MITOCHONDRIAL DNA PART A, 2017
https://doi.org/10.1080/24701394.2017.1413364

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Comparative analyses of species delimitation methods with molecular data in snappers (Perciformes: Lutjaninae)

Raimundo da Silva[a], Pedro L. V. Peloso[b], Marcelo José Sturaro[c], Ivana Veneza[a], Iracilda Sampaio[d], Horacio Schneider[d] and Grazielle Gomes[a]

[a]Laboratório de Genética Aplicada, Universidade Federal do Pará, Bragança, Brazil; [b]Coordenação de Zoologia, Museu Paraense Emílio Goeldi, Belém, Brazil; [c]Laboratório de Herpetologia, Museu Paraense Emílio Goeldi, Belém, Brazil; [d]Laboratório de Genética e Biologia Molecular, Universidade Federal do Pará, Bragança, Brazil

ABSTRACT

The integration of approaches that allow the incorporation of stochasticity of gene histories with phylogenetic methods resulted in new approaches for the old issue of species delimitation. Nevertheless, coalescent methods seem problematic for taxa with large effective population size and shallow temporal diversification (like marine fishes). Here, we investigate the performance of single-locus (cytochrome oxidase 1, commonly used in DNA barcoding initiatives) methods for molecular species delimitation in snappers of Lutjaninae from the Western Atlantic and Pacific Eastern. Our results show incongruences among methods. ABGD, PTP and mPTP trend towards a lower number of estimated species. Phylogenetic-coalescent methods with single threshold were majority congruent for a same number of lineages. On the other hand, algorithms with multiple thresholds tend to estimate a higher number of potential species. We do not endorse the use of single-locus for species delimitation, but we do reinforce that single-locus data is sufficient to flag many problems.

ARTICLE HISTORY
Received 10 June 2017
Accepted 1 December 2017

KEYWORDS
Beast; COI; GMYC; mtDNA; species delimitation

Introduction

Species are the fundamental entity in many areas of comparative biology and methodologies designed for their proper identification have a great importance in all the biological sciences. With current low pricing for the generation of DNA sequence data, the usage of molecular based criteria for species identification and delimitation is becoming routine (Carstens et al. 2013; Flot 2015). Propelled by global initiatives for DNA based identification (e.g. DNA barcode), methods for species delimitation based on a single locus have increased in number and are now widely used (Hebert et al. 2003; Schoch et al. 2012). Several authors have highlighted the difficulty and alleged inaccuracy of delimitation based on a single locus, mostly due to several natural processes that can mislead phylogenetic signal (e.g. incongruences between gene histories and species trees) (Carstens et al. 2013; Dowton et al. 2014; Hubert and Hanner 2015). Nonetheless, single locus species delimitation methods continue to be used and its impact in inferences of species diversity must be extensively evaluated.

DNA based species delimitation: a very brief history

There are many methods available with the main purpose of elaborating hypotheses of species delimitation from DNA sequence data. Each of these methods relies extensively on a number of simplifying assumptions that relate to parameter space these methods are allowed to explore (Carstens et al. 2013).

Early approaches for species delimitation include mainly methods that are non tree based (see Sites and Marshall 2003 for a review). The most notorious non tree based approach is the Population Analysis Aggregation (Davis and Nixon 1992). The PAA methodology is based in the following principle: if individuals from different populations share the same profile (genotype, morphological character), this is an indicative of conspecificity (Sites and Marshall 2003).

Derived from PAA, the character haplotype aggregation (CHA) (Brower 1999) is another very popular approach used in species delimitation. CHA is a tree based method, and uses the profiling of haplotypes in a network for discriminate species, hence shared identical haplotypes are used as indications of conspecificity.

With respect to current methods for DNA-based species delimitation, Hubert and Hanner (2015) roughly arranged these approaches into categories: phylogenetic methods, coalescent methods and phylogenetic-coalescent methods.

