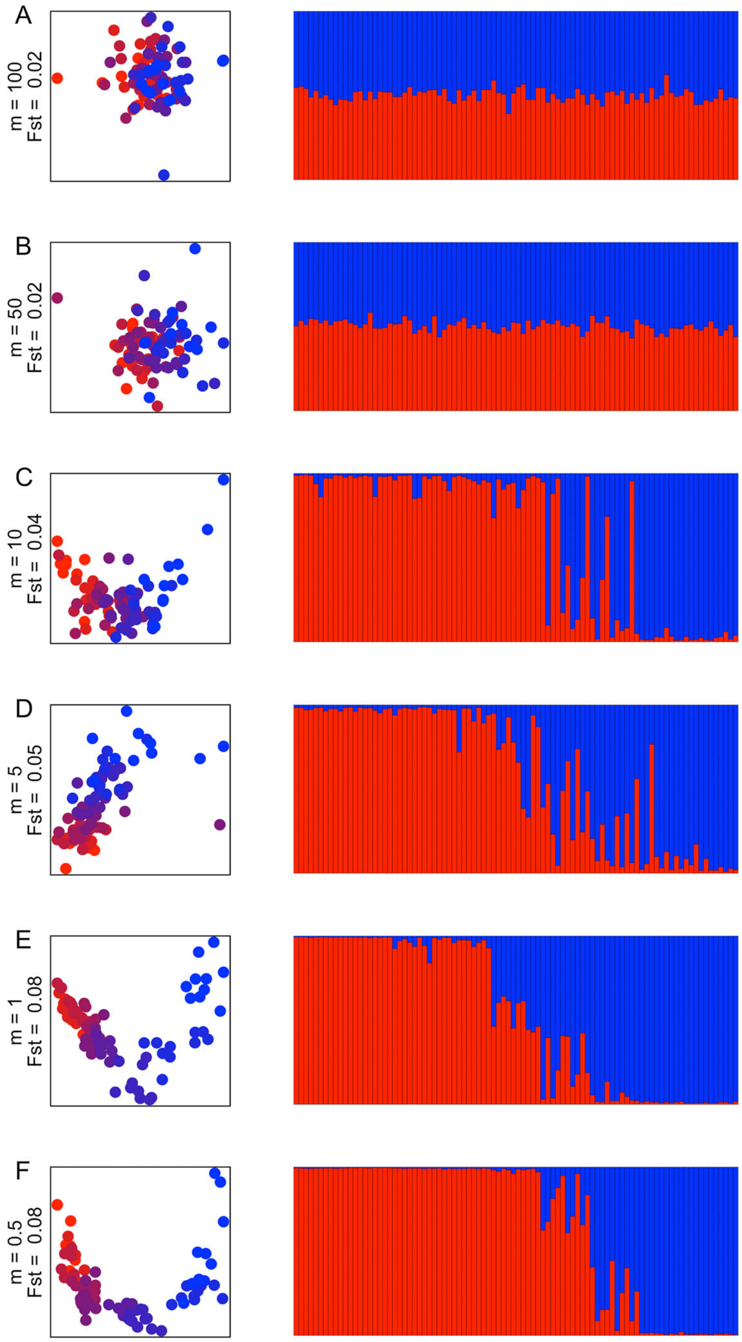

Fig. 1. Simulated data generated in msLandscape. Scatterplots of PC1 (*x* axis) and PC2 (*y* axis) are shown for each level of gene flow. The parameter '*m*' corresponds to the variable used in msLandscape to adjust the level of gene flow among adjacent populations and equals 4 $N_e m$. The simulated landscape included nine populations situated along a straight-line transect that can only exchange genes with adjacent populations. Simulated individuals are colored along a gradient from blue to red according to which side of the simulated landscape they are sampled from. We also show the output from Structure with the number of population clusters $K = 2$.

transect. The third and fourth scenarios each omitted an additional population from either side of the center of the transect. We kept the total number of individuals in each analysis as consistent as possible across hypothetical species delimitation scenarios ($n = 32$ or 36 'ingroup' individuals and 8 outgroup individuals). At the same time, we kept the number of individuals sampled per population equal ($n = 4$, 4, 6, and 9 individuals per populations for geographic sampling scenarios that include for 9, 8, 6, and 4 populations, respectively)

