A multi-tier species delimitation approach resolves conflicts in delineating the primitively segmented spider genus *Heptathela* endemic to Japanese islands

Xin Xu[1,2*+], Matjaž Kuntner [2,3,4,5+], Jason E. Bond[6], Hirotsugu Ono[7], Fengxiang Liu[2], Long Yu[2], Daiqin Li[8*]

[1]*College of Life Sciences, Hunan Normal University, Changsha, Hunan, China*
[2]*State Key Laboratory of Biocatalysis and Enzyme Engineering, Centre for Behavioural Ecology and Evolution, School of Life Sciences, Hubei University, Wuhan, Hubei, China*
[3]*Evolutionary Zoology Laboratory, Department of Organisms and Ecosystems Research, National Institute of Biology, Ljubljana, Slovenia*
[4]*Evolutionary Zoology Laboratory, Biological Institute ZRC SAZU, Ljubljana, Slovenia*
[5]*Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, D.C, USA*
[6]*Department of Entomology and Nematology, University of California at Davis, Davis, California, USA*
[7]*Department of Zoology, National Museum of Nature and Science, 4-1-1 Amakubo, Tsukuba-shi, Ibaraki-ken, 305-0005, Japan*
[8]*Department of Biological Sciences, National University of Singapore, Singapore*

+ Xin Xu and Matjaž Kuntner should be considered joint first authors
* Corresponding authors: xuxin_09@163.com; dbslidq@nus.edu.sg

ABSTRACT

Accurate species delimitation impacts all biological inferences, but taxonomic studies often face conflicts within/among datasets or among species delimitation methods. Here we use a multi-tier analytical strategy to handle conflicts in species delimitation on a group of primitively segmented spiders, genus *Heptathela* endemic to Japanese islands. Tier 1 analysis uses a suite of quick species delimitation methods to test the initial species hypothesis (ISH) and to identify fully congruent lineages. Testing ISH of 19 species based on morphological taxonomic evidence presented elsewhere, tier 1 analysis subjects a molecular dataset of 180 original *Heptathela* samples to distance- and tree-based species delimitation methods, and recovers 16 fully congruent species plus 3, 4 or 6 conflicting lineages. Given these conflicting lineages, tier 2 analysis tests, only within these lineages, alternative species hypotheses (ASH) via multi-locus, coalescent-based species delimitation methods on enriched dataset. We add additional molecular markers only for 35 samples from conflicting lineages, then evaluate three ASH using coalescent-based species delimitation methods (BP&P and BFD). While BP&P lacks delimitation power, BFD best supports 6-species without rejecting the 4-species hypothesis. Because incongruence persists, tier 3 analysis then uses additional operational criteria for identifying diagnosable lineages as valid species. Reevaluating morphological and DNA evidence for 4- vs. 6-species hypothesis, it reveals a DNA barcoding gap supporting 4-species. Our multi-tier approach to resolving conflicts in species delimitation by enriching data selectively for conflicting lineages is a fast and efficient strategy to delimit species in an integrative taxonomic framework.

Keywords: Integrative taxonomy, DNA barcodes, Liphistiidae, *Heptathela*, multi-tier species delimitation.

1. Introduction

Accurate species delimitation is a core assumption in biology that affects nearly every biological subdiscipline (Agnarsson and Kuntner, 2007; Camargo and Sites, 2013; Hedin, 2015; Kress et al., 2015; Sites and Marshall, 2003; Wheeler et al., 2004). However, the accuracy of species delimitation is often impeded by divergent species concepts (De Queiroz, 2007; Freudenstein et al., 2017) and varied practices, examples being models and assumptions in computational analyses (Carstens et al., 2013; Schlick-Steiner et al., 2014, 2010), as well as differing views on the reality and importance of cryptic species (Fišer et al., 2018; Heethoff et al., 2018; Struck et al., 2018). Sequence-based species delimitation methods may overcome some of these problems under an integrative taxonomic framework (Blair and Bryson, 2017; Cardoso et al., 2009; Carstens et al., 2013; Dayrat, 2005; Eberle et al., 2019; Karanovic et al., 2016; Leaché et al., 2014; Pante et al., 2015; Rannala, 2015; Satler et al., 2013). DNA species delimitation is a particularly welcome arbiter in the study of taxa that are well differentiated genetically but not morphologically (Bickford et al., 2007; Bond et al., 2001; Derkarabetian and Hedin, 2014; Leavitt et al., 2015; Xu et al., 2017, 2015), or in those that are, conversely, well differentiated morphologically but not genetically, as is sometime the case in adaptive radiations (Moyle et al., 2009; Wagner et al., 2012).

Species-level taxonomy has historically been often labeled to be a subjective biological discipline (but, see Agnarsson and Kuntner, 2007). This allegation may be due to the assumption that taxonomic species delimitation in many cases relies on one classical set of characters (usually morphology), and does not consider alternative species hypotheses (but, see Schlick-Steiner et al., 2010). This is an acute problem considering the commonly encountered conflict in evidence between morphological species and those delimited based on molecular data (Eberle et al., 2016; Ortiz and Francke, 2016; Vitecek et al., 2017). Even when alternative species hypotheses are considered from differing data sources or from different delimitation methods on the same data set, the resolution of conflicts is often haphazard and effectively devoid of stringent hypothesis testing (Derkarabetian and Hedin, 2014; C. A. Hamilton et al., 2014).

Numerous studies have not arrived at a general consensus how to handle conflicts, particularly when adding dense samples or more molecular markers or genomic data for all the taxa/specimens is not feasible (e.g., Abdelkrim et al., 2018; Jacobs et al., 2018; Satler et al., 2013). After detecting conflicts among single-locus delimitation methods, increasing sample size by adding as many taxa/individuals per taxon and loci as possible is considered as a general "good practice" of mitigating conflicts in delimitation (Abdelkrim et al., 2018; Blaimer et al., 2015; Carstens et al., 2013; Lemmon and Lemmon, 2013; McCormack et al., 2013; Ruane, 2015). However, increasing taxa and individuals may not be possibe for the understudied and/or hyperdiverse taxa (Abdelkrim et al., 2018; Lim et al., 2012; Xu et al., 2017, 2016). Furthermre, increasing loci, for example, genomic scale data, is still time-consuming and expensive, and computationally demanding (Flouri et al., 2018; Leaché et al., 2018; Noguerales et al., 2018). For these reasons, empirical studies, especially those dealing with taxa with no clear apriori taxonomic hypotheses, often start with DNA barcoding analyses. When such studies face the difficulty of conflicted species delimitation results among methods, rarely do they handle conflicts by increasing taxon and data coverage for the conflicted lineages.

In this paper, we focus on an empirical approach towards resolving such conflicts via a multi-tier analytical strategy using combinations of species delimitation schemes. First, we briefly overview recent methodologies in sequence-based species delimitation, a rapidly evolving field, then propose a multi-tier analytical strategy to handle conflict in species delimitation schemes. We then outline how the results from multiple data sources and multi-

tier analytical approaches can be fast and effectively evaluated against the background hypotheses based on classical sources. We use our multi-tier analyses to delimit species boundaries in a group of primitively segmented spiders, *Heptathela* (Mesothelae: Liphistiidae) endemic to Japanese islands that are known for their low vagility and inclination towards population structuring at relatively small to moderate geographic scales.

DNA species delimitation methods are numerous and can be divided into single- versus multi-locus techniques, or into genetic distance- versus tree-based techniques (Carstens et al., 2013; Goldstein and DeSalle, 2011). DNA barcode gap, the most commonly used single-locus, distance-based technique (Barrett and Hebert, 2005; Čandek and Kuntner, 2015; Hamilton et al., 2014; Hebert et al., 2004, 2003; Hebert and Barrett, 2005; Toussaint et al., 2015; Xu et al., 2017, 2015), uses threshold or cut-off values to differentiate inter- from intraspecific divergences. Although such a straightforward analysis of COI barcodes is fast and cost-effective, it may suffer from inaccuracy, or may even lead to erroneous results (Barrett and Hebert, 2005; Elias et al., 2009; Hamilton et al., 2014; Hebert et al., 2003; Hebert and Barrett, 2005; Hedin, 2015; Moritz and Cicero, 2004; Prendini, 2005; Rubinoff and Holland, 2005; Song et al., 2008; Spooner, 2009). Furthermore, a clear cut-off DNA barcode gap between intra- and interspecific genetic distances, which are used as a threshold to delimit species boundaries (e.g., Barrett and Hebert, 2005; Hebert et al., 2004; Meyer and Paulay, 2005; Weigand et al., 2011), may be elusive. Instead, substantial overlap is often detected between intra- and interspecific divergences in diverse taxonomic groups (Čandek and Kuntner, 2015; Dang et al., 2016; Hickerson et al., 2006; Kvist, 2016; Meier et al., 2006; Meyer and Paulay, 2005). Because of this challenge, researchers have developed other single-locus, distance-based species delimitation methods, including the automatic barcode gap discovery (ABGD) algorithm (Puillandre et al., 2012). However, as is true for barcoding gap analyses, ABGD ignores the phylogenetic utility of barcodes, relying instead on an automatically identified gap to delimit species (Hamilton et al., 2011; Hebert and Gregory, 2005).