Phylogenetic methods rely on the branching pattern of a given tree topology (without the modelling of coalescent process within species) to define thresholds of sequence divergence to determine individual clusters which are then recognized as species (Hubert and Hanner 2015). An example of a widely used implementation of phylogenetic species

CONTACT Grazielle Gomes graziellefeg@gmail.com Instituto de Estudos Costeiros, Universidade Federal do Pará – Bragança, Pará CEP 68.600-000, Brazil
Supplemental data for this article can be accessed here.

delimitation is the Automatic Barcode Gap Discovery – ABGD (Puillandre et al. 2012). In ABGD, the primary hypothesis of species is defined after the assumption of the existence of a gap for the distance values within and between groups of individuals, after successive tests of clustering and thresholds (Puillandre et al. 2012).

With the popularization of the coalescent theory, the use of genealogy-based criteria (i.e. coalescent methods and phylogenetic-coalescent methods) for species delimitation has grown in popularity (Dowton et al. 2014).

Among methods of species delimitation that account for phylogenetic-coalescent processes, and which are based on single locus, the General Mixed Yule Coalescent – GMYC: (Pons et al. 2006) is one of the most popular algorithms. The GMYC identifies species by optimization of a point of exchange between the coalescent process and a most depth pattern of branching, characteristic of the arising of species (Pons et al. 2006), i.e. similar to the tokogenetic and the phylogenetic process of Hennig (1966). The GMYC, nevertheless is only suitable for ultrametric trees (all paths from the root to a leaf have the same length). Here, ultrametric trees are commonly obtained as consensus trees from the software that implement Bayesian searches. However, in these situations, the stochasticity present in the Markov Chain Monte Carlo search is not exploited of an adequate way (Reid and Carstens 2012). This issue is solved with methods that allows for the use of multiple trees from the posterior distribution of trees from Bayesian analyses (Reid and Carstens 2012).

To that end, the multiple GMYC (mGMYC) is an implementation that allows most credibility to GMYC including the use of multiple temporal thresholds for the differentiation of the process of shallow branching (coalescent process) and deeper branching (inter-specific diversification) in a phylogeny (see Monaghan et al. 2009).

All of the algorithms discussed above use genealogies whose branch lengths are given in time unities. However, Zhang et al. (2013) implemented the Poisson tree processes (PTP) as an alternative to GMYC. In the PTP, the diversification process of speciation is obtained considering the number of substitutions (in place of time as in GMYC) in a given branch, in principle the number of substitutions seems to be higher between species than within species (Zhang et al. 2013). In a new implementation of PTP (multi-rate Poisson tree processes, mPTP), the distinct values of intraspecific divergence caused by differences in the evolutionary process are accounted, and should confer more credibility for species delimitation (Kapli et al. 2017).

Species delimitation in snappers

Herein, we test the influence of multiple methods of species delimitation, using a single locus, in the evaluation of taxa diversity within a large group of fishes (Perciformes: Lutjanidae: Lutjaninae). To that end, we applied a selection of different discovery operations to delimit species within Lutjaninae, which represent popularly used algorithms. All of these based on explicit optimality criteria to the definition and number of species in a given clade. We compare the results of our analyses to the currently adopted taxonomy of the group, one that is largely based on morphology.

Lutjanidae, commonly known as snappers, is a widely distributed family of perciform fishes, composed of generally large bodied and mostly marine taxa (rarely estuarine) (Nelson et al. 2016). Lutjanidae is composed of 110 species recognized in 17 genera and four subfamilies: Etelinae, Apsilinae, Paradicichthyinae and Lutjaninae (Nelson et al. 2016 and their references). Lutjanidae seems an excellent model for comparative analyses of molecular methods of species delimitation, because the family is widely distributed and adults of most species in Lutjanidae are easily distinguished from other species with the use of morphology alone (Allen 1985).

Moreover, recent molecular analyses suggest that some pairs of close species may have arisen after known events of alopatric speciation, e.g. by closure of isthmus of Panama (*Lutjanus guttatus*/*L. synagris*; *L. inermis*/*O. chrysurus*; *L. argentiventris*/*L. jocu*; *L. peru*/*L. campechanus*) (Gold et al. 2011, 2015). Additionally, Gomes et al. (2008, 2012) using DNA sequences from the mitochondrial control region, showed that red snappers *L. campechanus* and *L. purpureus* are probably a single widespread species of red snapper.