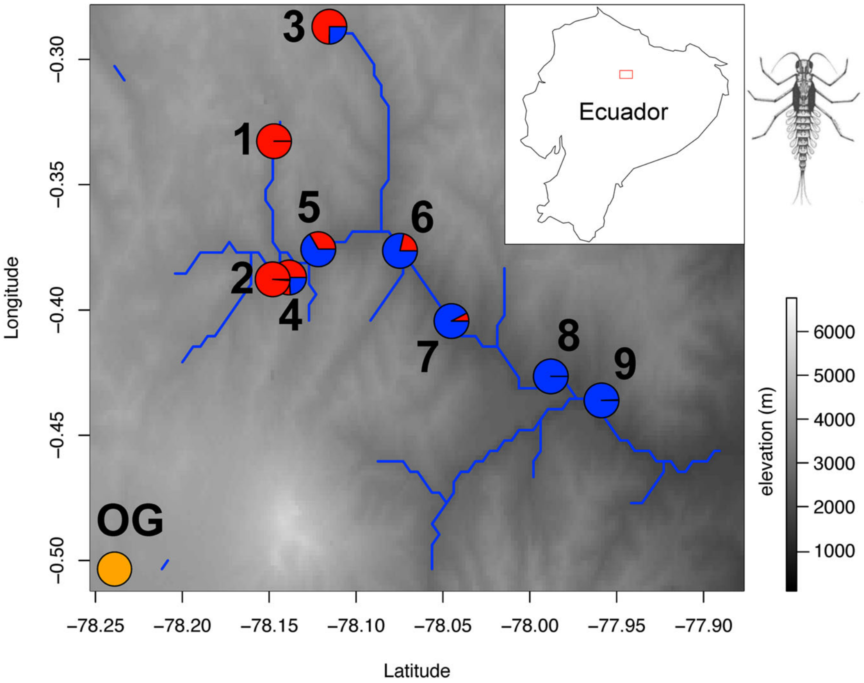

Fig. 2. Empirical data set of *Andesiops* mayflies from Ecuador. Sampling sites are shown, in which the proportion of colors for each pie chart represents the number of individuals assigned to one of two genetic clusters in STRUCTURE. The locality in orange is the outgroup population of *Andesiops* from the adjacent Antisana drainage. Elevation is shown in grayscale and rivers in blue.

so that each analysis had either 36 or 32 ingroup samples with an even number of individuals per sampling location.

Within each population and each geographic sampling scenario, we assigned ingroup individuals to putative species groups for subsequent analyses based on their individual Q scores from STRUCTURE (Pritchard, Stephens, & Donnelly, 2000) with $K = 2$. We then performed Bayes factor species delimitation by calculating the difference in marginal likelihoods between species trees in which the transect is treated either as a single species or as two species (BFD*; Grummer et al., 2014; Leaché, Fujita, et al., 2014) for each geographic sampling scenario. We implemented the 'stepping-stone' algorithm to calculate marginal likelihoods in SNAPP (Bryant et al., 2012), which is available as a plugin for BEAST v2.5.2 (Bouckaert et al., 2014). We fixed the mutation rates u and v to 0.68 and 1.83, respectively, based on posterior estimates from preliminary analyses. We set the prior for effective population sizes (θ) corresponding to each node using a gamma distribution $\theta \sim G(1, 250)$ with a mean of α/β=0.004 following documentation and tutorials provided by the software developers (Bryant et al., 2012). We assigned a gamma hyperprior for the speciation rate parameter lambda $(\lambda) \sim G(2, 200)$ with a mean of $\alpha \times \beta$= 400. We ran the path-sampling analysis for 24 steps with α=0.3, each of which included an MCMC chain with 200,000 burn-in generations followed by 500,000 generations. We repeated each analysis three times with random starting seeds to confirm that marginal likelihood estimates were consistent. A difference in Bayes factors (2 ln(marginal likelihood$_{\text{model 1}}$ − marginal likelihood$_{\text{model 2}}$)) of more than 10 suggests 'decisive' support in favor of one model over another (Kass & Raftery, 1995).

Empirical data

To compare trends in our simulated data set to a real-world, empirical system, we examined the effect of geographic sampling on species delimitation using a data set of SNPs genotyped for *Andesiops* mayflies collected along an elevational transect on the eastern slope of the Ecuadorian Andes (Polato et al., 2018). Our *Andesiops* sampling focused on nine sampling sites (wadeable montane streams) that spanned an approximately northwest-to-southeast transect along the Rio Papallacta drainage of northern Ecuador (Fig. 2). These sites represent the upper and lower elevational limits of the *Andesiops* species complex within this drainage based on DNA barcoding (Gill et al., 2016; Polato et al., 2018). Our working assumption for this study is that we are sampling a single species with population structure along a river drainage. Nonetheless, we acknowledge that mayflies often form species complexes that

