Received: 5 November 2017 | Revised: 23 May 2018 | Accepted: 24 May 2018

DOI: 10.1111/1755-0998.12908

RESOURCE ARTICLE

WILEY MOLECULAR ECOLOGY RESOURCES

The hitchhiker's guide to single-locus species delimitation

Simon Dellicour[1,2] | Jean-François Flot[3,4]

[1]Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Leuven, Belgium

[2]Spatial Epidemiology Lab (SpELL), Université libre de Bruxelles, Bruxelles, Belgium

[3]Evolutionary Biology & Ecology, Université libre de Bruxelles, Bruxelles, Belgium

[4]Interuniversity Institute of Bioinformatics in Brussels – (IB)^2, Brussels, Belgium

Correspondence
Simon Dellicour, Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Herestraat 49, 3000 Leuven, Belgium.
Email: simon.dellicour@kuleuven.be
and
Jean-François Flot, Evolutionary Biology & Ecology, Université libre de Bruxelles, C.P. 160/12, Avenue F. D. Roosevelt 50, 1050 Bruxelles, Belgium.
Email: jflot@ulb.ac.be

Abstract

Molecular approaches to species delimitation are increasingly used to ascertain the number of species in a sample prior to taxonomic, ecological or physiological studies. Although multilocus approaches are gaining fast in popularity, single-gene methods still predominate in the literature. However, available simulation benchmarks of these methods focus exclusively on species-poor samples and/or tree-based approaches: as a result, travellers in the land of single-locus species delimitation lack a comprehensive "hitchhiker's guide" highlighting the sweet spots and dangers on their road. To fill this gap, we compared the performances of distance-based (ABGD, "automatic barcode gap discovery"), allele sharing-based (haplowebs) and tree-based approaches (GMYC, "generalized mixed Yule-coalescent" and PTP, "Poisson tree processes") to detect interspecific boundaries in samples of 6, 60 and 120 simulated species with various speciation rates, effective population sizes, mutation rates and sampling patterns. We found that all approaches performed poorly when population sizes and speciation rates were large, with haplowebs yielding best results followed by ABGD then tree-based approaches. The latter's error type was mostly oversplitting, whereas ABGD chiefly overlumped and haplowebs leaned either way depending on simulation parameters: such widely divergent error patterns suggest that, if all three types of methods agree, then the resulting delimitation is probably correct. Perfect congruence being quite rare, travellers in search of a one-size-fit-all approach to single-locus species delimitation should forget it; however, our hitchhiker's guide raises hope that such species delimitation's Holy Grail may be found in the relatively uncharted nearby land of multilocus species delimitation.

KEYWORDS
automatic barcode gap discovery, generalized mixed Yule-coalescent, haploweb, molecular markers, Poisson tree processes, species delimitation

1 | INTRODUCTION

Species delimitation methods are increasingly used in integrative taxonomic and systematic studies to delineate and identify species-level entities (Flot, 2015; Sites & Marshall, 2003). Indeed, species delimitation is a crucial prerequisite to, for example, ecological, population genetic or phylogeographic analyses, in which failure to ensure that all samples analysed are conspecific can lead to erroneous conclusions (Bortolus, 2008). In the last years, various single-locus and, more recently, multilocus species delimitation methods based on molecular data have been developed (Fontaneto, Flot, & Tang, 2015). While multilocus approaches have the appeal of simultaneously considering several unlinked loci when delineating species, they are, on the other hand, computationally very demanding when dealing with large data sets (Fujisawa, Aswad, & Barraclough, 2016). Most single-locus approaches are considerably faster to run, and

single-locus data are also much cheaper to acquire, especially when surveying large numbers of individuals. Besides, there are situations in which single-locus approaches are the only available option: for instance, when analysing existing barcode data obtained from the sequencing of a single, often mitochondrial, gene fragment (Ratnasingham & Hebert, 2007), or when surveying museum specimens for which the quality and quantity of DNA recovered is often limited. Finally, even when analysing multilocus data sets, it can also be interesting to compare results obtained from each locus considered independently in order to look at their level of congruence, for instance by means of a conspecificity matrix (Debortoli et al., 2016). For all these reasons, single-locus delimitation methods are still widely used and will likely continue to be so for many years. They also remain an active field of study, with new methods being regularly proposed and published.

Single-locus species delimitation methods fall under several broad categories (Flot, 2015; Fontaneto et al., 2015): distance-based approaches such as DNA barcoding and ABGD (automatic barcode gap discovery; Puillandre, Lambert, Brouillet, & Achaz, 2012); allele sharing-based approaches such as haplowebs (Flot, Couloux, & Tillier, 2010); and tree-based approaches that may be rooted in generalized mixed Yule-coalescent (GMYC) models (Pons et al., 2006) or in Poisson tree processes (PTP; Zhang, Kapli, Pavlidis, & Stamatakis, 2013). Even within a given category, there exists a wide variety of algorithms and implementations. For instance, there are at least three different implementations of GMYC models: single-threshold (Pons et al., 2006), multiple-threshold (Monaghan et al., 2009) and Bayesian (Reid & Carstens, 2012), that can yield strikingly different results when applied to the same data set (Dellicour & Flot, 2015). Rigorous, reproducible benchmarks using simulated data are therefore needed to quantify and compare the performances of the various approaches available under a wide range of parameters. However, nearly all simulation-based benchmarks have so far focused exclusively on tree-based approaches while neglecting distance-based and allele sharing-based methods (Esselstyn, Evans, Sedlock, Anwarali Khan, & Heaney, 2012; Fujisawa & Barraclough, 2013; Lohse, 2009; Papadopoulou et al., 2008; Reid & Carstens, 2012; Tang, Humphreys, Fontaneto, & Barraclough, 2014; Zhang et al., 2013). As a result, travellers in the land of species delimitation are left without a comprehensive hitchhiker's guide to inform them of the sweet spots and dangers on their road.

In 2015, we performed for the first time a benchmark (Dellicour & Flot, 2015) including approaches from the three main groups outlined above, namely: barcode gaps, three implementations of GMYC and haplowebs. For computational reasons, we only tested these methods on relatively species-poor simulated data sets comprising 1, 3 or 6 species. As ABGD is geared towards species-rich data sets (Puillandre et al., 2012), we did not include it in our benchmark; instead, the ability of distance-based approaches to delineate species correctly was assessed by comparing directly the distributions of intraspecific and interspecific distances, whereas the performance of haplowebs and GMYC-based approaches was analysed by computing the number of pairs of individuals correctly assigned as conspecific or heterospecific. One of the main conclusions of our article was that, among the methods we tested, only haplowebs can be used to detect species boundaries (or rather, the lack thereof) in data sets comprising a single species, whereas the overall performance of all methods improved when the number of species increased. This is not surprising as most of these methods (with the exception of haplowebs) were designed with species-rich data sets in mind; however, even for data sets of six species the various methods tested only worked well in the "sweet spot" where the effective population sizes and speciation rates of our simulated species were low.

