CORRESPONDENCE

Comment on Falade *et al*. (2016) DNA-barcoding of *Clarias gariepinus*, *Coptedon zillii* and *Sarotherodon melanotheron* from Southwestern Nigeria [version 1; peer review: 1 approved, 1 approved with reservations]

Jeremy A. Miller[1], Isolde van Riemsdijk[1,2], Mohd Zacaery Khalik[1,2], David J. Scager[1,2], Menno Schilthuizen[1,2]

[1]Naturalis Biodiversity Center, Leiden, The Netherlands
[2]Institute Biology Leiden, Leiden University, Leiden, The Netherlands

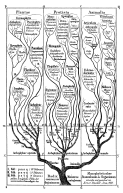
v1 **First published:** 09 Nov 2016, **5**:2654 (https://doi.org/10.12688/f1000research.9829.1)
Latest published: 09 Nov 2016, **5**:2654 (https://doi.org/10.12688/f1000research.9829.1)

Abstract

A publication by Falade *et al*. was selected for discussion by a Naturalis Biodiversity Center-Leiden University Journal Club. The study focused on the identification of fish from Southwestern Nigeria using a DNA barcoding approach. Questions raised during the discussion led to a reanalysis and reinterpretation of the data presented. The authors characterize the process of deriving a taxonomic identification from their sequence data as straightforward, but we were concerned that their approach made it nearly impossible to fail to obtain a taxonomic name for each sequence. The value of sophisticated DNA taxonomy, as well as the pitfalls of its naïve application, are discussed. We suggest that journal discussion groups may be an untapped resource for expanding rigorous peer review, particularly for journals that have adopted an open review model.

Keywords

Barcode , DNA taxonomy , FISH-BOL , peer review , scientific publishing



This article is included in the Phylogenetics collection.

Open Peer Review

Reviewer Status ? ✓

	Invited Reviewers	
	1	**2**
version 1 published 09 Nov 2016	? report	✓ report

1 **Shadi Shokralla**, University of Guelph, Guelph, Canada

2 **Dora Davies**, Universidad Nacional de Salta, Salta, Argentina
Carolina Davies, Universidad Nacional de Salta, Salta, Argentina

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Menno Schilthuizen (menno.schilthuizen@naturalis.nl)

Competing interests: No competing interests were disclosed.

Grant information: The Endless Forms Research Group (Naturalis) budget footed the bill for the journal club drinks at Meneer Jansen in Leiden.
lvR is supported by the 'Nederlandse organisatie voor Wetenschappelijk Onderzoek' (NWO Open Programme 824.14.014).

How to cite this article: Miller JA, van Riemsdijk I, Khalik MZ *et al.* **Comment on Falade *et al*. (2016) DNA-barcoding of *Clarias gariepinus*, *Coptedon zillii* and *Sarotherodon melanotheron* from Southwestern Nigeria [version 1; peer review: 1 approved, 1 approved with reservations]** F1000Research 2016, **5**:2654 (https://doi.org/10.12688/f1000research.9829.1)

First published: 09 Nov 2016, **5**:2654 (https://doi.org/10.12688/f1000research.9829.1)

DNA sequence data has become widely accepted as a useful tool for taxonomic determination and discovery[1–3]. But the potential pitfalls of DNA taxonomy in operation have been forewarned for some time[4–10].

The DNA barcode itself is simply a standard region selected to facilitate comparison[11]. A library built of many such sequences and based on a gene evolving at a rate that minimizes variation within and maximizes variation between species becomes a powerful taxonomic resource[5]. But the journey from DNA barcode sequence to species determination still requires critical application, particularly when applied to taxa or regions that are not currently well represented in sequence databases.

Falade *et al.*[12] obtained DNA sequences for sixteen individual fish from Southwest Nigeria, a region with relatively sparse coverage in sequence databases. Such data are valuable because broad geographic and taxonomic representation provide insight into genetic diversity within taxonomic groups and help us to refine hypotheses of species circumscription and phylogenetic relationships.

Falade *et al.*,[12] sequenced each specimen for the standard animal DNA barcode region cytochrome oxidase I (COI) and a region of the 16S mitochondrial ribosome. The authors queried their sequences against both the BOLD Systems (RRID: SCR_004278; boldsystems.org/index.php/IDS_OpenIdEngine) and NCBI GenBank (RRID: SCR_004860; BLASTN, RRID: SCR_001598; blast.ncbi.nlm.nih.gov/Blast.cgi) databases (because BOLD does not include 16S, these sequences were only compared to GenBank). Although the authors claim that "this resulted in straightforward identification", we take a more nuanced view on their results.