Lutjanidae is a good model for comparative analyses of molecular methods of species delimitation and their congruence with the previous classification by mean of morphological criteria. Herein, we use members of the subfamily Lutjaninae inhabiting the Western Atlantic and Eastern Pacific (EP) to test the effectiveness of species delimitation methods based on a single mitochondrial locus standardized for DNA barcode for all metazoa (cytochrome C oxidase subunit I – Cox1).

Materials and methods

Sampling

We obtained sequences of 652 bp corresponding to fragment of the Cox1 for 114 specimens of Lutjaninae and three of other Lutjanidae (Appendix I). Most sequences used herein were previously published (Gold et al. 2011, 2015; Veneza et al. 2014) and were obtained from Genbank (accession numbers and metadata for samples included are as Supplementary Data). We generated new sequences for *L. buccanela* and *L.* cf. *alexandrei*, using the primers FishF1/FishR1 (Ward et al. 2005), following the protocols used in Veneza et al. (2014).

The dataset compiled for this study include 22 species, without outgroups (four species), representing nearly 90% of the nominal taxa in the Western Atlantic (WA)/EP region. Here we also include *L. notatus* because of your close relationship with *L. viridis* (EP) (Gold et al. 2015).

Species delimitation analyses and parameters

The sequence alignment method used in all phylogenetic analyses was the same, although two different datasets were compiled from the alignment: (1) full dataset; (2) reduced dataset. The full dataset alignment included all available

sequences and was obtained in ClustalW 2.1 (Larkin et al. 2007) using default parameters. To exclude redundancies in the matrix, the full dataset was trimmed to an alignment with a single individual per haplotype – this was performed in the software Alter (Glez-Peña et al. 2010). All DNA sequences used here are available in Genbank (see Appendix I).

At first, we used cladistic haplotype analysis – CHA (Brower 1999) based on a network made in Haploviewer (Salzburger et al. 2011) following the parameters recommended by the software's authors.

For inference of the level of barcode gap, we used the software ABGD (Puillandre et al. 2012), using the following settings: (Pmin = 0.001, Pmax = 0.1, Steps = 10, NBins = 20, Relative gap = 1, JC69).

The guide trees used for GMYC and bGMYC analyses were obtained in Beast v. 2.3.1 (Bouckaert et al. 2014) using a strict clock and empirical frequency base priors. We performed multiple runs using different random seeds. Each run was performed with 10^7 steps and topologies were sampled each 1000 generations. We discarded 10% of the genealogies as burn-in. The choice of the evolutionary model that best fit the data was performed in Kakusan v. 4 (Tanabe 2007). Convergence was verified in TRACER v. 1.6 (available at: http://beast.bio.ed.ac.uk/tracer). All ESS values remained above 200 for GMYC/bGMYC. Moreover, we also tested the influence of different tree priors ('Coalescent', 'Yule' and 'Birth Death') on species delimitation scenarios.

The GMYC with a single threshold (sGMYC) (Pons et al. 2006) and multiple thresholds (mGMYC) (Monaghan et al. 2009) were conducted in the Splits package, available in the software R v. 3.2 (R Core Team 2012). For both approaches (sGMYC and mGMYC), we used a tree summarized from the Beast analyses. The function 'multi2di' was used in order to avoid branch-lengths of zero.

The bGMYC allows for the use of a set of trees in order to consider the stochasticity in the posterior distribution. Thus, for bGMYC we sampled a set of 100 trees of posterior distribution of the Beast runs, using the script BurnTrees (J. Nylander; https://github.com/nylander/Burntrees). The bGMYC was also performed in R, using the bGMYC package (Reid and Carstens 2012), following the settings recommended by the authors.

For species delimitation via PTP and mPTP, we used a phylogenetic tree of maximum likelihood made in the RaxML v. 8.1 (Stamatakis 2014), using the fast method of bootstrap (100 replications). The analyses were conducted in the PTP and mPTP web servers (http://species.h-its.org/and http://mptp.h-its.org/#/tree).

Results

The phylogeny of Lutjaninae

Final datasets included 600 bp for 117 individuals (full database) and 67 individuals (database collapsed). Both maximum likelihood and Bayesian inference topologies consistently recover *Lutjanus* to be paraphyletic with respect to *Ocyurus chrysurus* and *Rhomboplites aurorubens*, as shown previously (Gold et al. 2011; Frédérich and Santini 2017).