Here, we present a new benchmark dedicated to relatively "species-rich" data sets comprising 6, 60 and 120 species. The 60-species and 120-species cases were chosen based on a quick survey of the usual number of species included in articles attempting single-locus species delimitation; although some articles include more species, we limited ourselves to a maximum of 120 species for computational reasons. In addition to the methods tested in our 2015 article, this time we included ABGD as well as PTP in our benchmark: although PTP is, like GMYC, a tree-based method, it does not require ultrametric trees because it models speciation rate using directly the number of substitutions (Zhang et al., 2013), which makes it faster and easier to run than GMYC while being, according to a previous simulation study (Tang et al., 2014), at least as effective. In order to simulate sampling patterns closer to those typical of real-world biodiversity studies (Ahrens et al., 2016; Lim, Balke, & Meier, 2012), we departed from the equal sampling we had used in our 2015 study and instead compared the results of two sampling schemes: a "short-tail" sampling patterns with many abundant species and a few rare ones, and a "long-tail" sampling scheme with a few abundant species and many rare ones. Last but not least, to include synthetic charts in our hitchhiker's guide we crafted condensed, graphically intuitive representations of our results in the form of ternary plots highlighting the relative influence of each simulation parameter on the performance of the approaches we tested. The resulting benchmark applies not only to single-locus studies sensu stricto but also to studies based on separate analyses of several markers (in contrast with multilocus approaches that analyse several markers simultaneously: such approaches, being computationally very intensive especially when dealing with species-rich data sets, were not included in our study).

2 | MATERIAL AND METHODS

2.1 | Data sets simulations

Simulations were performed using the python package DENDROPY 3.12.0 (Sukumaran & Holder, 2010). We first simulated trees of either 6, 60 or 120 species using a Yule model with three different birth rates (0.1, 1 and 10 births per lineage per million generations), then used standard coalescent conditions with population sizes of 10^4, 10^5 and 10^6 to simulate ten replicate genealogies for several sampling patterns. The selected birth rate values correspond to the range of speciation rates reported for a variety of organisms and the

effective population sizes to a range of biologically relevant values, as detailed in Esselstyn et al. (2012) as well as in Dellicour and Flot (2015). Although migration (i.e., gene flow between populations) may potentially affect the performance of the different methods we tested, we limited ourselves here to considering a single panmictic population for each simulated species.

We implemented two different sampling patterns in our simulations: For the "short-tail" pattern, we simulated 10 sequences per species (as obtained when sequencing five diploid individuals) for five-sixths of the species and 2 sequences per species (as obtained when sequencing a single diploid individual) for the remaining one-sixth; whereas for the "long-tail" pattern, we simulated 10 sequences per species for one-sixth of the species and two sequences per species for the remaining five-sixths. As in Dellicour and Flot (2015), each simulated genealogy was turned into three alignments of 1,000-bp sequences using mutation rates of 10^{-8}, 10^{-7} and 10^{-6} mutations per generation and per base. The lowest figure is the overall mutation rate reported for the human nuclear genome by Roach et al. (2010), and we also simulated mutation rates two orders of magnitude higher to make our benchmark relevant to the highly variable markers favoured in species delimitation studies (such as mitochondrial genes, intervening transcribed spacers and nuclear gene introns). As haplowebs require as input pairs of sequences from diploid individuals, we simulated them by forming random pairs of conspecific sequences (without replacement) from the alignments obtained above, thereby mimicking the process of sexual reproduction in panmictic populations (which is commonly modelled as picking randomly pairs of genes without replacement).

2.2 | Species delimitation methods

As in Dellicour and Flot (2015), we started by evaluating the potential for DNA barcoding to delineate accurately species by comparing the distributions of intraspecific and interspecific distances ("mismatch distributions"); the rationale being that, for distance-based approaches to work, intraspecific distances should be markedly smaller than interspecific ones (Fontaneto et al., 2015). Mismatch distributions and the corresponding Sarle's bimodality coefficients (a measure of bimodality that varies between 0 for perfectly unimodal data and 1 for perfectly bimodal data; bi- or multimodal distributions have Sarle's coefficient values between 5/9 and 1, whereas unimodal distribution have values below 5/9; Ellison, 1987; Pfister, Schwarz, Janczyk, Dale, & Freeman, 2013) were, respectively, computed using the R packages "pegas" (Paradis, 2010) and "moments" (Komsta & Novomestky, 2012).

Barcode gap analyses were performed using the command-line version of the ABGD (automatic barcode gap discovery) program (Puillandre et al., 2012). ABGD requires users to provide several parameters: the choice of a distance metrics (e.g., Jukes–Cantor or p-distances), a prior limit to intraspecific diversity (P) and a proxy for the minimum gap width (X). For each set of user parameters, it returns two delimitations: a "primary partition" and a "recursive one" obtained after applying its algorithm recursively. To decide which one of these many possible delimitations to follow, ABGD authors advise using other sources of information such as morphology. However, as our benchmark required a single delimitation output for each method, we ran ABGD under many different sets of parameters and chose the output most frequently obtained. Specifically, we first ran ABGD under its default set of parameters (Jukes–Cantor distances, P = 0.001000, 0.001668, 0.002783, 0.004642, 0.007743, 0.012915, 0.021544, 0.035938, 0.059948 et 0.100000, X = 1.5); when ABGD outputted "Only one partition found with your data. Nothing to output. You should try to rerun with a lower X" we redid the analysis with X = 1.0 then X = 0.5 (if ABGD still outputted the same message for X = 0.5, we concluded that it considered the data set as monospecific). For about one-third of our simulations, ABGD failed with the error message "The matrix is not symmetric", probably because some pairs of sequences in those data sets differed by more than 75% (making calculation of the distance impossible): hence, to be able to analyse all our data sets, we ran them using p-distances as well, then considered both the outputs obtained using Jukes–Cantor distances (when it worked) and p-distances when selecting the most frequent delimitation obtained for each data set. A script automatizing this approach is available on https://github.com/jflot/ABGDconsensus.

In contrast to distance-based approaches, haploweb analyses do not take into account the number of mutations separating two haplotypes but simply aggregate individuals into putative species on the basis of the haplotypes they share (Flot et al., 2010). Following an earlier proposal by Doyle (1995; reviewed in Sites & Marshall, 2003), putative species obtained using this method are called "fields for recombination" (in short, FFRs), and their nonoverlapping sets of haplotypes, "gene pools". Haploweb analyses were performed using an improved version of the Perl script countFFRs.pl used in Dellicour and Flot (2015); the new version of this script is available at https://github.com/jflot/countFFRs2. It takes as input a FASTA alignment of simulated sequences with names in the form "Txxx_yyy" (where xxx is the species ID and yyy is the sequence ID) and randomly forms diploid genotypes under the hypothesis that each species is panmictic. It then performs haploweb-based species delimitation on the resulting data sets and returns a list of individuals marked according to the species they have been assigned to.

To investigate the impact of phylogenetic reconstruction on the performance of GMYC and PTP methods, we performed these analyses both directly on our simulated genealogies and on the phylogenetic trees inferred from the DNA sequence alignments simulated from these genealogies. The GMYC method delineates species by fitting models of inter- and intraspecific processes to ultrametric phylogenetic trees in order to identify the limit between the two regimes (Fujisawa & Barraclough, 2013). To do so, it optimizes a mixed model that combines diversification between species (the Yule model, Yule, 1925) and genealogical branching within species (the neutral coalescent model, Hudson, 1990). The PTP method also aims to identify the transition points between inter- and intraspecific processes (modelled as independent Poisson processes) but specifically requires that tree branch lengths be proportional to the number of substitutions (rather than to time as for the GMYC approach).