The BOLD identification engine and BLASTN comparison with GenBank work differently and were created for different purposes[13–15]; only BOLD is specifically intended to be used as a taxonomic identification tool, while BLASTN assesses sequence similarity. BLASTN will always return the most similar sequences in GenBank. BOLD is more discriminating, since it is limited to a handful of specific loci and uses similarity thresholds to assess whether or not a query sequence can be matched to identified sequences in the database with high confidence. BOLD will alert the user when it determines that no confident identification could be made. DNA-based identification is complicated by the fact that both BOLD and GenBank include misidentified sequences[16].

BOLD failed to identify with confidence any of the sixteen COI sequences. Eight were classified as probably belonging to one of a handful of possible species, while the rest received no hit. From this, we infer that Falade *et al.* made their taxonomic determinations based almost entirely on BLASTN results. As reported (Table 1), all but one of these were scored as 98–99% identical to their top GenBank hit with the remaining sequence (KX231778; Coptodon_zilli_odooba_1) scoring 86% identical.

To view the results in context, we downloaded from BOLD all COI sequences identified as one of the three species specified by Falade *et al.* [search 'Taxonomy' for *Clarias gariepinus*, *Sarotherodon melanotheron*, and *Coptodon zillii* (the latter also under the synonym *Tilapia zillii*)]. These sequences were combined with the Falade *et al.* data and initially aligned using MAFFT version 7.187[17] with manual adjustments made using Mesquite version 3.10[18] (mesquiteproject.wikispaces.com/). A phylogenetic analysis was performed using RAxML version 8.2.8[19]. Initial alignment and phylogenetic analysis were performed through the CIPRES Science Gateway version 3.3[20] (RRID: SCR_008439; phylo.org/). Alignment required reversing or reverse-complementing some of the sequences from Falade *et al.* The problematic sequence KX231778 could not be satisfactorily aligned with the others and had to be excluded from the tree. The remaining COI sequences did cluster with other GenBank sequences in such a way as to suggest the remaining taxonomic determinations reported by Falade *et al.* are credible.

Another anomalous sequence is KX243287 (Clarias_gariepinus_asejire_12), a 16S sequence approximately twice the length of the others. We have no explanation for this.

The evidence presented by Falade *et al.* is not sufficient to determine at least the COI sequence KX231778. The method applied by Falade *et al.* made it nearly impossible to fail to obtain a taxonomic name for each sequence. This is a scientific flaw, and an example of the uncritical application of DNA taxonomy.

This paper was discussed as part of a regular journal discussion group offered by the Endless Forms research group at Naturalis Biodiversity Center, which involves students in the Evolution, Biodiversity, and Conservation program at Leiden University. Similar journal-article-based discussion groups can be found at many universities and Natural History Museums. We support the rationale behind open review journals (blog.f1000research.com/2014/05/21/what-is-open-peer-review/) and therefore decided to share the sense of our discussion with the broader community. We would like to encourage other journal discussion groups to include open review articles in their literature discussions, and consider sharing summaries of their discussions as article comments. Healthy science literature depends on a robust pool of potential reviewers[21]. We see journal discussion groups as an untapped resource for providing feedback on scientific literature, and also as incubators for developing student-scientists into constructive and rigorous peer reviewers.

Dataset 1. Aligned COI sequence data

http://dx.doi.org/10.5256/f1000research.9829.d141383

(FASTA format)

Table 1. Sequences from Falade *et al.*[12] queried using the BOLD and GenBank databases. Top BOLD hit and BOLD identification note summarize results from BOLD. Top Blast hit and Sequence name specify the best match in GenBank (excluding the Falade *et al.* sequences) according to BLASTN, with the Blast metrics Query cover and Ident. See also Table 2 in Falade *et al.* Note that BOLD contains no 16S data, so these sequences are listed as NA (not applicable).