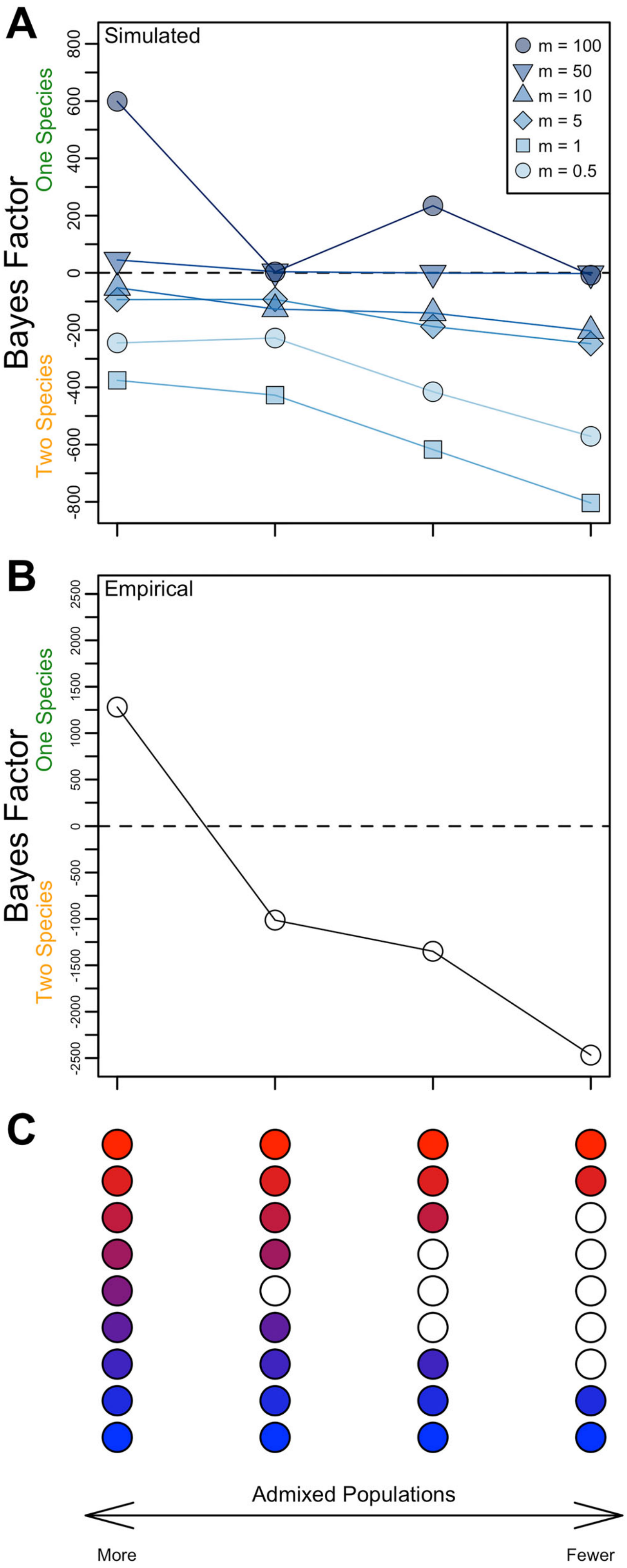

Fig. 3. Effects of geographic sampling on the output of the species delimitation pipeline for both simulated (A) and empirical (B) data for various geographic sampling scenarios (C). Positive Bayes factors indicate support for a single-species delimitation model, while negative Bayes factors support a two-species delimitation model. Output from simulated data sets with different strengths of isolation by distance are displayed with different symbols (A). The parameter 'm' corresponds to the variable used in msLandscape to adjust the level of gene flow among adjacent populations and equals 4 N$_e$*m*.

differentiate along elevational gradients, such that samples from high and low-elevation sites may represent separate evolutionary lineages or species (Gill et al., 2016). SNP genotyping was performed using double digest Restriction-Site Associated DNA sequencing (ddRAD-Seq; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). We demultiplexed reads, performed alignments, and called SNPs in Stacks v 1.19 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) using previously published parameters (Polato et al., 2017). We removed individuals that had more than 50% missing data and retained loci that had at least one individual present for each putative species, as the number of species can influence marginal likelihood values among competing delimitation scenarios (Leaché et al., 2018), and then randomly selected 100 loci that had a minor allele frequency greater than 0.05. Thus, we were unable to restrict our analyses to only SNPs with no missing data (as is common with ddRAD data sets), but we did ensure that each species delimitation scenario had equal numbers of SNPs as the size of the alignment is known to affect marginal likelihood estimates (Leaché et al., 2018).

To generate an 'idealized' scenario of isolation by distance to examine the impacts of geographic sampling in the presence of admixed individuals along a transect, we further filtered our empirical dataset by preferentially selecting genetically 'pure' individuals from the geographic extremes of the river drainage and genetically admixed individuals toward the center (Fig. S1). We first assigned individuals to one of two population clusters and examined patterns of admixture and individual assignment to populations using STRUCTURE with $K = 2$ and all nine populations. For the populations on either geographic extreme of the transect (populations 1, 2, 8, and 9 in Fig. 2), we preferentially selected individuals that had a high assignment probability to their respective end of the drainage transect. For populations toward the center of the transect (populations 3, 4, 5, 6, and 7 in Fig. 2), we selected individuals with Q scores closest to 0.5, which have higher levels of admixture.