For the GMYC and PTP methods, phylogenetic analyses were performed using the maximum-likelihood (ML) approach implemented in FASTTREE 2.1.8 (Price, Dehal, & Arkin, 2010). As GMYC methods require trees to be rooted, we used the mid-point rooting implemented in PATHD8 (Britton, Anderson, Jacquet, Lundqvist, & Bremer, 2007) to transform our unrooted ML trees into rooted ultrametric. We then used the R package "splits" (Ezard, Fujisawa, & Barraclough, 2013) to perform single-threshold GMYC (ST-GMYC) analyses, whereas analyses using the Bayesian implementation of GMYC (bGMYC) were performed by running the eponymous R package. For each analysis, bGMYC was run over 110,000 generations, discarding the first 10,000 as burn-in and sampling every 100 generations afterwards. Because bGMYC returns a probability of conspecificity for each pair of sequences in the data set, its output is not directly comparable to the groupings of individuals into putative species returned by other approaches. We therefore added as in Dellicour and Flot (2015) a discretization step (during which we retained only pairs of individuals that had a probability of conspecificity higher than 0.95) followed by a transitivization step (during which individuals were aggregated into species: individuals X and Y were considered conspecific if X was conspecific with Z and Y conspecific with Z, even if the probability of conspecificity of X and Y was lower than the aforementioned 0.95 threshold). In the text below, "bGMYC-1" and "bGMYC-2" refer respectively to the raw and discretized–transitivized results of bGMYC. For Poisson tree processes, we tested both the original method (PTP; Zhang et al., 2013) method and its Bayesian implementation (bPTP; http://species.h-its.org/ptp/). We initially planned to include the multiple-threshold GMYC (MT-GMYC) approach (Monaghan et al., 2009) in our comparison as in Dellicour and Flot (2015), but it failed to run properly when more than 6 species were simulated, producing outputs only for a small fraction of our simulated data sets (data not shown). Also, because of the computational challenge of simulating and analysing large number of sequences, we could not complete the benchmark of 120 species with the short-tail sampling pattern: hence, for the 120-species case, we report only results obtained using the long-tail sampling pattern.

2.3 | Performance assessment

To assess the performances of the different species delimitation methods tested, we calculated error rates following two distinct approaches: our original "pairwise sequence assessment" approach (Dellicour & Flot, 2015) and an adaptation of the methods used by Ratnasingham and Hebert (2013) and Eme et al. (2017) to compare the results of various species delimitation approaches, called henceforth "species-unit assessment".

For the pairwise sequence assessment approach, we calculated the percentage of conspecific pairs of sequences wrongly returned as heterospecific (hereafter referred to as %oversplitting) and the percentage of heterospecific pairs of sequences mistaken as conspecific (%overlumping). In the case of bGMYC, we also computed the percentage of pairs of sequences that were neither confidently split nor lumped, that is, for which the probability of conspecificity was between 0.05 and 0.95 (%indecision). Finally, we computed the percentage of pairs of sequences correctly determined (%success = 100−%oversplitting−%overlumping−%indecision).

For the species-unit assessment approach, we reported, for each simulation, the frequency of the following events: (a) match—all the sequences of a "true" species are detected as conspecific, and the corresponding inferred species does not include any other sequence (%match), (b) lump—the sequences in the inferred species belong to more than one actual species and include all the sequences inferred for those species (%lump), (c) split—the sequences in the inferred species belong to a single actual species but do not include all the sequences of that species (%split), and (d) reshuffle—the sequences in the inferred species belong to more than one actual species and do not include all the sequences of those species (%reshuffle). As the species-unit assessment approach compares the inferred species entities with the actual ones, it was not possible to use it for assessing the raw results of bGMYC (bGMYC-1), which reports only pairwise conspecificity probabilities. As a consequence, only the error rates of the discretized–transitivized output of bGMYC (bGMYC-2) were included in our species-unit assessment.

For each assessment approach, the different percentages were averaged over 10 replicate data sets simulated for each set of simulation parameters. In the case of haplowebs, diploid genotypes were simulated for each replicate data set by picking up randomly pairs of conspecific sequences, without replacement. This was performed 10 times per replicate data set, yielding a total of 100 replicate diploid populations per set of parameters.

For the "species-unit assessment", we also performed an overall comparison of the species delimitation results based on ternary plots. These triangular graphs were generated by plotting the % match, %split and (%lump + %reshuffle) on the three axes. On these plots, each dot corresponds to the results presented in a single barplot and thus corresponds to one 10-replicate average (or 100-replicate average in the case of haplowebs). Five ternary plots were generated for each delimitation method (one per simulation parameter) to help visualize these global trends.

3 | RESULTS AND DISCUSSION

3.1 | Pairwise sequence assessment vs. species-unit assessment

Using the pairwise sequence assessment approach, the performance of haplowebs, GMYC and PTP appeared to increase dramatically with the number of species in the data set, to the point that the influence of the various parameters could not be visually distinguished on the barplots for 120 species (Supporting Information Figures S1–S3). This is because, when there are many species, most pairs of heterospecific individuals attributed to the wrong species are still correctly assigned to the "heterospecific" category. By contrast, the plots obtained using the species-unit assessment approach were much less influenced by the number of species in the data sets,

allowing a precise dissection of the influence of each parameter even when dealing with 120 species (Figures 1–4): hence, in what follows we based ourselves solely on the results of the species-unit assessment. In contrast to the pairwise sequence assessment method, the species-unit assessment approach showed very little visible variation in the performance of the methods when dealing with 6, 60 or 120 species: this suggests that our result obtained using this approach may be extrapolated to even more species-rich situation comprising hundreds of species.

3.2 | Performance of distance-based species delimitation

Although the average number of pairwise differences was often lower within species than between them, intraspecific and interspecific mismatch distributions overlapped for most sets of parameters (Supporting Information Figures S4–S8). The only case in which they did not overlap (i.e., when a clear barcode gap existed) was for six species with low effective population sizes, low speciation rates and high mutation rates (Supporting Information Figures S4–S5); in the 60-species and 120-species cases, the distributions overlapped under all tested conditions. The distributions of Sarle's bimodality coefficients revealed that intraspecific distances were frequently multimodal, especially when the number of species was low; whereas the global distribution of mismatches was quite often unimodal, especially when the number of species was high and the mutation rate of the marker was low (Supporting Information Figure S9). The above observations were not influenced by the sampling pattern chosen ("long-tail" vs. "short-tail").

Congruent with these observations, we found that the main factors affecting the performance of ABGD were the effective population size and the speciation rate: this method was most effective for small population sizes and low speciation rates, whereas the success rate decreased dramatically for large populations speciating fast. Less dramatic was the influence of mutation rate (the higher, the better), whereas ABGD performed slightly worse on more speciose data sets and on simulations using the long-tail sampling pattern (Figures 1–4 and Supporting Information Figure S10). Overall the main type of error produced by ABGD was overlumping, notably when mutation rate was low and/or speciation rate was high; however, high mutation rates produced lots of oversplitting as well, whereas elevated speciation rates and effective population sizes led predominantly to reshuffling.

3.3 | Performance of allele sharing-based species delimitation

Compared with other approaches, haplowebs yielded results that were less variable (Supporting Information Figure S11), resulting in points tightly grouped (or even piled on top of each other) in the ternary diagrams. Most points were grouped in the top part of the triangle (Figure 4), highlighting a generally high success rate of species delimitation using this approach (Figures 1–3); however, some data sets were oversplit or overlumped, resulting in data points respectively near the lower left or lower right corners of the plot (Figure 4). If species delimitations were erroneous, it was usually because of overlumping when mutation rates, effective population sizes and/or speciation rates were low; whereas oversplitting was more prominent when using markers with high mutation rates, when dealing with species with high effective population sizes and/or when speciation rates were high (Figure 4 and Supporting Information Figure S10). The number of species did not influence the performance of the method, but the sampling pattern had a strong impact, with data points from "long-tail" sampling resulting in success or lumping and data points from "short-tail" sampling leaning more towards splitting (Figure 4).

