Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 40 (2006) 900–902
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Letter to the editor

Sicyopterus lagocephalus: Widespread species, species complex, or neither? A critique on the use of molecular data for species identification

In a recent study, Keith et al. (2005) analyzed partial cytochrome *b* sequences for 55 terminals belonging to the Indo-Pacific sicydiine (rock-climbing) goby genus *Sicyopterus* and concluded that the specimens they refer to as "*Sicyopterus lagocephalus*" formed a single widespread species that "occurs throughout 18,000 km from the west of the Indian Ocean... to the east of the Pacific Ocean." They based their "single widespread species" hypothesis on the monophyly of an assemblage comprising the specimens they referred to as "*Sicyopterus lagocephalus*" and a cursory discussion of pairwise sequence divergences. What Keith et al. fail to explain (or even address) is how they justify such a conclusion in light of the evidence presented and what hypothesis they have actually tested. The stated goal of their study was to analyze "representative taxa of the genus *Sicyopterus*, and specially [sic] the ubiquity of the species *S. lagocephalus*, in order to determine whether its huge range is really occupied by only one species, and to describe its possible genetic subdivisions." Unfortunately, they failed to test whether or not this particular clade of *Sicyopterus* comprises a single widespread species, a species complex, or neither. We will briefly discuss each of these possibilities in light of the evidence presented by Keith et al. (2005) and previous studies of *Sicyopterus*, and conclude with a critique on the use of molecular methods for taxonomic purposes.

Single widespread species. By the authors' own account, the conclusion that their "*Sicyopterus lagocephalus* clade" comprised but a single species was only justified by "the tree topology together with the small genetic distances measured between haplotypes." Given that species are historical individuals, it is, therefore, unnecessary for species to be monophyletic (e.g., Skinner, 2004). As a result, the monophyly of this assemblage provides no evidence with regard to species recognition, although it may serve as a heuristic toward species delimitation. With regard to the "small" pairwise sequence divergences for partial cytochrome *b* sequences that the authors present as evidence for a single species, Keith et al. fail to address the critical questions: can pairwise sequence divergences be used to diagnose species and, if they can diagnose species, what amount of cytochrome *b* sequence divergence is required for species recognition? With regard to the first question, we maintain that species diagnoses must rely on evidence that is in the form of character-based apomorphies (= character-state transformations). Pairwise divergences are necessarily distance-based, not character-based, so they are incapable of identifying the character-state transformations necessary for diagnosing any clade or terminal. With regard to the second question, we do not believe that there is a single immutable pairwise divergence value that is diagnostic for species recognition, particularly given wide variation in the rates of molecular evolution among lineages (even those closely related), not to mention issues of taxon sampling density. Nevertheless, the burden of proof falls on the authors, who fail to provide an answer.

Species complex. Prior to Watson et al. (2000), *Sicyopterus lagocephalus* was restricted to islands of the western Indian Ocean (see below, and Sparks and Nelson, 2004 for additional commentary on the validity of this taxon). Watson et al. (2000) placed 12 species into the synonymy of *S. lagocephalus*, greatly expanding the range of this species to include Pacific forms. It is clear that there is substantial hierarchical structure evident for "*S. lagocephalus*" in the cladograms presented by Keith et al. (2005) (their Figs. 2 and 3); such structure is suggestive of reproductive isolation and a lack of gene flow. Furthermore, their own molecular clock estimates suggest that some "genetic subdivisions" of *S. lagocephalus* have been reproductively separated longer than many other *Sicyopterus* species that they consider to be valid (their Fig. 3). From the evidence (= topologies) presented, both the French Polynesian and Mascarene-Comoros clades of "*S. lagocephalus*" are relatively old (their Fig. 3). Indeed, the French Polynesian clade of "*S. lagocephalus*" is estimated to be older than *S. rapa* and *S. marquesensis*, both of which they consider to be valid species; however, this problematic and conflicting result is not discussed. Moreover, a maximal "intraspecific" divergence of 2.6% is reported for "*S. lagocephalus*" between New Caledonian and Mauritian samples; whereas, interspecific divergences of 3.7% are reported between both *S. marquesensis* and "*S. lagocephalus*", and *S. marquesensis* and *S. aiensis* (Keith et al., 2005). Yet, there is no accompanying discussion as to why 3.7% divergence in partial cytochrome *b* sequences implies distinct species, whereas 2.6% does not. In short, the authors' decision to recognize distinct species in the former (3.7% divergence) and a single species in the latter (2.6% divergence) is entirely arbitrary

1055-7903/$ - see front matter
doi:10.1016/j.ympev.2006.05.016

and unsupported by the divergence time estimates presented.