Single-locus, tree-based species delimitation methods, in contrast, combine population genetics and phylogenetics to delimit evolutionary lineages. Among single-locus methods are the general mixed Yule-coalescent (GMYC) methodology (Pons et al., 2006), the species delimitation plugin P ID(Liberal) (Masters et al., 2011), the Bayesian Poisson tree processes (bPTP) (Zhang et al., 2013), and the multi-rate Poisson tree processes (mPTP) (Kapli et al., 2017). GMYC uses likelihood to test for species boundaries by detecting the transition point of interspecific versus intraspecific rates of lineage coalescence. There are the single (GMYCs) and multiple (GMYCm) threshold models (Monaghan et al., 2009; Pons et al., 2006). Since GMYC often overestimates species in the primitively segmented spiders (Xu et al., 2017, 2015) and in other taxa (Esselstyn et al., 2012; Hamilton et al., 2014; Miralles and Vences, 2013; Paz and Crawford, 2012; Talavera et al., 2013), we exclude it from our analysis in this study. P ID(Liberal) tests different species boundary hypotheses by enabling the user to a priori assign taxa to putative species groups on a phylogenetic tree (Masters et al., 2011). The bPTP model, an updated version of the original PTP with Bayesian support values, requires a rooted input tree. It assumes two independent exponential distributions (one for speciation and another for coalescence) to model branch lengths (Zhang et al., 2013). bPTP seems simplistic because it may ignore stochastic variation among species that differ in population sizes and demographic histories (Blair and Bryson, 2017). The newly developed mPTP, on the other hand, accounts for among-species stochastic variation by fitting multiple independent exponential distributions to each delimited species (Kapli et al., 2017). Both bPTP and mPTP seem to be consistent methods (Ortiz and Francke, 2016; Song et al., 2018; Zhang et al., 2013), but are sensitive to the accuracy of the input tree, population size, divergence time, or the ratio of population size to divergence time, and to ongoing gene flow, and tend to oversplit species (Luo et al., 2018; Tang et al., 2014).

Alternatively, Bayesian Phylogenetics and Phylogeography (BP&P, Yang, 2015), and Bayes Factor Delimitation (BFD, Grummer et al., 2014) are multi-locus, coalescent-based methods. These methods are often used to resolve deeper species complexes and to test alternative models of species delimitation. BP&P needs a resolved guide tree, and it is reported to be effective in delimiting closely related species using information gleaned from gene tree information (Flouri et al., 2018; Luo et al., 2018; Rannala and Yang, 2013; Yang, 2015; Yang and Rannala, 2017, 2010). However, BP&P may be prone to error when the guide tree is inaccurate (Rannala and Yang, 2013) or when the ratio of population size to divergence time is relatively high (Luo et al., 2018). Furthermore, BP&P tends to detect population splits rather than species divergence, thus potentially introducing taxonomic error (Jackson et al., 2017; Leaché et al., 2019; Sukumaran and Knowles, 2017). BFD that compares different species-tree models does not require a prior guide tree. It estimates the marginal likelihood (MLE), measured as log likelihoods, of each species delimitation model using path sampling (PS, (Lartillot and Philippe, 2006) Lartillot & Philippe, 2006), stepping stone (SS, Xie et al., 2011), harmonic mean estimation (HME, Newton and Raftery, 1994), and smoothed harmonic mean estimation (sHME, Newton and Raftery, 1994). After ranking competing models based on MLEs, the Bayes factors (BF) calculated for competing models are evaluated. However, BFD may be prone to over-splitting, especially if the BFD is designed and tested using the same data (Grummer et al., 2014; Leaché et al., 2014). In addition, missing data can cause a potential problem with BFD (Noguerales et al., 2018).

Combining both fast, single-, or two-locus species delimitation methods (DNA barcoding gap, GMYC, P ID(Liberal), ABGD, bPTP, mPTP, BP&P) for the full data matrix with multi-locus coalescent-based methods (BP&P and BFD) for only those conflicting lineages essentially delimits or "identifies" population structure. This strategy requires interpretation alongside morphological, ecological, and genetic data (Sukumaran and Knowles, 2017; Luo et al., 2018) to make informed decisions about which lineages should be regarded formally as nominal species.

Results from numerous species delimitation approaches often conflict, leaving it to researchers to justify using the results from one method over the others. This is more often than not a subjective decision, although several papers have proposed more objective and repeatable approaches (Andújar et al., 2014; Bond and Stockman, 2008; Goldstein and DeSalle, 2011; Kekkonen and Hebert, 2014; Schlick-Steiner et al., 2010). We outline a multi-tier, integrative taxonomic approach (Fig. 1) where species, defined as diagnosable lineages, are considered as hypotheses engaged in a process of validation or modification (Barberousse and Samadi, 2010; De Queiroz, 2007). This context allows simultaneous tests of any number of initial species hypotheses (ISH) and their alternatives suggested by species delimitation methods by adding more data only for the conflicting lineages, and envisions additional tiers of refined hypothesis testing only for the incongruent lineages revealed from preceding tier analysis that also revisit the species diagnostic evidence for the alternative species hypothesis (ASH).

Using this approach, we explore the species delineation in the genus *Heptathela*. A parallel study outlines how morphological diagnostics arrives at the hypothesis (ISH) of 19 *Heptathela* species (Xu et al., accepted). In this paper, we use original data from the mitochondrial (COI) and the nuclear genomes (ITS2) from the Japanese islands' endemic set of species of the primitively segmented spider genus *Heptathela*. We first test ISH using a number of quick species delimitation methods using COI or/and ITS2 (the full data matrix) in tier 1 analysis. Fully recovered congruent lineages are directly included within the final species counts. Upon recovering conflicting lineages, we add more molecular makers only for the samples of those conflicting lineages and perform tier 2 analysis using multi-locus species delimitation methods to test the validity of competing species hypotheses derived from tier 1

analysis. Tier 3 analysis then reevaluates diagnostic evidence. Multi-tier hypothesis testing and re-diagnosis outlined in Figure 1 allow us to arrive at a plausibly corroborated species taxonomy.

2. Materials and methods

2.1. Taxon sampling

We carried out three extensive collection trips to Japanese islands, from Kyushu to the central Ryukyus (Fig. 2a). Our sampling is described in detail in the parallel study (Xu et al., accepted). Here, we select 180 *Heptathela* specimens as ingroups, and choose two species of the liphistiid genus *Ryuthela, R. nishihirai* (Haupt, 1979) and *R. shimojanai* Xu, Liu, Ono, Chen, Kuntner, & Li, 2017, also endemic to Japanese islands (Xu et al., 2017), as outgroups (Supplementary Table S1).

2.2. Molecular protocols and phylogenetic analyses

From leg muscles of specimens preserved in absolute ethanol we extracted genomic DNA using the Animal Genomic DNA Isolation Kit (Dingguo, Beijing, China). For 182 samples (in tier 1 analysis), we amplified and sequenced COI and ITS2 using the primer pairs LCO1490/HCO2198 (Folmer et al., 1994) and ITS-5.8S/ITS-28S (White et al., 1990), respectively. For 35 selected *Heptathela* samples of the incongruent lineages revealed in tier 1 analysis (see Results), we amplified and sequenced three more gene fragments, 16S rRNA, 28S rRNA, and histone 3 (H3), using the primer pairs 16Sar/16Sbr (Huber et al., 1993), 28S-O/28S-C (Hedin and Maddison, 2001), and H3aF/H3aR (Colgan et al., 1998), respectively. We followed previously reported standard protocols for all the gene fragments (Xu et al., 2015). We manually edited and aligned sequence data in Geneious v6.1.8 (Biomatters Ltd., 2012).