3.4 | Performance of tree-based species delimitation

The data points for the four tree-based approaches tested (GMYC-ST, bGMYC, PTP and bPTP) were very dispersed in the ternary plots, revealing a high variability of their outcome (Figure 4 and Supporting Information Figure S10). Still, the position of the average points for each parameter value (displayed as crosses on the ternary diagrams) revealed a stronger tendency towards oversplitting in the case of GMYC-ST and bGMYC vs. a slightly better overall performance for PTP and bPTP under the range of parameters used in our simulations. In the case of GMYC-ST and bGMYC, there was no obvious influence of the mutation rates of the markers, but lumping increased when the speciation rate went up and success was higher when the effective population size was small (Figure 4). The success rate also increased slightly when the number of species increased, and the results were better for the "short-tail" sampling pattern than for the "long-tail" one (especially for ST-GMYC; Figures 1–4).

The influence of the mutation rate was much more prominent in the case of PTP and bPTP, with low mutation rate markers yielding more overlumping whereas high mutation rates led to oversplitting. In terms of speciation rate, however, the trends for PTP and bPTP were nearly identical to those of GMYC-ST and bGMYC: increased speciation rates led to more frequent reshuffling (Figures 1–4 and Supporting Information Figure S10). PTP and bPTP mostly gave correct answers when effective population sizes were small, whereas even for small effective population sizes GMYC-ST and bGMYC yielded frequently oversplit results (Figure 4). Sampling pattern also impacted the performance of the four tree-based approaches tested here, with long-tail sampling patterns yielding better results than short-tail ones (see below).

3.5 | Impact of sampling pattern

The tree-based and allele sharing-based approaches we tested performed better on data sets simulated using a long-tail sampling pattern, whereas ABGD performed slightly better on data sets simulated using the short-tail pattern. In the case of GMYC and PTP, our findings corroborate the study of Ahrens et al. (2016) highlighting that rarely sampled species are not problematic per se for tree--based approaches to species delimitation. The impact of sampling

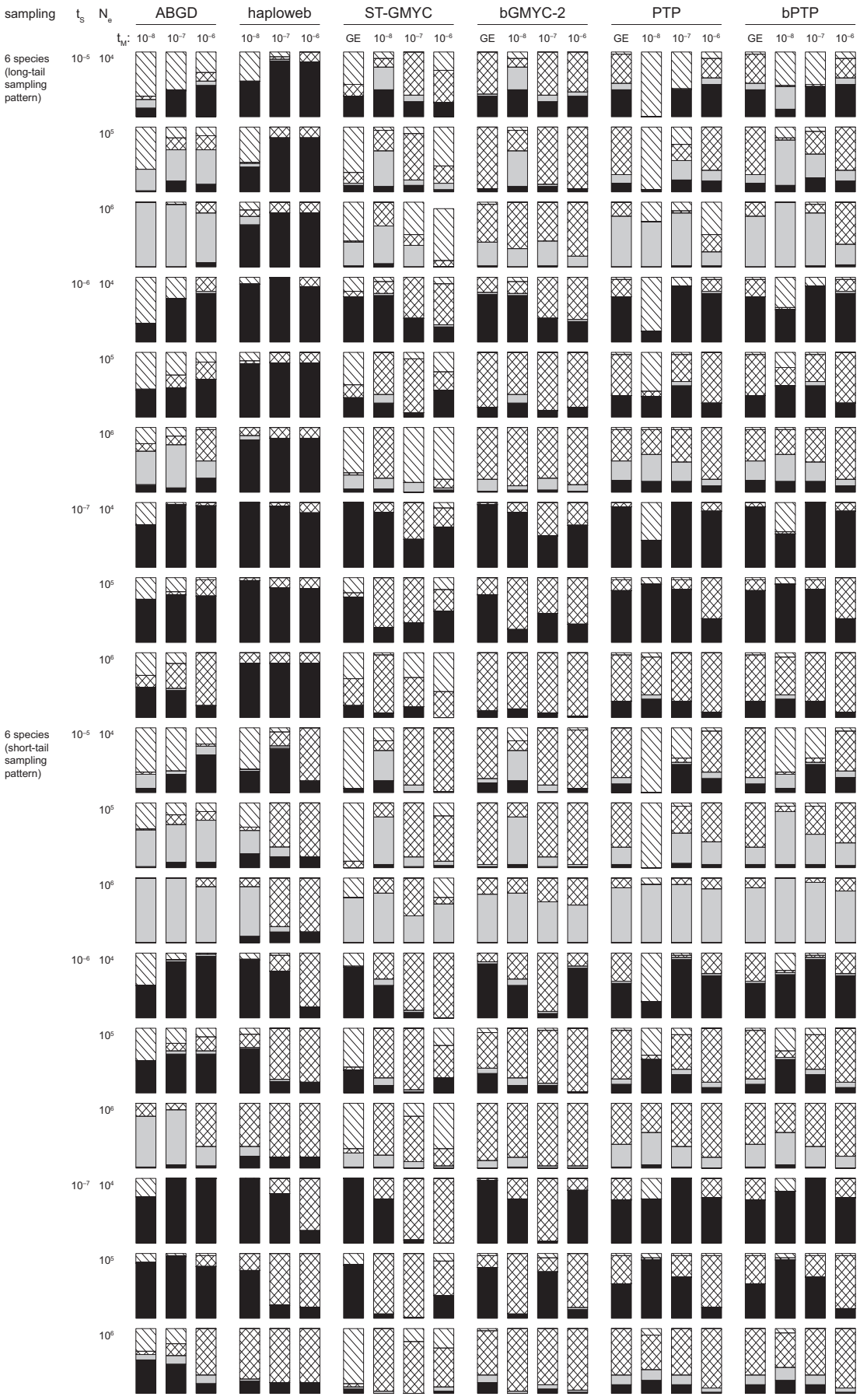
sampling
t_s
N_e
ABGD
haploweb
ST-GMYC
bGMYC-2
PTP
bPTP
t_M: 10^{-8} 10^{-7} 10^{-6} GE 10^{-8} 10^{-7} 10^{-6}
6 species (long-tail sampling pattern)
6 species (short-tail sampling pattern)
10^{-5}
10^{-6}
10^{-7}
10^4
10^5
10^6

FIGURE 1 Species delimitation results analysed with the "species-unit assessment" approach and based on the simulation of 6 species. Results are reported for both long-tail (5 species with 2 sampled sequences and 1 species with 10 sampled sequences) and short-tail sampling patterns (5 species with 10 sampled sequences and 1 species with 2 sampled sequences). t_S refers to the speciation rate, N_e to the effective (haploid) population size and t_M to the mutation rate (number of mutations per locus per generation). Results are reported in barplots with, from top to bottom, the percentages of lumping events (hatched), of splitting events (crosshatched), of reshuffling events (in grey) and of matching events (in black) detected when comparing the species delimitations with the true boundaries of the simulated species. For bGMYC, the reported percentages are those obtained after a discretization and transitivization step aimed at turning the probability matrices into species groupings (bGMYC-2, see details in text)

pattern was most dramatic for haplowebs: this is because the sampling of only two sequences per species is equivalent to sampling a single diploid individual, the two haplotypes of which are automatically considered conspecific using the haploweb approach (unlike tree-based or distance-based approaches that may consider them as heterospecific).

3.6 | Impact of species richness

Among the methods tested, ABGD distinguished itself by performing somewhat better on the less speciose data sets (Figures 1–4), whereas the performance of haplowebs was not affected significantly by the number of species (the averages for 6-species and 60-species data sets on Figure 4 are similar, whereas the performance for the 120-species data set appears higher but this is an artefact resulting from the fact that only results from the short-tail pattern were presented). Tree-based methods, in contrast, performed better on species-rich data sets, probably because such data sets allow better calibration of intraspecific vs. interspecific branching rates; species richness had a major influence on GMYC methods but was less dramatic for PTP approaches.