Neither. Although largely ignored by Keith et al. (2005), the availability of the name *Sicyopterus lagocephalus* has been refuted by Sparks and Nelson (2004), who argued that no Indo-Pacific sicydiine goby could properly be referred to as *Sicyopterus lagocephalus*. Briefly, the type specimen of *Gobius lagocephalus* Pallas, 1770 is lost, the original description and illustration lack diagnostic features, and the type was collected from "American Seas." Given that *Sicyopterus* is restricted to the Indo-Pacific and does not occur in American seas, and that Pallas' description and illustration can only be accurately classified to the subfamilial level, Sparks and Nelson (2004) contended that *Gobius lagocephalus* must be treated as a *nomen dubium*. Therefore, despite Watson et al.'s (2000) claim to the contrary, the name is not available for any Indo-Pacific *Sicyopterus* species.

We believe that the primary reason that Keith et al. failed to achieve their goal of testing the ubiquity of their "*Sicyopterus lagocephalus* clade" is because they uncritically accepted Watson et al.'s (2000) "determination" of the species limits of *S. lagocephalus* a priori as fact or background knowledge (Keith et al., 2005, 722). We stress that this acceptance was solely an appeal to authority or faith, in that Watson et al. (2000) failed to provide any apomorphic features (genotypic or phenotypic) or a unique combination of apomorphies that, in conjunction with a cladogram, would allow either Watson et al. (2000) or Keith et al. (2005) to conclude that what they had in their hand could be assigned to *S. lagocephalus* (see Sparks and Nelson, 2004). Perhaps this might explain why one of the putative *S. lagocephalus* specimens examined in the Keith et al. study (SIC840, AY940724) is instead identified on GenBank as a *Cotylopus*. We maintain that since Keith et al. assumed *a priori* that Watson et al.'s (2000) determination for species recognition was correct (i.e., not too narrowly or broadly restricted), and they had concluded prior to their own analysis that the specimens in question all belonged in *S. lagocephalus*, then they failed to test Watson et al.'s (2000) hypothesis. Instead, the appropriate test, especially given the significant hierarchical structure evident in Keith et al.'s cladograms (their Figs. 2 and 3), would have been to use these phylogenies to guide a comparative morphological examination of their voucher specimens or to target additional molecular studies to look for hidden diversity within this widespread clade.

The Keith et al. (2005) test was therefore weak, and the only potential falsification of Watson et al.'s (2000) hypothesis of the limits of *S. lagocephalus* (which was similarly unjustified [Sparks and Nelson, 2004]) would have been non-monophyly of their "*S. lagocephalus* clade." Since Keith et al.'s "*S. lagocephalus*" exemplars formed a clade, Watson et al.'s hypothesis of species limits within this monophyletic group was not tested any further. Keith et al. simply concluded that their "*S. lagocephalus* clade" comprised a single species using the nearly tautologous argument that *S. lagocephalus*, identified by Watson et al.'s (2000) morphometric and meristic determination, is *S. lagocephalus sensu* Watson et al. (2000). An equally tenable result of their study could have been that all *Sicyopterus* form a single species because they too are monophyletic; a line of reasoning that could be extrapolated out as far as one would like to take it. This example illustrates why a rational and defensible justification for species limits cannot be based on a topology alone (see also Davis and Nixon, 1992; Brower, 1999; DeSalle et al., 2005).

The use of molecular methods for taxonomic purposes. Most taxonomic studies still employ traditional comparative anatomical approaches, but there is a growing body of literature focusing on the integration of molecular data into taxonomy. A charge for the use of sequence data in species identification is being spearheaded by the Barcode of Life initiative (see October 2005 *Philosophical Transactions of the Royal Society*, *London*: *Biological Sciences* 360:1462 for the proceedings of the First International Barcode Conference). It is obvious from this symposium volume and the studies cited within, that the methods for DNA taxonomy alone, or its integration into taxonomy as a whole, are still a topic of considerable debate. Most studies focusing on barcoding as a means of species recognition have relied on distance-based methods (e.g., Hebert et al., 2003); however, such an approach is at odds with traditional character-based phenotypic taxonomy (DeSalle et al., 2005). Although we agree with DeSalle et al. (2005) that character-based phylogenetic methods are more appropriate than distance-based methods for establishing so-called DNA barcodes or addressing alpha-taxonomic problems, we believe that a combination of character-based phylogenies and non-tree based concepts or methods (e.g., population aggregation analysis [Davis and Nixon, 1992], cladistic haplotype aggregation [Brower, 1999]) should be employed to examine all available evidence for determining species limits if genotypic data alone is used.