We then incrementally omitted populations from our empirical data in the same manner as in our simulated datasets to assess the impact of geographic sampling: we compared the output of SNAPP species delimitation with Bayes factors among different sampling scenarios by incrementally removing one or more populations from the center of the transect and rerunning delimitation analyses (Fig. 3C). We used all the same settings as our simulated data sets for the stepping-stone analysis in SNAPP and BEAST to calculate marginal likelihoods for each level of gene flow and each geographic sampling scenario. We ran each stepping stone analysis three times to account for potential stochasticity in our Bayesian analyses and assessed consistency among our replicate runs.

Results

Bayes factors values varied among geographic sampling scenarios for both the empirical and simulated data sets. For the simulated data, we found that the amount of gene flow or isolation by distance interacted with the geographic sampling scenarios to impact species delimitation inferences (Table S1). When gene flow was low (i.e., $4\,N_em < 10$) a two-species delimitation scenario was consistently favored across all geographic sampling scenarios (Fig. 3A). In contrast, when gene flow was intermediate ($10 < 4\,N_em < 50$) or very high ($4\,N_em > 50$) geographic sampling impacts whether one or two species is inferred (Fig. 3A). Specifically, when we simulated landscapes with $4\,N_em = 50$ and $4\,N_em = 100$, a single species was "strongly favored" in both scenarios when all sampling sites were included (Bayes factor of 44.98 and 598.98 for $4\,N_em = 50$ and $4\,N_em = 100$, respectively; (Kass & Raftery, 1995). However, this switched to 'equivocal' or 'substantial' support in favor of two species when sampling sites with admixed individuals were omitted (Bayes factor of -2.52 and -8.12 for $4\ N_em = 50$ and $4\ N_em = 100$, respectively; Kass & Raftery, 1995). For the empirical data (Fig. 3B), we found that Bayes factors favored a single-species delimitation scenario under the geographic sampling scenario with all populations included (2LnBF = 1280.4), but favored two species under geographic sampling scenarios with one population removed (2LnBF$= -1013.1$), three populations removed (2LnBF$= -1348.8$), and five populations removed (2LnBF$= -2468.2$). Thus, inferences from this species delimitation pipeline are contingent on an interaction between the strength of isolation by distance and the completeness of geographic sampling among populations along a sampling transect.

Discussion

Geographic sampling and isolation by distance impacted species delimitation inferences in both the empirical and simulated data sets considered in this study. When we omitted populations with genetically admixed individuals from geographically central locations, support for two-species delimitation scenarios consistently increased for both simulated (Fig. 3A) and empirical (Fig. 3B) datasets. This result parallels theoretical and empirical studies on sampling design in geographic clines and hybrid zones and along ecological and spatial gradients in population genetics, in which continuous distributions appear bimodal when only the extremes are sampled (Barton, 1985; Mullen & Hoekstra, 2008; Nagylaki, 1976; Slatkin, 1973). Furthermore, when gene flow was restricted and isolation by distance was strong (i.e., $4\,N_e m \leq 10$), two species delimitation scenarios were favored even when all sites along the transect were included. With these findings in mind, we recommend that careful consideration must be given to the adequacy of geographic sampling and the impact of isolation by distance in coalescent-based species delimitation pipelines. Similar advice has been offered elsewhere (Carstens et al., 2013; Moritz & Cicero, 2004), but by quantifying the impact of missing populations in empirical and simulated data sets, our study highlights the sensitivity of coalescent-based species delimitation to geographic sampling and isolation by distance. Specifically, we find that omitting even a single geographically, genetically intermediate population can spuriously increase support for multiple species (Fig. 3). Thus, researchers should be aware for the propensity of SNAPP and related methods to potentially oversplit taxa when population genetic structure, such as isolation by distance, or large gaps in geographic sampling are present. It is noteworthy that in our study, simulated F_{ST} values between the geographically most distant populations were consistently low (<0.08), yet were still recognized as multiple species when intermediate sites were removed. While we have focused in this study on SNAPP, our findings likely translate to other coalescent-based methods. In fact, Barley et al. (2018) recently demonstrated how other multispecies coalescent methods, specifically BPP (Yang, 2015) and STACEY (Jones, 2017), are also sensitive to demographic violations of their underlying models, suggesting that the patterns observed in our study likely impact other methods beyond SNAPP.