3.7 | Impact of effective population size

Overall, all methods performed better when applied to species with small effective population sizes than to species with large effective population sizes. In the case of distance-based approaches, this is because the low amount of genetic drift in species with large effective population sizes is inefficient at pruning their ancestral polymorphism, resulting in genetic distances as large within species as between them. Similarly, in the case of tree-based approaches, large effective population sizes slow down lineage sorting, resulting in a greater time lag between speciation and the acquisition of monophyly that is required by these approaches to delineate species (Flot et al., 2010). Even though haplowebs do not require species monophyly nor genetic distances smaller within species than between them, this approach is still negatively impacted by a large effective population size for two reasons: first, it increases the probability that a few haplotypes shared between species persist, resulting in lumping; and second, it increases haplotypic diversity, resulting in splitting as more individuals have to be sequenced to reach the point when all the haplotypes sampled from a given species are connected by heterozygous individuals.

3.8 | Impact of speciation rate

Similarly, all approaches worked better when applied to species with low speciation rates than to species with high speciation rates. This is because species simulated using a low speciation rate are on average older than species simulated using a high speciation rate, and ancient species have had more time to reach mutual allelic exclusivity (in the case of allele sharing-based approaches), reciprocal allelic monophyly (in the case of tree-based approaches) and/or minimally overlapping distributions of intra- vs. interspecific distances (in the case of distance-based approaches). The fact that, among these three stages, reciprocal allelic monophyly is the one that takes always the longest time to reach (because it implies the two others; Flot et al., 2010) probably explains the relatively lower performance of tree-based approaches compared to distance-based and allele sharing-based approaches under the range of parameters tested here.

3.9 | Impact of mutation rate

Mutation rates had a very different impact depending on the method. Barcode gap detection and most tree-based approaches (except GMYC-ST) worked better with markers that had high mutation rates, whereas GMYC-ST's performance was higher when mutation rate was low. In the case of haplowebs, the success rate did not vary much depending on mutation rate but affected the type of error: the most frequent error type was overlumping when low-variation markers were used, vs. oversplitting in the case of highly variable markers.

4 | CONCLUSIONS AND RECOMMENDATIONS

All in all, our present study confirms and extends the main findings of our previous exploration of the "species-poor" case (Dellicour & Flot, 2015). The combination of a small effective population size and a low speciation rate is globally confirmed as a "sweet spot" for both species-poor and species-rich data sets. Our study also corroborates the results of previous articles focusing on tree-based approaches (Esselstyn et al., 2012; Fujisawa & Barraclough, 2013). Notably, small effective populations sizes and low speciation rates were also previously identified by Esselstyn et al. (2012) as good conditions for the GMYC approach.

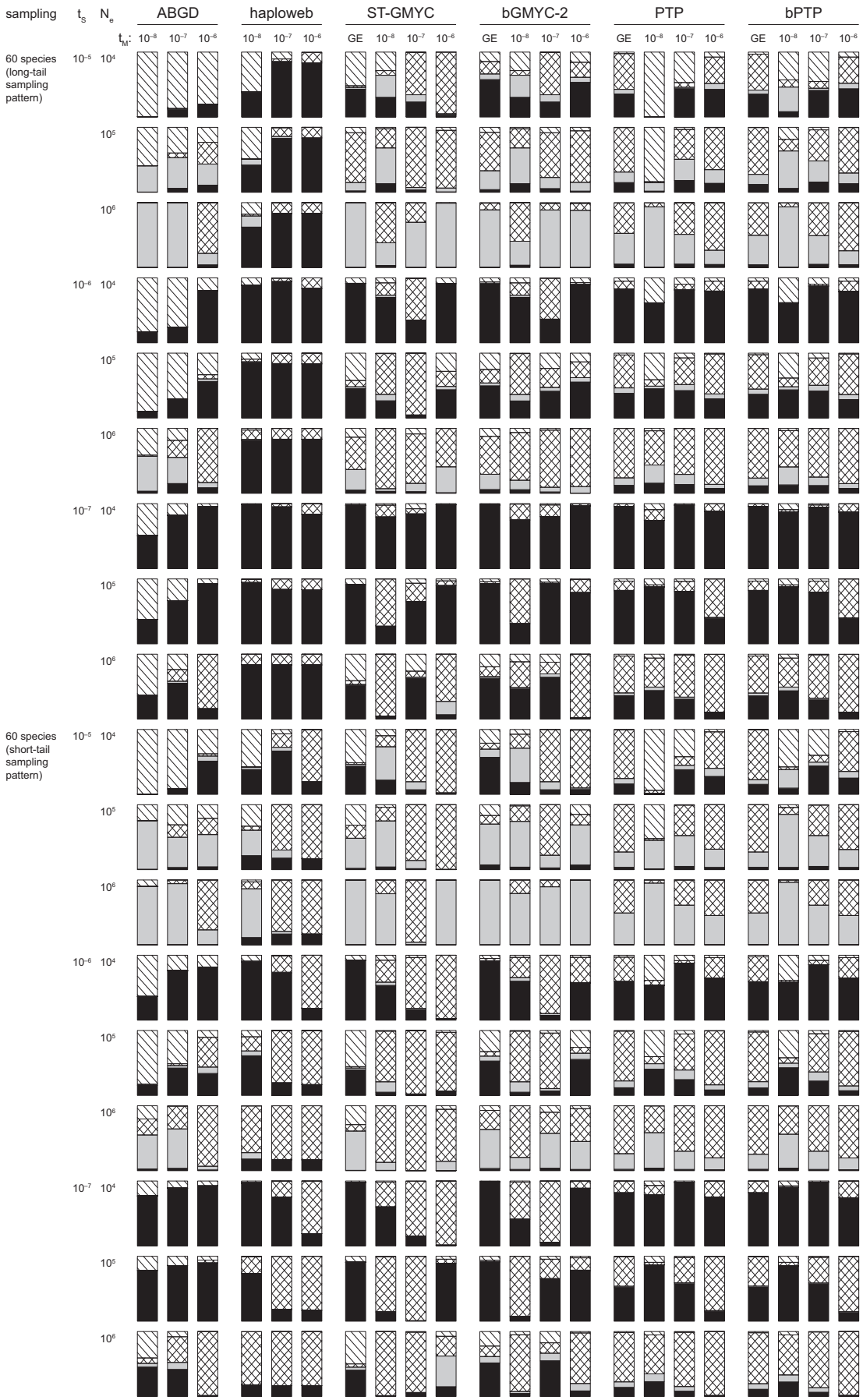
sampling
t_s
N_e
t_M:
ABGD
haploweb
ST-GMYC
bGMYC-2
PTP
bPTP
10^{-8}
10^{-7}
10^{-6}
GE
60 species (long-tail sampling pattern)
60 species (short-tail sampling pattern)
10^{-5}
10^{-6}
10^{-7}
10^4
10^5
10^6

FIGURE 2 Species delimitation results analysed with the "species-unit assessment" approach and based on the simulation of 60 species. Results are reported for both long-tail (50 species with 2 sampled sequences and 10 species with 10 sampled sequences) and short-tail sampling patterns (50 species with 10 sampled sequences and 10 species with 2 sampled sequences). t_S refers to the speciation rate, N_e to the effective (haploid) population size and t_M to the mutation rate (number of mutations per locus per generation). Results are reported in barplots with, from top to bottom, the percentages of lumping events (hatched), of splitting events (crosshatched), of reshuffling events (in grey) and of matching events (in black) detected when comparing the species delimitations with the true boundaries of the simulated species. For bGMYC, the reported percentages are those obtained after a discretization and transitivization step aimed at turning the probability matrices into species groupings (bGMYC-2, see details in text)