Species delimitation analyses

With respect to species delimitation with distance methods, the ABGD analysis recovered 18 putative species (Figure 1). After recursive partitions of the threshold values in the analyses, the number of groups varied between 18 and 46 putative species (Appendix II).

For CHA and phylogenetic-coalescent based methods, the number of species within Lutjaninae ranged from 16 to 25 (Figure 1, Table 1 and Appendix III), whereas we observed a higher number of putative species for methods with multiple thresholds. However, when we compared likelihood scores of single threshold methods with multiple thresholds, we found better scores of probabilities favouring methods with a single threshold (Table 1), which suggests a better performance of these approaches over multiple thresholds algorithms. With respect to the performance of GMYC with data sets collapsed to haplotypes and full data sets, only discrepancies for GMYC with multiple thresholds were observed (Table 1).

Their little difference between species delimitation related to tree priors (Yule, Coalescent and Birth Death) for the GMYC with a single threshold recovered the same 22 lineages. However, we observed differences in the bGMYC analyses, which recovered 24 lineages.

Discussion

Even with the knowledge of the limitations caused by stochasticity of the coalescent process, species delimitation methods based on DNA sequence data from one locus are still frequently used – special reference is given to identification of species and specimens by means of DNA barcode (e.g. Fujisawa and Barraclough 2013; Hubert and Hanner 2015). Our results suggest that the resulting species delimitation of snappers using a single locus varies widely.

In the present study, the ABGD was able to distinguish 18 groups, whose threshold for intraspecific distance was nearly 2%. Methods of species delimitation were based only on nucleotide divergence and do not include any information on evolutionary process (see Fujisawa and Barraclough 2013) – this approach has many flaws. Therefore, it is wise to support species delimitation methods that have an evolutionary basis (Yang and Rannala 2016).

In regard to phylogenetic-coalescent methods, the lower numbers of putative species recovered were obtained with PTP and mPTP (Figure 1). In other studies, these methods also lead to a lower number of species in comparison with other methods, such as GMYC (Aguila et al. 2017; Correa et al. 2017). In some of these cases, this fact is superficially treated as over splitting of lineages occurring in GMYC. However, the over splitting suggested for GMYC, and another methods, does not seem to be a likely scenario here.

The ABGD (with the initial partition) and the mPTP were unable to distinguish various pairs of sister species that are easily recognizable from morphology and which were likely isolated since the closure of the Isthmus of Panama (Lessios 2008) (*L. jocu* × *L. argentiventris*; *L. purpureus* × *L. campechanus* × *L. peru*; *L. novemfasciatus* × *L. cyanopterus*; *L. synagris* × *L. guttatus*). Even with maximum number of