Isolation by distance is prominent in nature and present to some extent in essentially all taxa (Meirmans, 2012; Slatkin, 1993). However, the strength of isolation by distance varies dramatically among organisms with different reproductive modes, life histories, and dispersal capacities (Bohonak, 1999; Bradbury, Laurel, Snelgrove, Bentzen, & Campana, 2008; Kinlan & Gaines, 2003). Our simulations suggest that when isolation by distance is strong and gene flow between neighboring populations is moderate or low ($4\,N_e m \leq 10$), SNAPP and BFD* consistently support the presence of multiple species over a single species (Fig. 3B). Yet even when gene flow is moderate or high ($4\,N_e m > 10$) and isolation by distance is correspondingly low, geographic sampling impacts the favored species delimitation scenario. SNAPP and BFD* have been frequently applied in tetrapod systems, many of which have levels of dispersal and gene flow among populations that fall within the range of migration rates that we considered in this study, including birds (Mason, Olvera-Vital, Lovette, & Navarro-Sigüenza, 2018; Mason & Taylor, 2015; Oswald et al., 2016), lizards (MacGuigan, Geneva, & Glor, 2017; Potter, Bragg, Peter, Bi, & Moritz, 2016), and frogs (French, Deutsch, Chávez, Almora, & Brown, 2019). More generally, empiricists should be aware of the impacts of IBD on BFD* and its proclivity to support taxonomic splits when geographic sampling is sparse across continuous or ambiguous distributions, as seen in river sharks (Li et al., 2015), finless porpoises (Zhou et al., 2018), and crocodile skinks (Rittmeyer & Austin, 2015). Empiricists should quantify the prevalence of IBD in their data sets in conjunction with other data on dispersal, especially in systems in which limited dispersal is suspected or that span large geographic distributions. In extreme cases, support for multispecies delimitation scenarios may simply reflect patterns of population structure, geographic sampling, or both rather than independently evolving lineages. Some may feel that oversplitting species may not impose a large problem for systematics or conservation initiatives when compared to the impacts of rampant habitat destruction and anthropogenic extinction events. However, spurious species splits may lead to misallocation of limited resources that could impose serious problems for mitigating species-level biodiversity loss (Isaac, 2004; Pillon & Chase, 2007). This problem is exacerbated in cryptic species complexes, in which observable phenotypic variation is minimal and cannot be used as an alternative line of evidence to evaluate the output of species delimitation pipelines and the potential for reproductive isolation.

Secondary contact between divergent lineages is common in nature and can produce admixed individuals through a different biological process than isolation by distance (Harrison, 1998). The origin of the admixed *Andesiops* populations we study here is unknown, but the simulated data that we generated followed a model of isolation by distance rather than secondary contact. Thus, we did not directly address the impacts of secondary contact in this study, but others have found that gene flow between previously isolated populations can impact estimates of population size and divergence times in coalescent models (Leaché, Harris, et al., 2014). Recent gene flow following protracted isolation generates large blocks of linked loci in first-generation hybrids and subsequent backcrosses, and selection in natural systems may restrict gene flow and introgression to certain genomic regions (Wu, 2001), thereby producing substantially different patterns of admixture across the genome than we have considered here. Future studies could expand on our work to assess the impact of gene flow via secondary contact on species delimitation following periods of reproductive isolation.

Exhaustive geographic sampling is exceedingly difficult for many systems with large, continuous geographic distributions. Empirical studies must therefore strike a balance between sampling adequacy and feasibility. As seen in our empirical data set on *Andesiops* mayflies, the inclusion or exclusion of even a single site containing genetically intermediate individuals can have a large effect on the inferred level of support for multispecies or a single-species delimitation scenarios. Studies based on inadequate intraspecific sampling may overestimate species richness if admixed or intermediate individuals are not included, especially when populations are distributed continuously across an expansive landscape. Biases imposed by inadequate geographic sampling may be diminished when systems have clear *a priori* designations of individuals based on subspecific taxonomy or phenotypic variation or when populations have allopatric distributions. Clearly, integrating these additional indicators of divergence or incipient speciation to guide collecting in admixture zones will reduce the chance of missing contact zones altogether.

Determining whether geographic sampling is adequate for applying species delimitation methods is not trivial: one must consider the geographic context of a species' distribution in light of what is known about dispersal ability and gene flow. Our study does not provide explicit advice on how many individuals or sites are necessary to have confidence in the output of coalescent species delimitation; this would be a fruitful avenue for future research in biodiversity and systematics. Nonetheless, our findings send a clear message to empirical systematists to use caution in applying coalescent-based species delimitation methods when geographic sampling is sparse or incomplete, and to focus efforts on sampling areas that are central in genetic clines created by isolation by distance. As others have suggested (e.g., Fujita et al., 2012), integrative taxonomy that incorporates multiple lines of evidence on reproductive isolation, morphological and ecological differentiation, and patterns of coalescent should be heavily favored over coalescent-based species delimitation alone.

Data accessibility

All simulated and empirical data sets generated in this study as well as scripts used to analyses are available via the Dryad Data Repository: https://doi.org/10.6078/D1Z41Q.

Acknowledgments

We thank L. Knowles and D. Hillis for organizing a symposium on species delimitation at the Portland 2017 Evolution meetings that inspired our study. We also thank M. Hahn, T. Reeder, A. Leache, and S. Harrington for useful feedback and input on earlier versions of our study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Supplemental data

Supplemental data for this article can be accessed here: https://doi.org/10.1080/14772000.2020.1730475.

Funding

This research was supported by National Science Foundation (NSF) Dimensions of Biodiversity Grants DEB-1046408 and DEB-1045960. N. Mason was supported by a NSF Postdoctoral Fellowship in Biology (DBI-1710739).