For phylogenetic inference using the data combined from both loci for 182 samples (tier 1 analysis), we produced two matrices, one partitioned by gene only, and the other by gene as well as codon position for COI. According to the greedy algorithm based on Akaike Information Criterion (AIC) in PartitionFinder v1.1.1 (Lanfear et al., 2012), the best substitution model for all partitions was GTR + I + G. For 37 samples (in tier 2 analysis), the best substitution models were HKY + G (16S), GTR + I + G (COI and ITS2), SYM + I + G (H3), and SYM + G (28S). For both datasets, we performed Bayesian-inference (BI) analyses in MrBayes v3.2.1 (Ronquist et al., 2012) by running Markov chain Monte Carlo (MCMC) for 50 million generations, and sampling trees every 5000 generations. We used TRACER v1.6 (Rambaut et al., 2014) to monitor stationarity, and FigTree v1.4.0 (Rambaut, 2012) to visualize and manipulate trees. Because both data partitions resulted in the same phylogenetic topology and comparable support values, we only used the matrix partitioned by gene in subsequent analyses.

We inferred phylogenetic trees using BI in MrBayes (as above), as well as maximum likelihood (ML) with autoMRE bootstrapping and a full ML search under the GTRGAMMA model in RAxML v8.2.11 (Stamatakis, 2014).

2.3. Multi-tier integrative taxonomic framework

2.3.1. Initial species hypothesis

Our parallel study provided the initial, 19-species hypothesis for all 180 specimens of

Heptathela species (for morphological details, see Xu et al., accepted). Alternatives to ISH are ASH from DNA species delimitation analyses (below).

2.3.2. Tier 1 analysis — single- or two-locus DNA species delimitation

In tier 1 analysis, we tested the initial hypothesis (ISH) using DNA barcoding, ABGD, P ID(Liberal), bPTP, mPTP and BP&P. In the DNA barcoding gap analysis, we used both Kimura two-parameter (K2P) and uncorrected *p*-distance (*p*-distance). In Mega v6.0.6 (Tamura et al., 2013) we calculated pairwise K2P and *p*-distances, mean intra- and interspecific K2P and *p*-distance for each putative species. Unlike DNA barcoding gap analysis, ABGD (Puillandre et al., 2012) does not require assigning terminals to putative species. ABGD calculates all pairwise distances in the dataset, evaluates intraspecific divergences, and then sorts the terminals into candidate species with calculated *p*-values. We performed ABGD analyses online (http://wwwabi.snv.jussieu.fr/public/abgd/), using three different distance metrics: Jukes-Cantor (JC69) (Jukes and Cantor, 1969), K2P (Kimura, 1980), and *p*-distance (Nei and Kumar, 2000). We analysed the data using two different values for the parameters Pmin (0.0001 and 0.001), Pmax (0.1 and 0.2), and relative gap width (X = 1 or 1.5), with other parameters at default values (see, Xu et al., 2017, 2015).

We obtained P ID(Liberal) statistical values from the species delimitation plugin (Masters et al., 2011) in Geneious v6.1.8 (Biomatters Ltd., 2012) as in our prior study (Xu et al., 2015). We used BI tree from the COI matrix as guide tree to calculate the mean probability of Intra/Inter genetic distance ratio for the initial 19 *Heptathela* species hypothesis.

We implemented bPTP analysis on its online server (http://species.h-its.org/ptp/; Zhang et al., 2013). This method is based on the model of Poisson tree processes, which relies on the intra- and interspecific substitution events (Zhang et al., 2013). BI tree from the COI matrix were input as guide tree and default parameter settings were used to explore the species hypotheses. We ran the analysis for 500,000 generations with a thinning of 500 and burn-in of 0.1, both with and without the outgroups (*R. nishihirai* and *R. shimojanai*). We performed mPTP analysis on ML tree from the COI matrix, and based on both ML and MCMC delimitations using the mPTP v0.2.4 (Kapli et al., 2017). Both analyses used the default –*multi* option, which incorporates differences in rates of coalescence among species, and the default minimum branch length of 0.0001. We ran MCMC analyses for 100 million generations, sampling every 10000, burn-in of the first 2 million generations. Analyses started from the ML species delimitation estimate, random delimitation and null delimitation, gave the same result.

BP&P v3.3 (Yang, 2015; Yang and Rannala, 2010) uses reversible-jump Markov chain Monte Carlo (rjMCMC) to calculate the posterior probabilities (PP) of different species delimitation models under a multispecies coalescent model (MSC) using the prior parameter settings, population size (theta, θ) and divergence times (tau, τ). The topology of the BI trees of 180 samples based on COI and COI + ITS2 were input as the guide tree in the rjMCMC species delimitation method. For 19 species hypothesis, we explored four different prior distributions for both two genes and single locus on the ancestral population size θ and root age τ_0 following Leaché and Fujita, 2010: (1) large population size, $\theta \sim$ G(1,10), and deep divergence time, $\tau_0 \sim$ G(1, 10); (2) small population size, $\theta \sim$ G(2, 2000), and shallow divergence time, $\tau_0 \sim$ G(2, 2000); (3) large population size, $\theta \sim$ G(1,10), and shallow divergence time, $\tau_0 \sim$ G(2, 2000); (4) small population size, $\theta \sim$ G(2, 2000), and deep divergence time, $\tau_0 \sim$ G(1, 10). We used the rjMCMC algorithm 0 with the fine-tuning parameter $\varepsilon = 15$, and alternatively algorithm 1 with $\alpha = 2$ and $m = 1$ to check for convergence (Hime et al., 2016). Analyses were run for 100,000 generations, with sampfreq = 5 and burnin = 20,000. Species tree nodes with PP > 0.95 were considered as evidence for

supporting species delimitation, whereas those with PP < 0.95 were considered as evidence for collapsing a species tree node.

2.3.3. Tier 2 analysis — testing competing species hypotheses derived from tier 1 analysis

In tier 2 analyses, we used two independent multi-locus coalescent-based species delimitation methods, BP&P (Yang, 2015; Yang and Rannala, 2010) and BFD (Grummer et al., 2014) to validate the incongruent lineages containing 35 *Heptathela* samples among 3-, 4- and 6-species hypotheses using five genes. BP&P analyses based on five gene fragments were used to validate these three competing species hypotheses as described above.

We performed BFD analyses in *BEAST v1.8.4 (Drummond et al., 2012). *BEAST analyses were run with a relaxed lognormal molecular clock, a yule process was set for species tree prior and piecewise linear & constant root for population size model, the lognormal priors for species.popMean and species.yule.birthRate (Starrett et al., 2018). Analyses were run for 50 million generation sampling every 10,000 generations, and the first 10% of trees were discarded as burn-in. We used TRACER v1.6 (Rambaut et al., 2014) to access convergence. We estimated the marginal likelihood (MLE) for each of three competing species models using PS and SS, as PS and SS outperform HME and sHME (Baele et al., 2012; Grummer et al., 2014). MLEs of both PS and SS were estimated with a chain length of 1000,000 generations and 100 path steps. We ranked the competing models by their MLEs, and used 2*ln*Bfs, as calculated as 2 × the difference in MLE between the best-fitting and alternative models, to assess the support for each model relative to the model with the highest ranking. We evaluated the models following Kass & Raftery, 1995: a 2*ln*Bf = 0-2 means "not worth more than a bare mention", 2*ln*Bf = 2-6 means "positive" support, 2*ln*Bf = 6-10 means "strong" support, and 2*ln*Bf > 10 provides "decisive" support in distinguishing between competing species delimitation hypotheses. Grummer et al. (2014) recommended a 2*ln*Bf > 10 for distinct lineages and we followed these guidelines in this study.

2.3.4. Tier 3 analysis — integrative taxonomic framework

Our framework (Fig. 1) allows for multi-tier testing of species hypotheses within an integrative taxonomic framework to redefine species diagnoses. Because in our test case some delimitation methods suggested ASH (20, or 22 species corresponding to 4-, or 6-species hypothesis in tier 2 analysis) to the ISH (19-species, or 3-species hypothesis in tier 2 analysis), an additional tier of testing simultaneously evaluated three competing hypotheses only within the recovered incongruent lineages from tier 2 analysis.