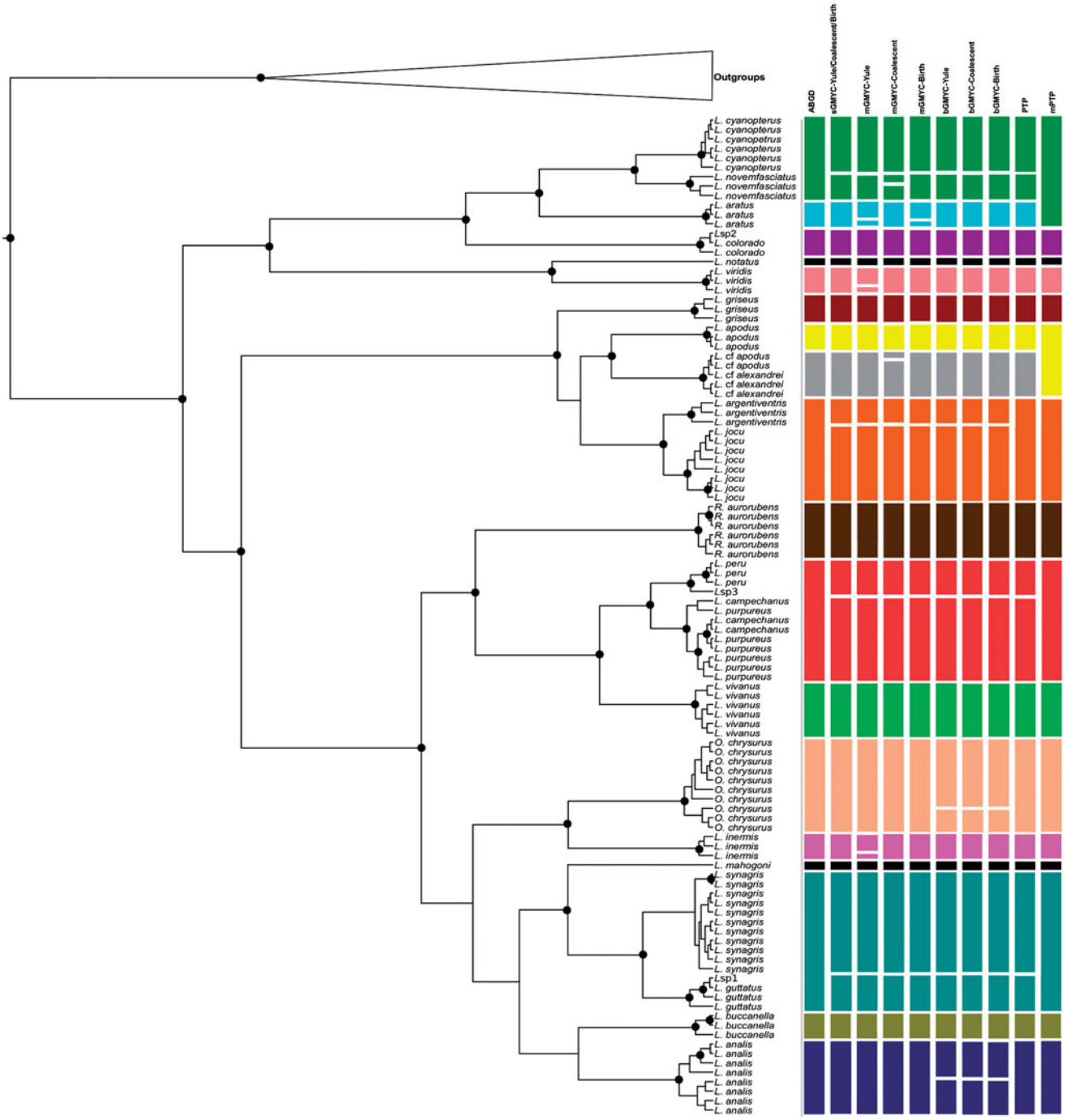

Figure 1. Ultrametric tree showing different scenarios of species delimitation is encountered here. Vertical bars correspond to each lineage own as a potential species. Dots above branches represent values of posterior probability (>0.95). Generic name abbreviations are as follows: *L.* (*Lutjanus*), *O.* (*Ocyurus*), R. (*Rhomboplites*), E. (*Etelis*), A. (*Apsilus*), and P. (*Pristipomoides*). ABGD 1, ABGD 2, ABGD 3, represent different thresholds for intraspecific divergence.

recursive partitions (see Figure 1), ABGD was not able to distinguish between the pair of species *L. jocu* and *L. argentiventris*.

The species listed above are isolated by a minimum of 3 millions of years (Gold et al. 2011) and present several diagnostic phenotypic characters, which support recognition of them as distinct taxa. Marine fishes have higher effective population size (McCusker and Bentzen 2010), therefore in loci whose evolution is driven by neutral process, the time necessary for establishment of larger values of nucleotide divergence or even lineage sorting can be much higher than a few millions of years (Yang and Rannala 2016). This might be one of the causes of under-splitting observed in some of our analyses.

In many of the tree based on species delimitation, species are distinguished by means of the optimization of thresholds between coalescent process, i.e. short branches typical of intraspecific diversification and deep branching, typical of speciation process (Pons et al. 2006; Esselstyn et al. 2012). Therefore, methods such as the GMYC (and related methods) are largely rooted on the phylogenetic species concept (Zhang et al. 2013), which support the notion that species should be unambiguously monophyletic. On the other hand, even for tree-based methods, the use of operations for

Table 1. Dataset, implementation/tree priors, number of entities, and log likelihoods for GMYC analyses (sGMYC + mGMYC) of the present study.