References

Alcaide, M., Scordato, E. S. C., Price, T. D., & Irwin, D. E. (2014). Genomic divergence in a ring species complex. *Nature*, *511*, 83–85. doi:10.1038/nature13285

Barley, A. J., Brown, J. M., & Thomson, R. C. (2018). Impact of model violations on the inference of species boundaries under the multispecies coalescent. *Systematic Biology*, *67*, 269–284. doi:10.1093/sysbio/syx073

Barrowclough, G. F., Cracraft, J., Klicka, J., & Zink, R. M. (2016). How many kinds of birds are there and why does it matter? *Public Library of Science One*, *11*, e0166307. doi: 10.1371/journal.pone.0166307

Barton, N. H. (1985). Analysis of hybrid zones. *Genetics*, *75*, 733–756.

Bauer, A. M., Parham, J. F., Brown, R. M., Stuart, B. L., Grismer, L., Papenfuss, T. J., … Inger, R. F. (2011). Availability of new Bayesian-delimited gecko names and the importance of character-based species descriptions. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 490–492. doi:10.1098/rspb.2010.1330

Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, *74*, 21–45. doi: 10.1086/392950

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *Public Library of Science Computational Biology*, *10*, e1003537. doi:10.1371/journal.pcbi.1003537

Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring continuous and discrete population genetic structure across space. *Genetics*, *210*, 33–52. doi:10.1534/genetics.118.301333

Bradbury, I. R., Laurel, B., Snelgrove, P. V. R., Bentzen, P., & Campana, S. E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proceedings of the Royal Society B: Biological Sciences*, *275*, 1803–1809. doi:10.1098/rspb.2008.0216

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology & Evolution*, *29*, 1917–1932. doi:10.1093/molbev/mss086

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*, 4369–4383. doi:10.1111/mec.12413

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. doi:10.1111/mec.12354

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*, 879–886. doi:10.1080/10635150701701083

Dupuis, J. R., Roe, A. D., & Sperling, F. A. H. (2012). Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology*, *21*, 4422–4436. doi:10.1111/j.1365-294X.2012.05642.x

Edwards, S. V., Shultz, A. J., & Campbell-Staton, S. C. (2015). Next-generation sequencing and the expanding domain of phylogeography. *Folia Zoologica*, *64*, 187–206. doi:10.25225/fozo.v64.i3.a2.2015

Ence, D. D., & Carstens, B. C. (2011). SpedeSTEM: a rapid and accurate method for species delimitation. *Molecular Ecology Resources*, *11*, 473–480. doi:10.1111/j.1755-0998.2010.02947.x

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, *143*, 1919–1927. doi:10.1016/j.biocon.2010.05.011

Frankham, R., Ballou, J. D., Dudash, M. R., Eldridge, M. D. B., Fenster, C. B., Lacy, R. C., … Ryder, O. A. (2012). Implications of different species concepts for conserving biodiversity. *Biological Conservation*, *153*, 25–31. doi:10.1016/j.biocon.2012.04.034

Frantz, A. C., Cellina, S., Krier, A., Schley, L., & Burke, T. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, *46*, 493–505. doi:10.1111/j.1365-2664.2008.01606.x

French, C. M., Deutsch, M. S., Chávez, G., Almora, C. E., & Brown, J. L. (2019). Speciation with introgression: phylogeography and systematics of the *Ameerega petersi* group (Dendrobatidae). *Molecular Phylogenetics & Evolution*, *138*, 31–42. doi:10.1016/j.ympev.2019.05.021

Fujita, M. K., & Leaché, A. D. (2011). A coalescent perspective on delimiting and naming species: a reply to Bauer et al. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 493–495. doi:10.1098/rspb.2010.1864

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*, 480–488. doi:10.1016/j.tree.2012.04.012

Gill, B. A., Kondratieff, B. C., Casner, K. L., Encalada, A. C., Flecker, A. S., Gannon, D. G., … Funk, W. C. (2016). Cryptic species diversity reveals biogeographic support for the 'mountain passes are higher in the tropics' hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20160553. doi:10.1098/rspb.2016.0553

Gill, F. (2014). Species taxonomy of birds: which null hypothesis? *The Auk*, *131*, 150–161. doi:10.1642/AUK-13-206.1

Gottscho, A. D., Wood, D. A., Vandergast, A. G., Lemos-Espinal, J., Gatesy, J., & Reeder, T. W. (2017). Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: Delimiting species in the presence of gene flow. *Molecular Phylogenetics & Evolution*, *106*, 103–117. doi:10.1016/j.ympev.2016.09.008

Grummer, J. A., Bryson, R. W., & Reeder, T. W. (2014). Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology*, *63*, 119–133. doi:10.1093/sysbio/syt069

Harrison, R. G. (1998). Linking Evolutionary Pattern and Process. In S. Berlocher & D. Howard (Eds.), *Endless forms: Species and speciation* (pp. 19–31). New York: Oxford University Press.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*, 313–321. doi:10.1098/rspb.2002.2218