3. Results

3.1. Tier 1 analysis

Our original COI matrix of 180 *Heptathela* individuals of 676 bp had 248 variable and 239 parsimony informative sites, and ITS2 matrix of 359 bp had 20 variable and 19 parsimony informative sites. The Bayesian analyses on the concatenated data using both partition schemes agree on the topological details (Fig. 2b).

Barcoding gap analysis, ABGD, P ID(Liberal), bPTP, and mPTP were done on COI only. Most of these approaches (though not bPTP and mPTP) support ISH (i.e., 19 species) (Fig. 2b). A distinct gap between intra- and interspecific genetic distances for 19 hypothetical *Heptathela* species was detected, ranging from 4.2 to 5.7% for K2P and from 4.0 to 5.4% for *p*-distance (Fig. 3). The lowest mean interspecific and highest mean intraspecific distances

were 6/5.7% and 2.1/2.0% (K2P/uncorrected *p*-distance), respectively. ABGD analyses corroborated 19 species partitions based on different parameter combinations (Fig. 2b; Table S2). P ID(Liberal) also supported 19 species based on COI gene tree (Table S3 in supplementary). However, both bPTP and mPTP produce conflicted results: bPTP supported 22 species, whereas mPTP supported 20 species. Both methods with or without outgroups gave the same results.

BP&P, based on both COI and COI + ITS2 datasets, supported the 19-species hypothesis with very high support (PP > 0.95) of speciation events for most of nodes under small population size prior regardless whether divergence time was shallow or deep (i.e., the second and fourth prior settings), except when using the condition of large population size and deep divergence time (i.e., the first and third prior settings) (Table S4). The analyses using empirically estimated priors of two genes yielded similar results as the second prior setting, and the analyses using empirically estimated priors of single locus produced similar results as the second and fourth prior settings.

Despite the conflict among the various species delimitation methods in tier 1 analysis, 16 putative species were recognized by all the methods, supported by morphological diagnosis in our parallel taxonomic study (Xu et al., accepted). Three, four or six putative species were incongruent among different methods, deriving three ASH (Fig. 2b).

3.2. Tier 2 analysis

BP&P and BFD were used to test three competing species hypotheses and the results showed that BP&P supports all three species hypotheses (Table S4), indicating the lack of power of that method to delimit species within our dataset. Species delimitation analyses using BFD showed that the rankings based on MLEs of both PS and SS were in exact agreement with one another (Table 1). Both PS and SS determined that the 6-species model was best supported and significantly better than the 3-species model (2*ln*Bf: 90.55 and 92.53), but was only marginally better than the 4-species model (2*ln*Bf: 0.44 and 2.29).

3.3. Tier 3 analysis

Our tier 3 analysis — the diagnostic stage (Fig. 1) — is a reevaluation of morphological and/or DNA barcode diagnostic evidence for validating 4- or 6- species model supported by BFD. In our case, additional scrutiny revealed no additional morphological species diagnostics in the problem set of terminals (these are marked as spp. 17-19 under ISH in Fig. 2b). However, a closer examination of the COI barcode gaps within this set of 35 terminals rejected the 6-species hypothesis (barcoding gap: 1–1.51% based on K2P model), but supported the 4-species hypothesis (barcoding gap: 3.93–4.11% based on K2P model). Because the alternative hypothesis (20 species: 16 species delimited in tier 1 analysis, plus 4 species in tiers 2 & 3 analyses) is diagnostically supported, the species taxonomy is resolved under the assumption that DNA barcode diagnostic differences warrant species delimitation.

4. Discussion

Integrative taxonomy is a quickly evolving field, and many studies use innovative approaches (reviewed by Schlick-Steiner et al., 2010). Our multi-tier analysis strategy combines fast, single- or two-locus delimitation methods based on large, full data matrix, with multi-locus coalescent-based methods to validate only incongruent lineages revealed from tier 1 analysis by adding more molecular markers only for those conflicting lineages. Our results confirm the taxonomic utility of the COI barcoding region by detecting no conflict with the

equivalent two-locus analyses. Despite the theoretical advantages of using multi-locus, tree-based methods for more accurate species delimitations compared with single-locus approaches, the field has not arrived at any single species delimitation method that would be preferred for its reliability, cost-effectiveness, robustness, and congruence. Because this conflict among the approaches also emerges from our analyses, this gives the credibility to our approach to use as many methods as possible simultaneously, then objectively choose among them.

Single-locus, distance- and tree-based species delimitation methods are routinely used to delimit taxa, which are well differentiated genetically but not morphologically such as liphistiid and mygalomorph spiders (Derkarabetian and Hedin, 2014; Giarla et al., 2014; Satler et al., 2013; Xu et al., 2017, 2015). However, as phylogenetics widely acknowledges potential discordance among gene and species trees due to incomplete lineage sorting (e.g., Harrington and Near, 2012; McGuire et al., 2007), species inferences based on a single locus alone can be misleading. Therefore, tree-based species delimitation methods on multiple loci can overcome the difficulties encountered in single-locus, tree-based species delimitation methods as they uncouple gene trees and species trees, and furthermore allow gene tree coalescences to be older than species tree coalescences. Our results from all different analytical approaches suggest that this may not be a problem when using COI in species delimitation in tier 1 analysis to quickly test the initial hypothesis, at least in liphistiid spiders. Because all the methods in tier 1 analysis produced the same results using a single- versus double-locus datasets, we can conclude that COI may be informative enough, confirming its overall utility in arthropod taxonomy. One can and should of course use additional markers, however, ITS2 does not seem to be the best choice, in tier 1 analysis, at least for liphistiid spiders.

A comparison of the performance of the six sequence-based species delimitation approaches in tier 1 analysis is not trivial. Our prior studies on liphistiid species limits in other genera (Xu et al., 2017, 2015) have largely relied on the outcomes of three, DNA barcoding gap, ABGD, and P ID(Liberal). In this study, they all support the initial, morphology-based hypothesis (ISH, 19 species). On the other hand, mPTP supports more species (20) and bPTP even more (22). BP&P, designed to take the data of multiple genes and coalescent based, is used here based on both COI and COI + ITS2 also support 19 species in tier 1 analysis. BP&P is also used in tier 2 analysis to test the validity of three competing hypotheses using five genes, but it lacks the power to delimit as it supports the three alternative hypotheses derived from tier 1 analysis. Theoretical studies have shown that species delimitation using BP&P may capture population splits rather than species divergences (Jackson et al., 2017; Sukumaran and Knowles, 2017), thus leading to oversplitting when the amount of data (the number of loci) increases (Leaché et al., 2019). Therefore, more research is needed to evaluate the performance of this method relative to the others.

Although the phylogenetic approach applied to the sub-dataset composing of five genes for 35 samples of incongruent lineages from tier 1 analysis is unable to delimit these conflicting lineages (i.e., 3-, 4- or 6 species), the BFD method is not decisive as it best supports 6-species hypothesis but does not reject 4-species hypothesis based on the sub-dataset. One possible explanation is that the additional loci are not informative for a reliable test in this group of liphistiid spiders because there are very few unique nucleotides among the samples. Thus, more informative loci or genomic data are needed for the use of coalescent-based delimitation methods such as BFD in delimiting *Heptathela* and other liphistiids. As shown in a few theoretical and empirical studies, perhaps BFD is also prone to over-splitting, especially if the BFD is designed and tested using the same data (Grummer et al., 2014; Leaché et al., 2018, 2014). Given our sampling design and the propensity of BFD to over-splitting, our species delimitation results should be retested using informative, genomic

data in the future. Integrating all lines of evidence is also appropriate to avoid overestimation of species number.

In conclusion, our paper formalizes the necessary tiers of species delimitation analyses within an integrative taxonomic framework (Fig. 1). Our empirical study on primitively segmented spiders of the genus *Heptathela* endemic to Japanese islands confirms that our multi-tier approach is sufficiently effective and robust to delimit species accurately, but stresses the need to evaluate numerous lines of taxonomic evidence for an objective, repeatable test of species limits. This multi-tier approach can apply to species delimitation in many other domains of life.