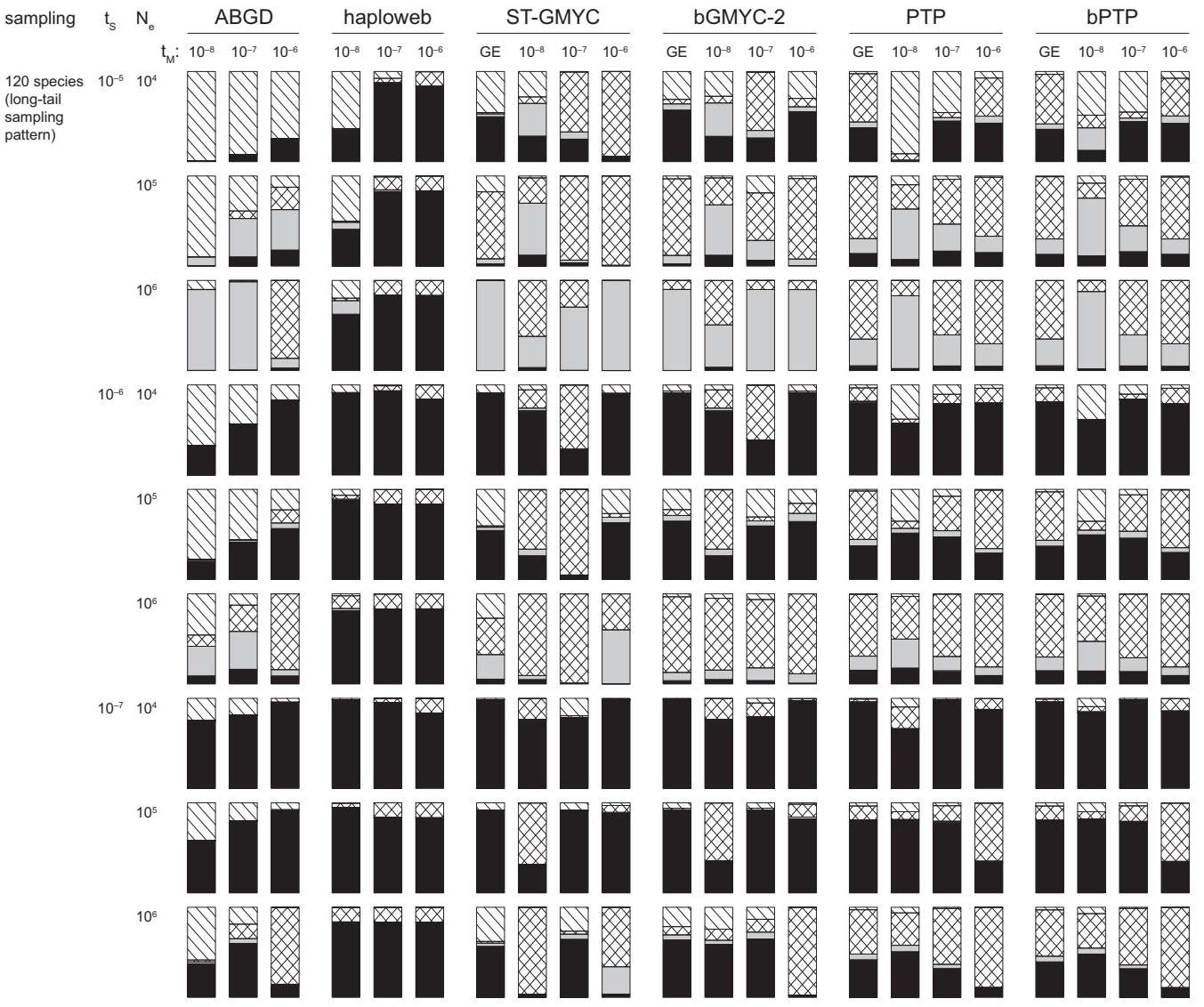

FIGURE 3 Species delimitation results analysed with the "species-unit assessment" approach and based on the simulation of 120 species (long-tail sampling pattern only: 100 species with 2 sampled sequences and 20 species with 10 sampled sequences; see text for further details). t_S refers to the speciation rate, N_e to the effective (haploid) population size and t_M to the mutation rate (number of mutations per locus per generation). Results are reported in barplots with, from top to bottom, the percentages of lumping events (hatched), of splitting events (crosshatched), of reshuffling events (in grey) and of matching events (in black) detected when comparing the species delimitations with the true boundaries of the simulated species. For bGMYC, the reported percentages are those obtained after a discretization and transitivization step aimed at turning the probability matrices into species groupings (bGMYC-2, see details in text)

Including for the first time in our benchmark distance-based (ABGD) and allele sharing-based (haploweb) approaches allowed us to compare them (Figures 1–3) and provide a synthetic "bird's eye view" of the relative performance of existing single-locus species delimitation approaches (Figure 4). Overall, the points for haplowebs are closer to the upper "match" vertex of the ternary graphs than those for ABGD,