Hedin, M. (2015). High-stakes species delimitation in eyeless cave spiders (*Cicurina,* Dictynidae, Araneae) from central Texas. *Molecular Ecology*, *24*, 346–361. doi:10.1111/mec.13036

Heller, R., Frandsen, P., Lorenzen, E. D., & Siegismund, H. R. (2013). Are there really twice as many bovid species as we thought? *Systematic Biology*, *62*, 490–493. doi:10.1093/sysbio/syt004

Hey, J. (2006). Recent advances in assessing gene flow between diverging populations and species. *Current Opinion in Genetics & Development*, *16*, 592–596. doi:10.1016/j.gde.2006.10.005

Hey, J., & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, *167*, 747–760. doi:10.1534/genetics.103.024182

House, G. L., & Hahn, M. W. (2018). Evaluating methods to visualize patterns of genetic differentiation on a landscape. *Molecular Ecology Resources*, *18*, 448–460. doi:10.1111/1755-0998.12747

Hudson, R. R. (2002). Generating samples under a Wright–Fisher neutral model of genetic variation. *Bioinformatics*, *18*, 337–338. doi:10.1093/bioinformatics/18.2.337

Irwin, D. E. (2005). Speciation by distance in a ring species. *Science*, *307*, 414–416. doi:10.1126/science.1105201

Isaac, N. (2004). Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution*, *19*, 464–469. doi:10.1016/j.tree.2004.06.004

Jackson, N. D., Carstens, B. C., Morales, A. E., & O'Meara, B. C. (2016). Species delimitation with gene flow. *Systematic Biology*, *66*, 799–812. doi:10.1093/sysbio/syw117

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*, 3594–3602. doi:10.1111/mec.14187

Jones, G. (2017). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, *74*, 447–467. doi:10.1007/s00285-016-1034-0

Kane, N. C., King, M. G., Barker, M. S., Raduski, A., Karrenberg, S., Yatabe, Y., … Rieseberg, L. H. (2009). Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. *Evolution*, *63*, 2061–2075. doi:10.1111/j.1558-5646.2009.00703.x

Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. doi:10.1080/01621459.1995.10476572

Kinlan, B. P., & Gaines, S. D. (2003). Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*, *84*, 2007–2020. doi:10.1890/01-0622

Knowles, L. L., & Carstens, B. C. (2007). Delimiting species without monophyletic gene trees. *Systematic Biology*, *56*, 887–895. doi:10.1080/10635150701701091

Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, *63*, 534–542. doi:10.1093/sysbio/syu018

Leaché, A. D., Harris, R. B., Rannala, B., & Yang, Z. (2014). The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, *63*, 17–30. doi:10.1093/sysbio/syt049

Leaché, A. D., McElroy, M. T., & Trinh, A. (2018). A genomic evaluation of taxonomic trends through time in coast horned lizards (genus *Phrynosoma*). *Molecular Ecology*, *27*, 2884–2895. doi:10.1111/mec.14715

Leaché, A. D., & Oaks, J. R. (2017). The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annual Review of Ecology, Evolution, & Systematics*, *48*, 69–84. doi:10.1146/annurev-ecolsys-110316-022645

Leaché, A. D., Zhu, T., Rannala, B., & Yang, Z. (2019). The spectre of too many species. *Systematic Biology*, *68*, 168–181. doi:10.1093/sysbio/syy051

Li, C., Corrigan, S., Yang, L., Straube, N., Harris, M., Hofreiter, M., … Naylor, G. J. P. (2015). DNA capture reveals transoceanic gene flow in endangered river sharks. *Proceedings of the National Academy of Sciences*, *112*, 13302–13307. doi:10.1073/pnas.1508735112

Lugo-Ortiz, C. R., & McCafferty, W. P. (1999). Three new genera of small minnow mayflies (Insecta: Ephemeroptera: Baetidae) from the Andes and Patagonia. *Studies on Neotropical Fauna & Environment*, *34*, 88–104. doi:10.1076/snfe.34.2.88.2102

Luo, A., Ling, C., Ho, S. Y. W., & Zhu, C. D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, *67*, 830–846. doi:10.1093/sysbio/syy011

MacGuigan, D. J., Geneva, A. J., & Glor, R. E. (2017). A genomic assessment of species boundaries and hybridization in a group of highly polymorphic anoles (*distichus* species complex). *Ecology & Evolution*, *7*, 3657–3671. doi:10.1002/ece3.2751

Mason, N. A., Olvera-Vital, A., Lovette, I. J., & Navarro-Sigüenza, A. G. (2018). Hidden endemism, deep polyphyly, and repeated dispersal across the Isthmus of Tehuantepec: diversification of the White-collared Seedeater complex (Thraupidae: *Sporophila torqueola*). *Ecology & Evolution*, *8*, 1867–1881. doi:10.1002/ece3.3799

Mason, N. A., & Taylor, S. A. (2015). Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. *Molecular Ecology*, *24*, 3009–3025. doi:10.1111/mec.13140

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, *21*, 2839–2846. doi:10.1111/j.1365-294X.2012.05578.x

Moritz, C., & Cicero, C. (2004). DNA barcoding: promise and pitfalls. *Public Library of Science Biology*, *2*, e354. doi:10.1371/journal.pbio.0020354

Mullen, L. M., & Hoekstra, H. E. (2008). Natural selection along an environmental gradient: A classic cline in mouse pigmentation. *Evolution*, *62*, 1555–1570. doi:10.1111/j.1558-5646.2008.00425.x

Nagylaki, T. (1976). Clines with variable migration. *Genetics*, *83*, 867–866.