Data set	Method/tree prior	Number of entities	Log likelihood
Full	sGMYC/Yule	22	−950.8829
Full	sGMYC/birth-death	22	−993.3962
Full	sGMYC/coalescent	22	−951.5797
Full	mGMYC/Yule	25	−952.3841
Full	mGMYC/birth-death	23	−994.3212
Full	mGMYC/coalescent	24	−951.6738
Collapsed	sGMYC/Yule	22	−429.7938
Collapsed	sGMYC/birth-death	22	−436.8823
Collapsed	sGMYC/coalescent	22	−436.7826
Collapsed	mGMYC/Yule	24	−431.2117
Collapsed	mGMYC/birth-death	21	−437.6848
Collapsed	mGMYC/coalescent	21	−437.2014

Number of entities does not include outgroup taxa.

species delimitation based on a unique locus has strong theoretical limitations. Special attention should be given to inference of false positives caused by incomplete lineage sorting or differences between gene trees and species tree. Nonetheless, approaches for species delineation with one locus can be very helpful in many situations, for example, in 'turbo-taxonomy' (Butcher et al. 2012; Dentinger and Suz 2014), or for next-generation environmental barcoding (Hajibabaei et al. 2011).

Other issues are also mentioned as causes of incongruence in tree-based methods for species delimitation. For example, the use of trees with branches of zero length has been mentioned as a relevant point for the correct identification of the transition point between coalescent and speciation processes (Tang et al. 2014). The process of the intraspecific diversification process, also can lead to phylogenetic arrangements with long branches and then these populations should be identified as potential species (Esselstyn et al. 2012).

The choice of tree priors for inferring the phylogenetic tree is another aspect relevant for the species delimitation process. As an example, under a Yule process prior there is an unrealistic assumption in regard to arising and extinction of lineages (i.e. Yule process assumes only the birth of a lineage, and the rate of extinction is zero) (Esselstyn et al. 2012). In this way, tree priors such as Birth-Death or a coalescent prior, seem to provide more adequate assumptions with regards to evolutionary history (Ceccarelli et al. 2012). Here, however, we did not observe differences between species delimitation in relation to tree priors for GMYC with a single threshold (Yule, Coalescent and Birth Death) and CHA, with all these methods converging in resolving 22 lineages.

However, for all bGMYC, we found a number of 24 lineages with a splitting within *L. analis* and *O. chrysurus*. Curiously, *L. analis* seems to exhibit a pattern of potential speciation between Caribbean and south Atlantic (Brazilian) individuals. Some studies have suggested restriction of gene flow between Caribbean and Brazilian localities, including *O. chrysurus* (Vasconcellos et al. 2008; da Silva et al. 2015) and *L. synagris* (Silva et al. 2018), however, no information is available in regard to levels of genetic connectivity between Brazil and Caribbean for *L. analis*.

We did not observe discrepancies between data sets collapsed and species delimitation scenarios with the full data set (i.e. including branches with length of zero), in accordance with (Talavera et al. 2013), which punctuate as the main factor for choice between full data sets and data set collapsed, seem to be the lower computational effort. The sGMYC and CHA provide species delimitation scenarios that are very similar to the currently accepted taxonomy of Lutjaninae. On the other hand, the GMYC with multiple thresholds leads to a higher number of molecular entities than the currently accepted method for the formal taxonomy of Lutjaninae. Moreover, methods like mGMYC allow many temporal thresholds for speciation, and therefore are most sensible to deep coalescent process at population level (Esselstyn et al. 2012), this can lead to overestimation of the species number.

Another aspect inherent to molecular species delimitation with one locus refers to the low accuracy of these algorithms in species with larger effective population sizes (Esselstyn et al. 2012), a very common scenario in marine teleosts (McCusker and Bentzen 2010). Therefore, in such situations, species delimitation with a single locus has been extensively criticized on many fronts, but these approaches can undoubtedly still be used as valid method for species delimitation in many taxa.