Acknowledgements

We thank Arong Luo and Chengmin Shi for analytical advice. This study was supported by the grants from the National Science Foundation of China (NSFC-31601850; NSFC-31272324), the Hunan Provincial Natural Science Foundation of China (2017JJ3202), and a bilateral China and Slovenia exchange project No. 12-8 from China, Bi-CN/18-20-022 from Slovenia, the Singapore Ministry of Education AcRF Tier 1 grant (R-154-000-A52-114), and the Slovenian Research Agency (P1-10236, J1-672). In part, analyses relied on the resources of the National Supercomputing Centre, Singapore (https://www.nscc.sg).

Appendix Supplementary material

DNA sequences are deposited in GENBANK. DNA alignments and TableS1-S4 as supplementary materials can be found online.

References

Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., Zaharias, P., Puillandre, N., 2018. Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. Mol. Ecol. 27, 4591–4611. https://doi.org/10.1111/mec.14882

Agnarsson, I., Kuntner, M., 2007. Taxonomy in a changing world: seeking solutions for a science in crisis. Syst. Biol. 56, 531–539. https://doi.org/10.1080/10635150701424546

Andújar, C., Arribas, P., Ruiz, C., Serrano, J., Gómez-Zurita, J., 2014. Integration of conflict into integrative taxonomy: fitting hybridization in species delimitation of *Mesocarabus* (Coleoptera: Carabidae). Mol. Ecol. 23, 4344–4361. https://doi.org/10.1111/mec.12793

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A. V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167. https://doi.org/10.1093/molbev/mss084

Barberousse, A., Samadi, S., 2010. Species from Darwin onward. Integr. Zool. 5, 187–197. https://doi.org/10.1111/j.1749-4877.2010.00204.x

Barrett, R.D.H., Hebert, P.D.N., 2005. Identifying spiders through DNA barcodes. Can. J. Zool. 83, 481–491. https://doi.org/10.1139/z05-024

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22, 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., Ward, P.S., 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. 15, 271. https://doi.org/10.1186/s12862-015-0552-5

Blair, C., Bryson, R.W., 2017. Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: Phrynosoma). Mol. Ecol. Resour. 17, 1168–1182. https://doi.org/10.1111/1755-0998.12658

Bond, J.E., Hedin, M.C., Ramirez, M.G., Opell, B.D., 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Mol. Ecol. 10, 899–910. https://doi.org/10.1046/j.1365-294X.2001.01233.x

Bond, J.E., Stockman, A.K., 2008. An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. Syst. Biol. 57, 628–646. https://doi.org/10.1080/10635150802302443

Camargo, A., Sites, J., 2013. Species delimitation: a decade after the renaissance, in: Pavlinov, I.Y. (Ed.), The Species Problem - Ongoing Issues. InTech, New York, pp. 225–247. https://doi.org/10.5772/52664

Čandek, K., Kuntner, M., 2015. DNA barcoding gap: reliable species identification over morphological and geographical scales. Mol. Ecol. Resour. 15, 268–277. https://doi.org/10.1111/1755-0998.12304

Cardoso, A., Serrano, A., Vogler, A.P., 2009. Morphological and molecular variation in tiger beetles of the *Cicindela hybrida* complex: is an "integrative taxonomy" possible? Mol. Ecol. 18, 648–664. https://doi.org/10.1111/j.1365-294X.2008.04048.x

Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation. Mol. Ecol. 22, 4369–4383. https://doi.org/10.1111/mec.12413

Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R., 1998. Histone H3 and U2 snRNA DNA

sequences and arthropod molecular evolution. Aust. J. Zool. 46, 419. https://doi.org/10.1071/ZO98048

Dang, N.X., Sun, F.H., Lv, Y.Y., Zhao, B.H., Wang, J.C., Murphy, R.W., Wang, W.Z., Li, J.T., 2016. DNA barcoding and the identification of tree frogs (Amphibia: Anura: Rhacophoridae). Mitochondrial DNA Part A 27, 2574–2584. https://doi.org/10.3109/19401736.2015.1041113

Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. Lond. 85, 407–415.

De Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886. https://doi.org/10.1080/10635150701701083

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. https://doi.org/10.1371/journal.pone.0104982

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973. https://doi.org/10.1093/molbev/mss075

Eberle, J., Bazzato, E., Fabrizi, S., Rossini, M., Colomba, M., Cillo, D., Uliana, M., Sparacio, I., Sabatinelli, G., Warnock, R.C.M., Carpaneto, G., Ahrens, D., 2019. Sex-biased dispersal obscures species boundaries in integrative species delimitation approaches. Syst. Biol. 68, 441–459. https://doi.org/10.1093/sysbio/syy072

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. https://doi.org/10.1186/s12862-016-0659-3

Elias, M., Joron, M., Willmott, K., Silva-BrandÃo, K.L., Kaiser, V., Arias, C.F., PiÑerez, L.M.G., Uribe, S., Brower, A.V.Z., Freitas, A.V.L., Jiggins, C.D., 2009. Out of the Andes: patterns of diversification in clearwing butterflies. Mol. Ecol. 18, 1716–1729. https://doi.org/10.1111/j.1365-294X.2009.04149.x

Esselstyn, J.A., Evans, B.J., Sedlock, J.L., Anwarali Khan, F.A., Heaney, L.R., 2012. Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. Proc. R. Soc. B Biol. Sci. 279, 3678–3686. https://doi.org/10.1098/rspb.2012.0705

Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486

Flouri, T., Jiao, X., Rannala, B., Yang, Z., 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol. Biol. Evol. 35, 2585–2593. https://doi.org/10.1093/molbev/msy147

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.

Freudenstein, J. V., Broe, M.B., Folk, R.A., Sinn, B.T., 2017. Biodiversity and the species concept - lineages are not enough. Syst. Biol. 66, 644–656. https://doi.org/10.1093/sysbio/syw098

Giarla, T.C., Voss, R.S., Jansa, S.A., 2014. Hidden diversity in the Andes: comparison of species delimitation methods in montane marsupials. Mol. Phylogenet. Evol. 70, 137–151. https://doi.org/10.1016/j.ympev.2013.09.019

Goldstein, P.Z., DeSalle, R., 2011. Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. BioEssays. https://doi.org/10.1002/bies.201000036

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). Syst. Biol. 63, 119–133. https://doi.org/10.1093/sysbio/syt069

Hamilton, C. a, Formanowicz, D.R., Bond, J.E., 2011. Species delimitation and phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosidae): cryptic diversity in North American tarantulas. PLoS One 6, e26207. https://doi.org/10.1371/journal.pone.0026207

Hamilton, C. a, Hendrixson, B.E., Brewer, M.S., Bond, J.E., 2014. An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). Mol. Phylogenet. Evol. 71, 79–93. https://doi.org/10.1016/j.ympev.2013.11.007

Harrington, R.C., Near, T.J., 2012. Phylogenetic and coalescent strategies of species delimitation in snubnose darters (Percidae: *Etheostoma*). Syst. Biol. 61, 63–79. https://doi.org/10.1093/sysbio/syr077

Hebert, P.D.N., Barrett, R.D.H., 2005. Reply to the comment by L. Prendini on "Identifying spiders through DNA barcodes". Can. J. Zool. 83, 505–506. https://doi.org/10.1139/z05-026

Hebert, P.D.N., Ratnasingham, S., DeWaard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. 270, S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M., 2004. Identification of birds through DNA barcodes. PLoS Biol. 2, e312. https://doi.org/10.1371/journal.pbio.0020312

Hebert, P.D.N.N., Gregory, T.R., 2005. The promise of DNA barcoding for taxonomy. Syst. Biol. 54, 852–859. https://doi.org/10.1080/10635150500354886

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

Hedin, M.C., Maddison, W.P., 2001. Phylogenetic utility and evidence for multiple copies of elongation factor-1α in the spider genus *Habronattus* (Araneae: Salticidae). Mol. Biol. Evol. 18, 1512–1521. https://doi.org/10.1093/oxfordjournals.molbev.a003937

Heethoff, M., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K., Liow, L.H., Nowak, M.D., Stedje, B., Bachmann, L., 2018. Cryptic species – conceptual or terminological chaos? A response to Struck et al. Trends Ecol. Evol. 33, 310. https://doi.org/10.1016/j.tree.2018.02.008

Hickerson, M.J., Meyer, C.P., Moritz, C., 2006. DNA barcoding will often fail to discover new animal species over broad parameter space. Syst Biol 55, 729–739. https://doi.org/Doi 10.1080/10635150600969898