In their justification for not using tree-based methods for their DNA "barcode engine," DeSalle et al. (2005) noted that species cannot be delimited by cladograms alone. First, any number of natural events (e.g., lineage sorting, hybridization, introgression) can result in the non-monophyly of a species (Avise, 2000). This has no bearing, however, on the historical individuality of a species (e.g., Frost and Kluge, 1994), but it does limit the effectiveness of cladograms for this purpose. Second, phylogenetic algorithms assume that there are genealogical relationships among the included terminals; thus, a hierarchy will be forced on terminals that are related only tokogenetically (Davis and Nixon, 1992; Brower, 1999). Specifically, a cladogram is inadequate for delimiting species precisely because it does not provide a defensible or objective method for separating intraspecific tokogenetic structure from interspecific phylogenetic structure. Despite these caveats, phylogenetic analysis can certainly aid in species delimitation because it facilitates the identification of the apomorphic characters that are fundamental to traditional species diagnosis.

If DNA taxonomy is to be integrated into traditional phenotypic taxonomy, we require a method that allows us to identify apomorphic characters for species diagnosis; such an approach precludes distance-based methods or pairwise sequence divergence values from having evidentiary value in species delineation. If Keith et al. (2005) had instead decided to look for apomorphic base pairs to diagnose species, they would have noted that their Mascarene-Comoros clade of "*S. lagocephalus*" was diagnosed by unique nucleotides (among examined taxa) at three positions (Adenine at position 213 [A-213], T-240, and A-288) and that their French Polynesian clade of "*S. lagocephalus*" was diagnosed by unique nucleotides at two positions (T-366 and A-540). According to many species concepts (e.g., the phylogenetic species concept), this would be sufficient evidence to recognize their *Sicyopterus lagocephalus* as a species complex, rather than a single widespread species. Furthermore, this species-complex view of their *S. lagocephalus* is equally consistent with their cladogram, which retained significant hierarchical structure, including the identification of Mascarene-Comoros and French Polynesian clades. Clearly, additional study that incorporates new methods being developed for character-based DNA taxonomy with additional data (morphological or molecular) is needed to resolve the species limits within *Sicyopterus*.

Keith et al. (2005), like Watson et al. (2000) before them, failed to sufficiently test species boundaries within an assemblage of rock-climbing gobies they referred to as "*Sicyopterus lagocephalus*." This failure stems from their inability to identify, or refusal to enumerate, a single unique morphological or molecular feature (= transformation) that diagnoses this clade, and which could be construed as providing evidence for a single widespread species. Whereas we sympathize with these authors in their attempt to work out species boundaries in a morphologically conservative (and challenging) clade, in both the study of Watson et al. (2000) and Keith et al. (2005), in which character-based approaches to species delimitation were not adopted, the "evidence" for species boundaries was based on authoritative conjecture and was not corroborated by apomorphy. Ultimately, the shortcomings of the Keith et al. (2005) study were irrelevant to the question we posed in the title of this paper. In fact, the answer is quite simple: *Sicyopterus lagocephalus* is neither a widespread species nor a species complex; it is a *nomen dubium*.

References

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

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

Davis, J.I., Nixon, K.C., 1992. Populations, genetic variation, and the delimitation of phylogenetic species. Syst. Biol. 41, 421–435.

DeSalle, R., Egan, M.G., Siddall, M., 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos. Trans. R. Soc. Lond. B 360, 1905–1916.

Frost, D.R., Kluge, A.G., 1994. A consideration of epistemology in systematic biology, with a special reference to species. Cladistics 10, 259–294.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B 270, 313–321.

Keith, P., Galewski, T., Cattaneo-Berrebi, G., Hoareau, T., Berrebi, P., 2005. Ubiquity of *Sicyopterus lagocephalus* (Teleostei: Gobioidei) and phylogeography of the genus *Sicyopterus* in the Indo-Pacific area inferred from mitochondrial cytochrome *b* gene. Mol. Phylogenet. Evol. 37, 721–732.

Skinner, A., 2004. Hierarchy and monophyly. Cladistics 20, 498–500.

Sparks, J.S., Nelson, D.W., 2004. Review of the Malagasy sicydiine gobies (Teleostei: Gobiidae), with description of a new species and comments on the taxonomic status of *Gobius lagocephalus* Pallas, 1770. Amer. Mus. Novitates 3440, 1–20.

Watson, R.E., Marquet, G., Pöllabauer, C., 2000. New Caledonia fish species of the genus *Sicyopterus* (Teleostei: Gobioidei: Sicydiinae). Aqua 4, 5–34.

John S. Sparks
William Leo Smith *
Department of Ichthyology,
Division of Vertebrate Zoology,
American Museum of Natural History,
Central Park West at 79th Street,
New York, NY 10024, USA
E-mail address: jsparks@amnh.org (J.S. Sparks),
leosmith@amnh.org (W.L. Smith).

Received 26 October 2005; revised 8 May 2006;
accepted 16 May 2006
Available online 22 May 2006

* Corresponding author. Fax: +1 212 769 5642.