In the present study, our dataset is limited, both in number of individuals and number of loci, and so our results should be evaluated carefully. Nonetheless, the congruence among most methods (often based on widely discrepant assumptions) suggests that there is some consistency of inferred species limits, despite obvious philosophical and operational differences among the implementations (Carstens et al. 2013). We do not, however, endorse the use of single loci approaches as the best practice in these kinds of studies. Merely, we reinforce that single locus data are sufficient to solve many systematic problems, including within Lutjanidae (see also Veneza et al. 2014). If method of analysis is robust and adequately chosen to fit the data at hand, when resources are available (e.g. samples, facilities, budget) – we do, however, reinforce that whenever possible more loci should be included.

Acknowledgements

An earlier version of this manuscript was presented as a class seminar at Museu Paraense Emílio Goeldi. We are grateful to all colleagues for ideas and discussions that improved our work.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

P.L.V.P. was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/BJT 313680/2014-0, 400252/2014-77).

ORCID

Marcelo José Sturaro http://orcid.org/0000-0002-3157-2488

References

Aguila RN, Genaro JA, Peréz-Asso ARHA. 2017. A new species of the *Hysius* species-group of *Calisto* Hübner (Lepidoptera, Nymphalidae,

Satyrinae) and insights into the status of different populations currently attributed to *C. grannus* Bates. Zootaxa. 4317:001–044.

Allen G-R. 1985. Snappers of the world. An annotated and illustrated catalogue of Lutjanid species known to date. Rome: FAO.

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu C, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol. 10:1–6.

Brower AVZ. 1999. Delimitation of phylogenetic species with DNA sequences: a critique of Davis and Nixon's population aggregation analysis. Syst Biol. 48:199–213.

Butcher BA, Smith MA, Sharkey MJ, Quicke DL. 2012. A turbo-taxonomic study of Thai Aleiodes (*Aleiodes*) and *Aleiodes* (*Arcaleiodes*) (Hymenoptera: Braconidae: Rogadinae) based largely on COI barcoded specimens, with rapid descriptions of 179 new species. Zootaxa. 3457:1–232.

Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. Mol Ecol. 22:4369–4383.

Ceccarelli FS, Sharkey MJ, Zaldívar-Riverón A. 2012. Species identification in the taxonomically neglected, highly diverse, neotropical parasitoid wasp genus *Notiospathius* (Braconidae: Doryctinae) based on an integrative molecular and morphological approach. Mol Phylogenet Evol. 62:485–495.

Correa C, Vásquez D, Castro-Carrasco C, Zúñiga-Reinoso Á, Ortiz JC, Palma RE. 2017. Species delimitation in frogs from South American temperate forests: the case of *Eupsophus*, a taxonomically complex genus with high phenotypic variation. PLoS ONE. 12:e0181026.

da Silva R, Veneza I, Sampaio I, Araripe J, Schneider H, Gomes G. 2015. High levels of genetic connectivity among populations of yellowtail snapper, *Ocyurus chrysurus* (Lutjanidae – Perciformes), in the Western South Atlantic revealed through multilocus analysis. PLoS One. 10:e0122173.

Davis JI, Nixon KC. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. Syst Biol. 41:421–435.

Dentinger BTM, Suz LM. 2014. What's for dinner? Undescribed species of porcini in a commercial packet. Peer J. 2:e570.

Dowton M, Meiklejohn K, Cameron SL, Wallman J. 2014. A preliminary framework for DNA barcoding, incorporating the multispecies coalescent. Syst Biol. 63:639–644.

Esselstyn JA, Evans BJ, Sedlock JL, Ali F, Khan A, Heaney LR. 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 Lond B Biol Sci. 279:3678–3686.

Flot J-F. 2015. Species delimitation's coming of age. Syst Biol. 64:897–899.

Frédérich B, Santini F. 2017. Macroevolutionary analysis of the tempo of diversification in snappers and fusiliers (Percomorpha: Lutjanidae). Belg J Zool. 147:17–35.

Fujisawa T, Barraclough TG. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. Syst Biol. 44:1–49.

Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D. 2010. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res. 38:14–18.