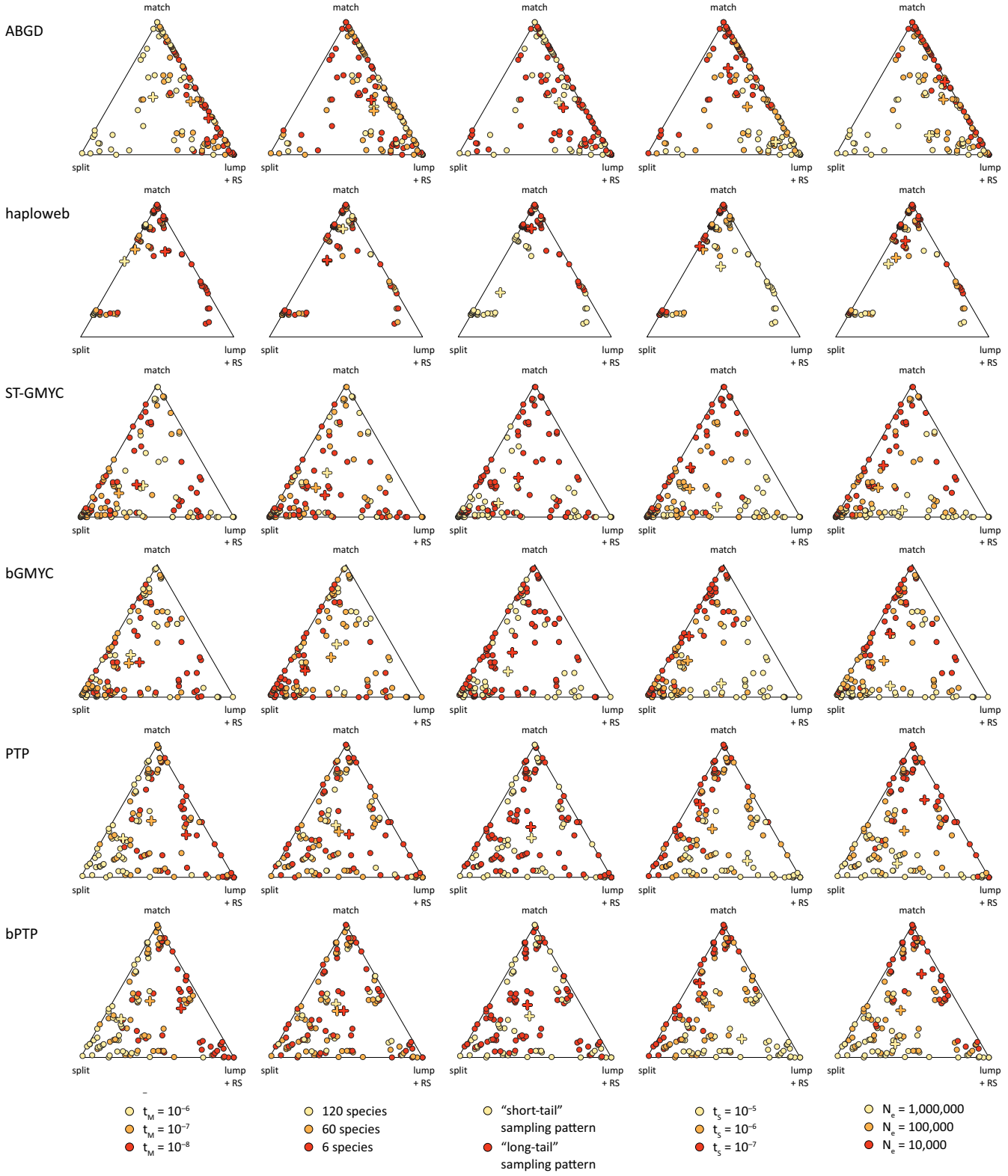

FIGURE 4 Bird's eye view of species delimitation results using the "species-unit assessment" method. The five ternary plots for each method display the same set of dots (one point per replicate) but with different colours highlighting the influence of one simulation parameter on the efficiency of the species delimitation method tested. On each triangle, the uppermost vertex corresponds to 100% match (i.e., each species is correctly delimited), the leftmost vertex corresponds to 100% splitting (i.e., all species are oversplit), and the rightmost vertex groups the remaining cases (overlumping and reshuffling). Crosses indicate the average position of each group of points

which are themselves higher in the plots compared to tree-based approaches. This suggests that haplowebs perform generally better than ABGD, which performs generally better than tree-based approaches. The same pattern is also observed for the dispersion of the points: data points in the ternary graphs are tightly grouped for haploweb, more dispersed for ABGD, and extremely dispersed for tree-based approaches (see also Supporting Information Figure S11). Hence, haploweb results are the most consistent, followed by ABGD then by tree-based approaches. Last but not least, the prominent type of error also varies from one method to the next: ABGD's main error type is overlumping, whereas the outcomes of tree-based approaches tend to cluster around the "oversplitting" vertex and haplowebs produce both oversplitting and overlumping depending on simulation parameters.

What advice may our hitchhiker's guide provide to fellow travellers in the land of species delimitation? If they are looking for a single, one-size-fit-all approach, we might return the same laconic statement as the 1972 edition of the Hitchhiker's Guide to Europe regarding Albania (then a communist dictatorship), p.42: "Forget it". On a more optimistic note, they may find comfort in the fact that the patterns of errors returned by ABGD, GMYC/PTP and haplowebs are widely different, suggesting that, if the three yield the same delimitation, then this delimitation should probably be correct. However, such congruence among approaches is rarely observed in actual studies (e.g., Miralles & Vences, 2013). Besides, although a "sweet spot" does exist where all methods tested perform relatively well, our barplots show that 100% species delimitation success is only achieved for 6-species data sets. For more speciose data sets, even an optimistic 99% per-species delimitation success rate implies 55% chances of correctly delimiting a 60-species data set, and 30% chances of delimiting correctly a 120-species data set. Hence, even for "ideal groups" characterized by small effective population sizes and low speciation rates (such as groundwater crustaceans; Flot et al., 2014; Copilaş-Ciocianu et al., 2017), one should always consider the results of any of the methods tested here with caution and cross-compare them with results obtained using other approaches (and/or using the same approach on other independent markers).

Using all three types of delimitation approaches on the same data set is only possible in the case of nuclear, diploid markers. To date, studies attempting to delineate species from single markers nearly always use haploid, organellar data sets (either chloroplastic or mitochondrial), as such markers are both easier to analyse (no double peaks) and have smaller effective population sizes than single-copy nuclear markers. However, ribosomal DNA markers such as intervening spacer sequences (ITSs) are also characterized by small effective population sizes due to their concerted mode of evolution, and their fast mutation rates make them ideal candidate markers for single-locus species delimitation. Although they remain rare, studies using single diploid sequence markers to delineate species are presently on the rise, either alone (Adjeroud et al., 2014; Daïnou et al., 2016) or in conjunction with an organellar marker such as COI (Copilaş-Ciocianu et al., 2017; Flot, Dahl, & André, 2013; Papakostas et al., 2016). Our take-home message may therefore be: if you sequence a single marker to delineate species, make it a nuclear one; and if you sequence two markers, include at least one nuclear marker (instead of piling up two mitochondrial or chloroplastic DNA regions, which behave as a single marker because they are linked together on the same molecule).

Indeed, the best way to overcome the limitations of single-locus approaches is to use information from several independent loci. Several tree-based multilocus methods have been already proposed, notably BPP (Rannala & Yang, 2013; Yang & Rannala, 2010), DISSECT (Jones, Aydin, & Oxelman, 2015), SpeDeSTEM (Ence & Carstens, 2011) and most recently PHRAPL (Jackson, Morales, Carstens, & O'Meara, 2017). Such approaches may be able to deal with high speciation rates and large population sizes by leveraging information provided by several independent markers, but given their very high computing requirements, their performances have not yet been thoroughly assessed. Yet, a first simulation-based assessment of BPP recently showed that this popular approach does not actually delineate species but differentiated populations (Sukumaran & Knowles, 2017). Given the relatively better performance of allele sharing-based and distance-based approaches over tree-based approaches observed in our benchmark of single-locus species delimitation, it seems reasonable to assume that multilocus approaches based on allele sharing and/or distances would perform better than BPP, but this hypothesis will have to be tested.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers for their helpful comments on this study. We are also grateful to Yann Spöri for his help improving the ABGDconsensus script. S.D. is funded by the Fonds Wetenschappelijk Onderzoek (FWO, Belgium) as well as the Fonds National de la Recherche Scientifique (FNRS, Belgium) and was previously funded by the Wiener-Anspach Foundation. Computational resources were provided by the Belgian High Performance Computing Centre cofunded by ULB and VUB (HPC cluster "Hydra").

AUTHORS' CONTRIBUTIONS

S.D. and J.-F.F. jointly conceived the study, designed its methodology, analysed the data and wrote the resulting manuscript.

DATA ACCESSIBILITY

The scripts used are available online at https://github.com/jflot/ABGDconsensus and https://github.com/jflot/countFFRs2.

ORCID

Simon Dellicour http://orcid.org/0000-0001-9558-1052
Jean-François Flot http://orcid.org/0000-0003-4091-7916

REFERENCES

Adjeroud, M., Guérécheau, A., Vidal-Dupiol, J., Flot, J. F., Arnaud-Haond, S., & Bonhomme, F. (2014). Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: A multi-scale

analysis in an insular, fragmented reef system. *Marine Biology*, *161*, 531–541. https://doi.org/10.1007/s00227-013-2355-9

Ahrens, D., Fujisawa, T., Krammer, H.-J., Eberle, J., Fabrizi, S., & Vogler, A. P. (2016). Rarity and incomplete sampling in DNA-based species delimitation. *Systematic Biology*, *65*, 478–494. https://doi.org/10.1093/sysbio/syw002

Bortolus, A. (2008). Error cascades in the biological sciences: The unwanted consequences of using bad taxonomy in ecology. *Ambio*, *37*, 114–118. https://doi.org/10.1579/0044-7447(2008)37[114:ECITBS]2.0.CO;2

Britton, T., Anderson, C. L., Jacquet, D., Lundqvist, S., & Bremer, K. (2007). Estimating divergence times in large phylogenetic trees. *Systematic Biology*, *56*, 741–752. https://doi.org/10.1080/10635150701613783