Hime, P.M., Hotaling, S., Grewelle, R.E., Neill, E.M.O., 2016. The influence of locus number and information content on species delimitation: an empirical test case in an endangered Mexican salamander. Mol. Ecol. 5959–5974. https://doi.org/10.1111/mec.13883

Huber, K., Haider, T., Mueller, M., 1993. DNA sequence data indicates the polyphyly of the family Ctenidae (Araneae). J. Arachnol. 194–201. https://doi.org/10.2307/3705482

Jackson, N.D., Carstens, B.C., Morales, A.E., O'Meara, B.C., 2017. Species delimitation with gene flow. Syst. Biol. 66, 799–812. https://doi.org/10.1093/sysbio/syw117

Jacobs, S.J., Kristofferson, C., Uribe-convers, S., Latvis, M., Tank, D.C., 2018. Incongruence in molecular species delimitation schemes: what to do when adding more data is difficult. Mol. Ecol. 27, 2397–2413. https://doi.org/10.1111/mec.14590

Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules, in: Munro, H. (Ed.), Mammalian Protein Metabolism. New York: Academic Press, pp. 21–132. https://doi.org/10.1016/B978-1-4832-3211-9.50009-7

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., Flouri, T., 2017.

Multi-rate poisson tree processes for single-locus species delimitation under Maximum Likelihood and Markov Chain Monte Carlo. Bioinformatics 33, 1630–1638. https://doi.org/10.1093/bioinformatics/btx025

Karanovic, T., Djurakic, M., Eberhard, S.M., 2016. Cryptic species or inadequate taxonomy? Implementation of 2D geometric morphometrics based on integumental organs as landmarks for delimitation and description of dopepod taxa. Syst. Biol. 65, 304–327. https://doi.org/10.1093/sysbio/syv088

Kekkonen, M., Hebert, P.D.N., 2014. DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. Mol. Ecol. Resour. 14, 706–715. https://doi.org/10.1111/1755-0998.12233

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120. https://doi.org/10.1007/BF01731581

Kress, W.J., García-Robledo, C., Uriarte, M., Erickson, D.L., 2015. DNA barcodes for ecology, evolution, and conservation. Trends Ecol. Evol. 30, 25–35. https://doi.org/10.1016/j.tree.2014.10.008

Kvist, S., 2016. Does a global DNA barcoding gap exist in Annelida? Mitochondrial DNA Part A 27, 2241–2252. https://doi.org/10.3109/19401736.2014.984166

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701. https://doi.org/10.1093/molbev/mss020

Lartillot, N., Philippe, H., 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55, 195–207. https://doi.org/10.1080/10635150500433722

Leaché, A.D., Fujita, M.K., 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. B Biol. Sci. 277, 3071–3077. https://doi.org/10.1098/rspb.2010.0662

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. https://doi.org/10.1093/sysbio/syu018

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

Leaché, A.D., Zhu, T., Rannala, B., Yang, Z., 2019. The spectre of too many species. Syst. Biol. 68, 168–181. https://doi.org/10.1093/sysbio/syy051

Leavitt, D.H., Starrett, J., Westphal, M.F., Hedin, M., 2015. Multilocus sequence data reveal dozens of putative cryptic species in a radiation of endemic Californian mygalomorph spiders (Araneae, Mygalomorphae, Nemesiidae). Mol. Phylogenet. Evol. 91, 56–67. https://doi.org/10.1016/j.ympev.2015.05.016

Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99–121. https://doi.org/10.1146/annurev-ecolsys-110512-135822

Lim, G.S., Balke, M., Meier, R., 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. Syst. Biol. 61, 165–169. https://doi.org/10.1093/sysbio/syr030

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. Syst. Biol. 67, 830–846. https://doi.org/10.1093/sysbio/syy011

Masters, B.C., Fan, V., Ross, H.A., 2011. Species delimitation - a geneious plugin for the exploration of species boundaries. Mol. Ecol. Resour. 11, 154–157. https://doi.org/10.1111/j.1755-0998.2010.02896.x

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013.

Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66, 526–38. https://doi.org/10.1016/j.ympev.2011.12.007

McGuire, J.A., Linkem, C.W., Koo, M.S., Hutchison, D.W., Lappin, A.K., Orange, D.I., Lemos-Espinal, J., Riddle, B.R., Jaeger, J.R., 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution (N. Y). 61, 2879–2897. https://doi.org/10.1111/j.1558-5646.2007.00239.x

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L., 2006. DNA barcoding and taxonomy in diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. 55, 715–728. https://doi.org/10.1080/10635150600969864

Meyer, C.P., Paulay, G., 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3, e422. https://doi.org/10.1371/journal.pbio.0030422

Miralles, A., Vences, M., 2013. New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. PLoS One 8, e68242. https://doi.org/10.1371/journal.pone.0068242

Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J.G., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst. Biol. 58, 298–311. https://doi.org/10.1093/sysbio/syp027

Moritz, C., Cicero, C., 2004. DNA barcoding: promise and pitfalls. PLoS Biol. 2, e354. https://doi.org/10.1371/journal.pbio.0020354

Moyle, R.G., Filardi, C.E., Smith, C.E., Diamond, J., 2009. Explosive pleistocene diversification and hemispheric expansion of a "great speciator". Proc. Natl. Acad. Sci. U. S. A. 106, 1863–1868. https://doi.org/10.1073/pnas.0809861105

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford University Press, New York.

Newton, M.A., Raftery, A.E., 1994. Approximate bayesian inference with the weighted likelihood bootstrap. J. R. Stat. Soc. Ser. B 56, 3–48. https://doi.org/0035-9246/94/56003

Noguerales, V., Cordero, P.J., Ortego, J., 2018. Integrating genomic and phenotypic data to evaluate alternative phylogenetic and species delimitation hypotheses in a recent evolutionary radiation of grasshoppers. Mol. Ecol. 27, 1229–1244. https://doi.org/10.1111/mec.14504

Ortiz, D., Francke, O.F., 2016. Two DNA barcodes and morphology for multi-method species delimitation in *Bonnetina* tarantulas (Araneae: Theraphosidae). Mol. Phylogenet. Evol. 101, 176–193. https://doi.org/10.1016/j.ympev.2016.05.003

Pante, E., Schoelinck, C., Puillandre, N., 2015. From integrative taxonomy to species description: one step beyond. Syst. Biol. 64, 152–160. https://doi.org/10.1093/sysbio/syu083

Paz, A., Crawford, A.J., 2012. Molecular-based rapid inventories of sympatric diversity: a comparison of DNA barcode clustering methods applied to geography-based vs clade-based sampling of amphibians. J. Biosci. 37, 887–896. https://doi.org/10.1007/s12038-012-9255-x

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamoun, S., Sumlin, W., Vogler, A., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609. https://doi.org/10.1080/10635150600852011

Prendini, L., 2005. Comment on "Identifying spiders through DNA barcodes". Can. J. Zool. 83, 498–504. https://doi.org/10.1139/z05-025

Puillandre, N., Lambert, a, Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol. Ecol. 21, 1864–77. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Rambaut, A., 2012. FigTree v. 1.4.3. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. TRACER v1.6. Retrieved from http://beast.bio.ed.ac.uk/Tracer

Rannala, B., 2015. The art and science of species delimitation. Curr. Zool. 61, 846–853. https://doi.org/10.1503/cmaj.109-4864

Rannala, B., Yang, Z., 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics 194, 245–253. https://doi.org/10.1534/genetics.112.149039

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Ruane, S., 2015. Using geometric morphometrics for integrative taxonomy: an examination of head shapes of milksnakes (genus *Lampropeltis*). Zool. J. Linn. Soc. 174, 394–413. https://doi.org/10.1111/zoj.12245

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54, 952–961. https://doi.org/10.1080/10635150500234674

Satler, J.D., Carstens, B.C., Hedin, M., 2013. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (mygalomorphae, antrodiaetidae, *Aliatypus)*. Syst. Biol. 62, 805–823. https://doi.org/10.1093/sysbio/syt041

Schlick-Steiner, B.C., Arthofer, W., Steiner, F.M., 2014. Take up the challenge! Opportunities for evolution research from resolving conflict in integrative taxonomy. Mol. Ecol. 23, 4192–4194. https://doi.org/10.1111/mec.12868