Gold JR, Voelker G, Renshaw MA. 2011. Phylogenetic relationships of tropical western Atlantic snappers in subfamily Lutjaninae (Lutjanidae: Perciformes) inferred from mitochondrial DNA sequences. Biol J Linn Soc. 102:915–929.

Gold JR, Willis SC, Renshaw MA, Buentello A, Puritz JB, Hollenbeck CM, Voelker G. 2015. Phylogenetic relationships of tropical eastern Pacific snappers (Lutjanidae) inferred from mitochondrial DNA sequences. Syst Biodivers. 13:1–12.

Gomes G, Sampaio I, Schneider H. 2012. Population structure of *Lutjanus purpureus* (Lutjanidae – Perciformes) on the Brazilian coast: further existence evidence of a single species of red snapper in the western Atlantic. An Acad Bras Cienc. 84:979–999.

Gomes G, Schneider H, Vallinoto M, Santos S, Orti G, Sampaio I. 2008. Can *Lutjanus purpureus* (South red snapper) be "legally" considered a red snapper (*Lutjanus campechanus*)? Genet Mol Biol. 31:372–376.

Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ. 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One. 6:e17497.

Hebert PDN, Ratnasingham S, DeWaard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B Biol Sci. 270:96–99.

Hennig W. 1966. Phylogenetic systematics. Urbana: University of Illinois Press.

Hubert N, Hanner R. 2015. DNA barcoding, species delineation and taxonomy: a historical perspective. DNA Barcodes. 3:44–58.

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.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23:2947–2948.

Lessios HA. 2008. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. Annu Rev Ecol Evol Syst. 39:63–91.

McCusker MR, Bentzen P. 2010. Positive relationships between genetic diversity and abundance in fishes. Mol Ecol. 19:4852–4862.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol. 58:298–311.

Nelson JS, Grande TC, Wilson MV. 2016. Fishes of the world. 5th ed. Hoboken (NJ): John Wiley & Sons.

Pons J, Barraclough T, Gomez-Zurita J, Cardoso A, Duran DP, Hazzel S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst Biol. 55:595–609.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol. 21:1864–1877.

R Core Team 2012. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Reid NM, Carstens BC. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol Biol. 12:1471–2148.

Salzburger W, Ewing GB, Von Haeseler A. 2011. The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. Mol Ecol. 20:1952–1963.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Chen W, Consortium FB. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci USA. 109:6241–6246.

Silva D, Martins K, Oliveira J, da Silva R, Sampaio I, Schneider H, Gomes G. 2018. Genetic differentiation in populations of lane snapper (*Lutjanus synagris* – Lutjanidae) from Western Atlantic as revealed by multilocus analysis. Fish Res. 198:138–149.

Sites JW, Jr, Marshall JC. 2003. Delimiting species: a renaissance issue in systematic biology. Trends Ecol Evol. 18:462–470.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Talavera G, Dinca V, Vila R. 2013. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. Methods Ecol Evol. 4:1101–1110.

Tanabe AS. 2007. kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. Mol Ecol Notes. 7:962–964.

Tang CQ, Humphreys AM, Fontaneto D, Barraclough TG. 2014. Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. Methods Ecol Evol. 5:1086–1094.

Vasconcellos AV, Vianna P, Paiva PC, Schama R, Solé-Cava A. 2008. Genetic and morphometric differences between yellowtail snapper (*Ocyurus chrysurus*, Lutjanidae) populations of the tropical West Atlantic. Genet Mol Biol. 31:308–316.

Veneza I, Felipe B, Oliveira J, Silva R, Sampaio I, Schneider H, Gomes G. 2014. A barcode for the authentication of the snappers (Lutjanidae) of the western Atlantic: rDNA 5S or mitochondrial COI? Food Control. 38:116–123.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci. 360:1847–1857.

Yang Z, Rannala B. 2016. Species identification by Bayesian fingerprinting: a powerful alternative to DNA barcoding. bioRxiv. 2016:041608.
Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. 29:2869–2876.

Appendix I. Code of access and reference for all sequences utilized in the present study.

*Out group.

**Nomenclature used in Gold et al. (2011).

Appendix II. Number of groups and prior intraspecific divergence (*p*) for ABGD.

Appendix III. Network showing the lineages observed in CHA.