O'Connell, K. A., & Smith, E. N. (2018). The effect of missing data on coalescent species delimitation and a taxonomic revision of whipsnakes (Colubridae: *Masticophis*). *Molecular Phylogenetics & Evolution*, *127*, 356–366. doi:10.1016/j.ympev.2018.03.018

Oswald, J. A., Harvey, M. G., Remsen, R. C., Foxworth, D. U., Cardiff, S. W., Dittmann, D. L., … Brumfield, R. T. (2016). Willet be one species or two? A genomic view of the evolutionary history of *Tringa semipalmata*. *The Auk*, *133*, 593–614. doi:10.1642/AUK-15-232.1

Perez, M. F., Franco, F. F., Bombonato, J. R., Bonatelli, I. A. S., Khan, G., Romeiro-Brito, M., … Moraes, E. M. (2018). Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Diversity & Distributions*, *24*, 1883–1889. doi:10.1111/ddi.12816

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *Public Library of Science One*, *7*, e37135. doi:10.1371/journal.pone.0037135

Pillon, Y., & Chase, M. W. (2007). Taxonomic exaggeration and its effects on orchid conservation. *Conservation Biology*, *21*, 263–265. doi:10.1111/j.1523-1739.2006.00573.x

Polato, N. R., Gill, B. A., Shah, A. A., Gray, M. M., Casner, K. L., Barthelet, A., … Zamudio, K. R. (2018). Narrow thermal tolerance and low dispersal drive higher speciation in tropical mountains. *Proceedings of the National Academy of Sciences*, *115*, 12471–12476. doi:10.1073/pnas.1809326115

Polato, N. R., Gray, M. M., Gill, B. A., Becker, C. G., Casner, K. L., Flecker, A. S., … Zamudio, K. R. (2017). Genetic diversity and gene flow decline with elevation in montane mayflies. *Heredity*, *119*, 107–116. doi:10.1038/hdy.2017.23

Potter, S., Bragg, J. G., Peter, B. M., Bi, K., & Moritz, C. (2016). Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. *Molecular Ecology*, *25*, 1367–1380. doi:10.1111/mec.13546

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, *164*, 1645–1656.

Rittmeyer, E. N., & Austin, C. C. (2015). Combined next-generation sequencing and morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: *Tribolonotus*). *Molecular Ecology*, *24*, 466–483. doi:10.1111/mec.13030

Rosenberg, N. A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J. K., & Feldman, M. W. (2005). Clines, clusters, and the effect of study design on the inference of human population structure. *Public Library of Science Genetics*, *1*, e70. doi:10.1371/journal.pgen.0010070

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*, 1219–1228.

Slatkin, M. (1973). Gene flow and selection in a cline. *Genetics*, *75*, 733–756.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*, 264–279. doi:10.1111/j.1558-5646.1993.tb01215.x

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*, 1607–1612. doi:10.1073/pnas.1607921114

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, *3*, 170. doi:10.1038/s41559-018-0777-y

Toews, D. P. L. (2014). Biological species and taxonomic species: will a new null hypothesis help? (A comment on Gill 2014). *The Auk.*, *132*, 78–81. doi:10.1642/AUK-14-138.1

Ulmer, G. (1920). Neue ephemeropteren. *Archiv Für Naturgeschichte*, *85*, 1–80.

Wang, J. (2017). The computer program STRUCTURE for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources*, *17*, 981–990. doi:10.1111/1755-0998.12650

Won, Y.-J., Sivasundar, A., Wang, Y., & Hey, J. (2005). On the origin of Lake Malawi cichlid species: a population genetic analysis of divergence. *Proceedings of the National Academy of Sciences*, *102*, 6581–6586. doi:10.1073/pnas.0502127102

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

Wu, C.-I. (2001). The genic view of the process of speciation: Genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*, 851–865. doi:10.1046/j.1420-9101.2001.00335.x

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*, 854–865. doi:10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *107*, 9264–9269. doi:10.1073/pnas.0913022107

Zhou, X., Guang, X., Sun, D., Xu, S., Li, M., Seim, I., … Yang, G. (2018). Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. *Nature Communications*, *9*, 1276. doi:10.1038/s41467-018-03722-x

Associate Editor: Dimitar Dimitrov