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. https://doi.org/10.1146/annurev-ento-112408-085432

Sites, J.W., Marshall, J.C., 2003. Delimiting species: a Renaissance issue in systematic biology. Trends Ecol. Evol. 18, 462–470. https://doi.org/10.1016/S0169-5347(03)00184-8

Song, C., Lin, X.L., Wang, Q., Wang, X.H., 2018. DNA barcodes successfully delimit morphospecies in a superdiverse insect genus. Zool. Scr. 47, 311–324. https://doi.org/10.1111/zsc.12284

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. 105, 13486–13491. https://doi.org/10.1073/pnas.0803076105

Spooner, D.M., 2009. DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. Am. J. Bot. 96, 1177–1189. https://doi.org/10.3732/ajb.0800246

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Starrett, J., Hayashi, C.Y., Derkarabetian, S., Hedin, M., 2018. Cryptic elevational zonation in trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus janus* complex) from the California southern Sierra Nevada. Mol. Phylogenet. Evol. 118, 403–413. https://doi.org/10.1016/j.ympev.2017.09.003

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S., Larsson, K., Liow, L.H., Nowak, M.D., Stedje, B., Bachmann, L., Dimitrov, D., 2018. Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33, 153–163. https://doi.org/10.1016/j.tree.2017.11.007

Sukumaran, J., Knowles, L.L., 2017. Multispecies coalescent delimits structure, not species.

Proc. Natl. Acad. Sci. 114, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Talavera, G., Lukhtanov, V.A., Rieppel, L., Pierce, N.E., Vila, R., 2013. In the shadow of phylogenetic uncertainty: the recent diversification of *Lysandra* butterflies through chromosomal change. Mol. Phylogenet. Evol. 69, 469–478. https://doi.org/10.1016/j.ympev.2013.08.004

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Tang, C.Q., Humphreys, A.M., Fontaneto, D., Barraclough, T.G., 2014. Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. Methods Ecol. Evol. 5, 1086–1094. https://doi.org/10.1111/2041-210X.12246

Toussaint, E.F.A., Morinière, J., Müller, C.J., Kunte, K., Turlin, B., Hausmann, A., Balke, M., 2015. Comparative molecular species delimitation in the charismatic Nawab butterflies (Nymphalidae, Charaxinae, *Polyura*). Mol. Phylogenet. Evol. 91, 194–209. https://doi.org/10.1016/j.ympev.2015.05.015

Vitecek, S., Kučinić, M., Previšić, A., Živić, I., Stojanović, K., Keresztes, L., Bálint, M., Hoppeler, F., Waringer, J., Graf, W., Pauls, S.U., 2017. Integrative taxonomy by molecular species delimitation: multi-locus data corroborate a new species of Balkan Drusinae micro-endemics. BMC Evol. Biol. 17, 129. https://doi.org/10.1186/s12862-017-0972-5

Wagner, C.E., Harmon, L.J., Seehausen, O., 2012. Ecological opportunity and sexual selection together predict adaptive radiation. Nature 487, 366–369. https://doi.org/10.1038/nature11144

Weigand, A.M., Jochum, A., Pfenninger, M., Steinke, D., Klussmann-Kolb, A., 2011. A new approach to an old conundrum-DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). Mol. Ecol. Resour. 11, 255–65. https://doi.org/10.1111/j.1755-0998.2010.02937.x

Wheeler, Q.D., Raven, P.H., Wilson, E.O., 2004. Taxonomy: impediment or expedient? Science. 303, 285. https://doi.org/10.1126/science.303.5656.285

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: PCR Protocols. Elsevier, pp. 315–322. https://doi.org/10.1016/B978-0-12-372180-8.50042-1

Xie, W., Lewis, P.O., Fan, Y., Kuo, L., Chen, M.H., 2011. Improving marginal likelihood estimation for bayesian phylogenetic model selection. Syst. Biol. 60, 150–160. https://doi.org/10.1093/sysbio/syq085

Xu, X., Liu, F., Chen, J., Li, D., Kuntner, M., 2015. Integrative taxonomy of the primitively segmented spider genus *Ganthela* (Araneae: Mesothelae: Liphistiidae): DNA barcoding gap agrees with morphology. Zool. J. Linn. Soc. 175, 288–306. https://doi.org/10.1111/zoj.12280

Xu, X., Liu, F., Chen, J., Ono, H., Agnarsson, I., Li, D., Kuntner, M., 2016. Pre-Pleistocene geological events shaping diversification and distribution of primitively segmented spiders on East Asian margins. J. Biogeogr. 43, 1004–1019. https://doi.org/10.1111/jbi.12687

Xu, X., Liu, F., Ono, H., Chen, J., Kuntner, M., Li, D., 2017. Targeted sampling in Ryukyus facilitates species delimitation of the primitively segmented spider genus *Ryuthela* (Araneae: Mesothelae: Liphistiidae). Zool. J. Linn. Soc. 181, 867–909. https://doi.org/10.1093/zoolinnean/zlx024

Xu X., Ono H., Kuntner M., Liu F., Li D. Inpress. A taxonomic monograph of the liphistiid spider genus *Heptathela,* endemic to Japanese islands. *Zookeys*.

Yang, Z., 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61, 854–865. https://doi.org/10.1093/czoolo/61.5.854

Yang, Z., Rannala, B., 2017. Bayesian species identification under the multispecies coalescent provides significant improvements to DNA barcoding analyses. Mol. Ecol. 26, 3028–3036. https://doi.org/10.1111/mec.14093

Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. 107, 9264–9269. https://doi.org/10.1073/pnas.0913022107

Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Table 1. Comparison of the three species models in tier 2 analysis using Bayes factor species delimitation (BFD).

Species models	**MLE (SS)**	rank	2*ln*BF	**MLE (PS)**	rank	2*ln*BF
3-species model	-5288.039	3	92.53	-5286.996	3	90.55
4-species model	-5241.994	2	0.44	-5242.867	2	2.29
6-species model	-5241.773	1		-5241.722	1	