Copilaș-Ciocianu, D., Fišer, C., Borza, P., Balázs, G., Angyal, D., & Petrusek, A. (2017). Low intraspecific genetic divergence and weak niche differentiation despite wide ranges and extensive sympatry in two epigean *Niphargus* species (Crustacea: Amphipoda). *Zoological Journal of the Linnean Society*, *181*, 485–499. https://doi.org/10.1093/zoolinnean/zlw031

Daïnou, K., Blanc-Jolivet, C., Degen, B., Kimani, P., Ndiade-Bourobou, D., Donkpegan, A. S. L., … Hardy, O. J. (2016). Revealing hidden species diversity in closely related species using nuclear SNPs, SSRs and DNA sequences – A case study in the tree genus *Milicia*. *BMC Evolutionary Biology*, *16*, 259. https://doi.org/10.1186/s12862-016-0831-9

Debortoli, N., Li, X., Eyres, I., Fontaneto, D., Hespeels, B., Tang, C. Q., … Van Doninck, K. (2016). Genetic exchange among bdelloid rotifers is more likely due to horizontal gene transfer than to meiotic sex. *Current Biology*, *26*, 723–732. https://doi.org/10.1016/j.cub.2016.01.031

Dellicour, S., & Flot, J. F. (2015). Delimiting species-poor data sets using single molecular markers: A study of barcode gaps, haplowebs and GMYC. *Systematic Biology*, *64*, 900–908. https://doi.org/10.1093/sysbio/syu130

Ellison, A. M. (1987). Effect of seed dimorphism on the density-dependent dynamics of experimental populations of *Atriplex triangularis* (Chenopodiaceae). *American Journal of Botany*, *74*, 1280–1288. https://doi.org/10.1002/j.1537-2197.1987.tb08741.x

Eme, D., Zagmajster, M., Delić, T., Fišer, C., Flot, J. F., Konecny-Dupré, L., … Malard, F. (2017). Do cryptic species matter in macroecology? Sequencing European groundwater crustaceans yields smaller ranges but does not challenge biodiversity determinants. *Ecography*, *41*, 424–436. https://doi.org/10.1111/ecog.02683

Ence, D. D., & Carstens, B. C. (2011). SPEDESTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources*, *11*, 473–480. https://doi.org/10.1111/j.1755-0998.2010.02947.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. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 3678–3686. https://doi.org/10.1098/rspb.2012.0705

Ezard, T., Fujisawa, T., & Barraclough, T. (2013). *R package splits: SPecies' LImits by Threshold Statistics, version 1.0-18/r45*. Retrieved from http://r-forge.r-project.org/projects/splits/

Flot, J. F. (2015). Species delimitation's coming of age. *Systematic Biology*, *64*, 897–899. https://doi.org/10.1093/sysbio/syv071

Flot, J. F., Bauermeister, J., Brad, T., Hillebrand-Voiculescu, A., Sarbu, S. M., & Dattagupta, S. (2014). *Niphargus-Thiothrix* associations may be widespread in sulfidic groundwater ecosystems: Evidence from southeastern Romania. *Molecular Ecology*, *23*, 1405–1417. https://doi.org/10.1111/mec.12461

Flot, J. F., Couloux, A., & Tillier, S. (2010). Haplowebs as a graphical tool for delimiting species: A revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evolutionary Biology*, *10*, 372. https://doi.org/10.1186/1471-2148-10-372

Flot, J. F., Dahl, M., & André, C. (2013). *Lophelia pertusa* corals from the Ionian and Barents seas share identical nuclear ITS2 and near-identical mitochondrial genome sequences. *BMC Research Notes*, *6*, 144. https://doi.org/10.1186/1756-0500-6-144

Fontaneto, D., Flot, J. F., & Tang, C. Q. (2015). Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity*, *45*, 433–451. https://doi.org/10.1007/s12526-015-0319-7

Fujisawa, T., Aswad, A., & Barraclough, T. G. (2016). A rapid and scalable method for multilocus species delimitation using Bayesian model comparison and rooted triplets. *Systematic Biology*, *65*, 759–771. https://doi.org/10.1093/sysbio/syw028

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, *62*, 707–724. https://doi.org/10.1093/sysbio/syt033

Hudson, R. R. (1990). Gene genealogies and coalescent process. In D. J. Futuyma, & J. Antonivics (Eds.), *Oxford surveys in evolutionary biology* (pp. 1–44). Oxford, UK: Oxford University Press.

Jackson, N. D., Morales, A. E., Carstens, B. C., & O'Meara, B. C. (2017). PHRAPL: Phylogeographic inference using approximate likelihoods. *Systematic Biology*, *66*, 1045–1053. https://doi.org/10.1093/sysbio/syx001

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: An assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, *31*, 991–998. https://doi.org/10.1093/bioinformatics/btu770

Komsta, L., & Novomestky, F. (2012). *MOMENTS: Moments, cumulants, skewness, kurtosis and related tests*. R package version 0.13. Available from http://cran.r-project.org/web/packages/moments/

Lim, G. S., Balke, M., & Meier, R. (2012). Determining species boundaries in a world full of rarity: Singletons, species delimitation methods. *Systematic Biology*, *61*, 165–169. https://doi.org/10.1093/sysbio/syr030

Lohse, K. (2009). Can mtDNA barcodes be used to delimit species? A response to Pons et al (2006) *Systematic Biology*, *58*, 439–442. https://doi.org/10.1093/sysbio/syp039

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., … Vogler, A. P. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, *58*, 298–311. https://doi.org/10.1093/sysbio/syp027

Papadopoulou, A., Bergsten, J., Fujisawa, T., Monaghan, M. T., Barraclough, T. G., & Vogler, A. P. (2008). Speciation and DNA barcodes: Testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 2987–2996. https://doi.org/10.1098/rstb.2008.0066

Papakostas, S., Michaloudi, E., Proios, K., Brehm, M., Verhage, L., Rota, J., … Declerck, S. A. J. (2016). Integrative taxonomy recognizes evolutionary units despite widespread mitonuclear discordance: Evidence from a rotifer cryptic species complex. *Systematic Biology*, *65*, 508–524. https://doi.org/10.1093/sysbio/syw016

Paradis, E. (2010). PEGAS: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*, 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pfister, R., Schwarz, K. A., Janczyk, M., Dale, R., & Freeman, J. (2013). Good things peak in pairs: A note on the bimodality coefficient. *Frontiers in Psychology*, *4*, 700.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., … Vogler, A. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*, 595–609. https://doi.org/10.1080/10635150600852011

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FASTTREE 2 – Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, *5*, e9490. https://doi.org/10.1371/journal.pone.0009490

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

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

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364.

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. *PLoS ONE*, *8*, e66213. https://doi.org/10.1371/journal.pone.0066213

Reid, N., & Carstens, B. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology*, *12*, 196. https://doi.org/10.1186/1471-2148-12-196

Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., … Galas, D. J. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, *328*, 636–639. https://doi.org/10.1126/science.1186802

Sites, J. W. Jr, & Marshall, J. C. (2003). Delimiting species: A renaissance issue in systematic biology. *Trends in Ecology & Evolution*, *18*, 462–470. https://doi.org/10.1016/S0169-5347(03)00184-8

Sukumaran, J., & Holder, M. T. (2010). DENDROPY: A Python library for phylogenetic computing. *Bioinformatics*, *26*, 1569–1571. https://doi.org/10.1093/bioinformatics/btq228

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

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 in Ecology and Evolution*, *5*, 1086–1094. https://doi.org/10.1111/2041-210X.12246

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

Yule, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *213*, 21–87. https://doi.org/10.1098/rstb.1925.0002

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Dellicour S, Flot J-F. The hitchhiker's guide to single-locus species delimitation. *Mol Ecol Resour*. 2018;00:1–13. https://doi.org/10.1111/1755-0998.12908