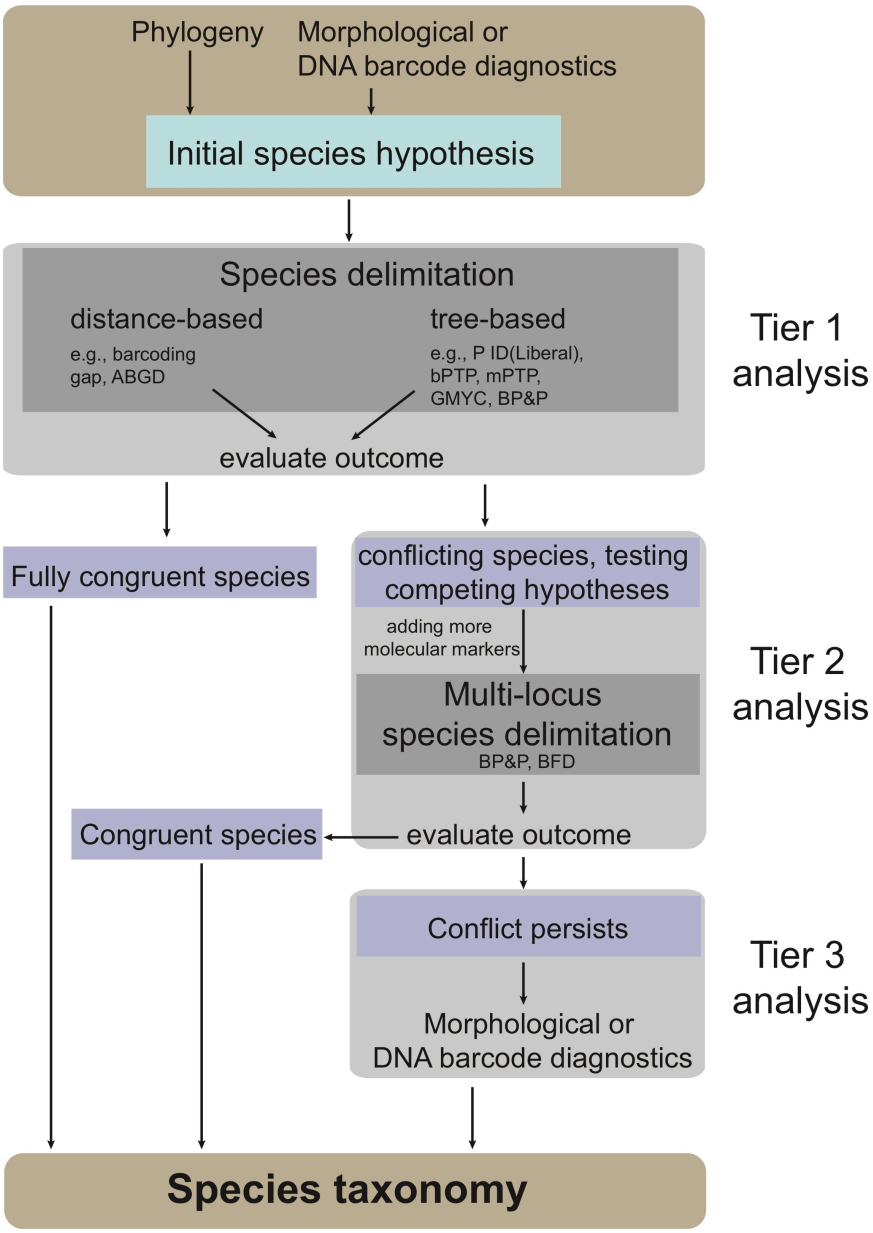

Fig. 1. A multi-tier species delimitation approach within an integrative taxonomy framework. Any number of initial species hypotheses (ISH) may derive from detailed morphological examination, DNA diagnostic features, geographical data, and topology from a specimen phylogeny. ISH is tested in tier 1 analysis using a variety of fast, single- or two-locus species delimitation methods. Fully congruent putative species supported by all the methods are included in the final species counts. If some of lineages are incongruent among these analyses, in tier 2 analysis, more molecular markers are added only for the samples of these conflicting lineages to form a new sub-dataset, which is subjected to multi-locus coalescent-based delimitation to validate the competing species hypotheses derived from tier 1 analysis. A species taxonomy can be finalized if one outcome is favored. In the case of further incongruence, a reevaluation of morphological or DNA barcode diagnostic evidence for alternative hypotheses (ASH) is performed in tier 3 analysis.

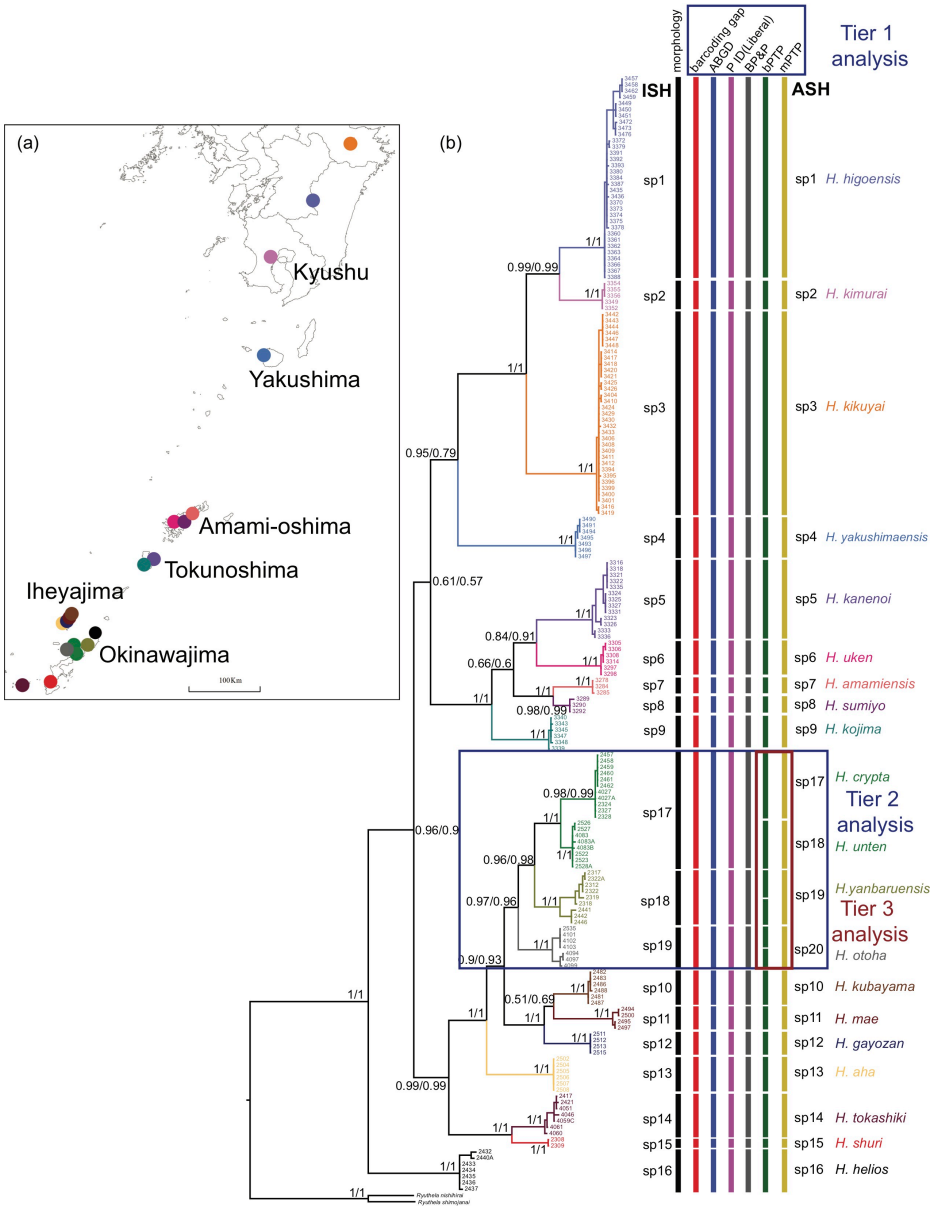

Fig. 2. A case study on the primitively segmented spiders. (a) Map showing representative sampling localities for species of *Heptathela* on Kyushu and Ryukyu archipelagos. (b) Bayesian specimen phylogeny on concatenated dataset for 180 *Heptathela* specimens and 2 *Ryuthela* outgroup specimens. Clades are color coded to match the map in A, clade supports are posterior probabilities for two different data partitions. ISH: initial species hypothesis (19 species), based on morphology, phylogenetic topology, and geographic information. Vertical bars corresponding to species numbers are estimations based on morphology, as well as on partial results of species delimitation methods. In tier 1 analysis, 16 species (spp. 1 – 16) are fully supported by all the delimitation methods, but three competing species hypotheses (i.e., 3, 4 or 6 species models) reflect conflicted clades. In tier 2 analysis, BP&P supports all three competing species hypotheses while BFD supports the 6- species hypothesis while not rejecting the 4-specices hypothesis. In tier 3 analysis, reevaluation of morphological and DNA barcode diagnostic evidence supports the 4-species hypothesis, which brings the total species taxonomy to 20 species. ASH: alternative species hypothesis.

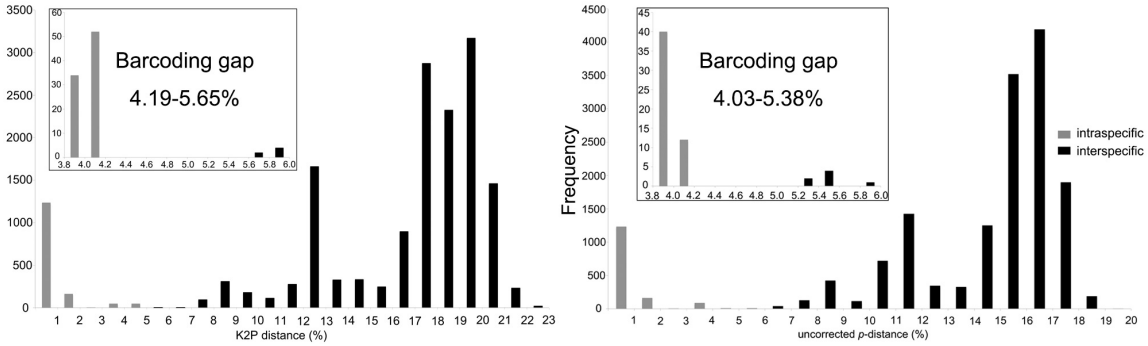

Fig. 3. DNA barcoding gap for 19 hypothetical *Heptathela* species.