Accession no.	Locus	Specimen voucher no.	Top BOLD hit	BOLD identification note	Top Blast hit	Sequence name	Query cover	Ident	Blast note
KX231778	COI	Coptodon_zilli_odooba_1	No Hit		JX173760.1	Tilapia zillii isolate MAB08 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	76	86	
KX231779	COI	Coptodon_zilli_odooba_2	Top Hit: Chordata - Cichliformes - Tilapia zillii (99.65%)	A species level match could not be made, the queried specimen is likely to be one of the following: Tilapia zillii, Coptodon zillii, Coptodon sp., Oreochromis mossambicus, Coptodon rendalli, Tilapia guineensis	KM658974.1	Coptodon zillii mitochondrion, complete genome	87	99	
KX231780	COI	Coptodon_zilli_odooba_3	Top Hit: Chordata - Cichliformes - Tilapia zillii (99.44%)	A species level match could not be made, the queried specimen is likely to be one of the following: Tilapia zillii, Coptodon zillii, Coptodon sp., Oreochromis mossambicus, Coptodon rendalli	KM658974.1	Coptodon zillii mitochondrion, complete genome	88	99	
KX231781	COI	Sarotherodon_ melanotheron_odooba_4	No Hit		JF894132.1	Sarotherodon melanotheron mitochondrion, complete genome	92	98	
KX231782	COI	Sarotherodon_ melanotheron_odooba_5	No Hit		JF894132.1	Sarotherodon melanotheron mitochondrion, complete genome	92	98	
KX231783	COI	Clarias_gariepinus_ odooba_6	No Hit		KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	99	
KX231784	COI	Clarias_gariepinus_ odooba_7	Top Hit: Chordata - Siluriformes - Clarias gariepinus (99.62%)	A species level match could not be made, the queried specimen is likely to be one of the following: Clarias gariepinus, Clarias sp. NM-2010	KT001082.1	Clarias gariepinus mitochondrion, complete genome	90	99	
KX231785	COI	Clarias_gariepinus_ odooba_8	No Hit		KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	99	
KX231786	COI	Clarias_gariepinus_ odooba_9	No Hit		KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	99	
KX231787	COI	Clarias_gariepinus_ odooba_10	Top Hit: Chordata - Siluriformes - Clarias gariepinus (100%)	A species level match could not be made, the queried specimen is likely to be one of the following: Clarias gariepinus, Clarias sp. NM-2010	KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	99	
KX231788	COI	Clarias_gariepinus_ odooba_11	No Hit		KT001082.1	Clarias gariepinus mitochondrion, complete genome	90	98	

Accession no.	Locus	Specimen voucher no.	Top BOLD hit	BOLD identification note	Top Blast hit	Sequence name	Query cover	Ident	Blast note
KX231789	COI	Clarias_gariepinus_asejire_12	No Hit		KT001082.1	Clarias gariepinus mitochondrion, complete genome	90	98	
KX231790	COI	Clarias_gariepinus_asejire_13	Top Hit: Chordata - Siluriformes - Clarias gariepinus (99.84%)	A species level match could not be made, the queried specimen is likely to be one of the following: Clarias gariepinus, Clarias sp., Clarias magur, Clarias cf. stappersii, Clarias ngamensis	JQ699203.1	Clarias gariepinus isolate CLGP5 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	93	99	
KX231791	COI	Clarias_gariepinus_asejire_14	Top Hit: Chordata - Siluriformes - Clarias gariepinus (100%)	A species level match could not be made, the queried specimen is likely to be one of the following: Clarias gariepinus, Clarias sp., Clarias magur, Clarias cf. stappersii, Clarias ngamensis	KX619412.1	Clarias gariepinus cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	93	99	
KX231792	COI	Clarias_gariepinus_asejire_15	Top Hit: Chordata - Siluriformes - Clarias gariepinus (100%)	A species level match could not be made, the queried specimen is likely to be one of the following: Clarias gariepinus, Clarias sp., Clarias magur, Clarias cf. stappersii, Clarias ngamensis	JQ699201.1	Clarias gariepinus isolate CLGP3 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	93	99	
KX231793	COI	Clarias_gariepinus_asejire_16	No Hit		KX619412.1	Clarias gariepinus cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial	92	99	
KX243276	16S	Coptodon_zilli_odooba_1	NA		KM658974.1	Coptodon zillii mitochondrion, complete genome	93	99	Top three hits from source publication, fourth hit reported
KX243277	16S	Coptodon_zilli_odooba_2	NA		KM658974.1	Coptodon zillii mitochondrion, complete genome	93	99	Top hit this sequence, second hit reported
KX243278	16S	Coptodon_zilli_odooba_3	NA		GQ168017.1	Tilapia aff. zillii 'Kisangani' isolate J72 16S ribosomal RNA gene, partial sequence; mitochondrial	90	99	Top three hits from source publication, fourth hit reported
KX243279	16S	Sarotherodon_melanotheron_odooba_4	NA		JF894132.1	Sarotherodon melanotheron mitochondrion, complete genome	93	99	Top two hits from source publication, third hit reported
KX243280	16S	Sarotherodon_melanotheron_odooba_5	NA		JF894132.1	Sarotherodon melanotheron mitochondrion, complete genome	89	99	Top two hits from source publication, third hit reported

Accession no.	Locus	Specimen voucher no.	Top BOLD hit	BOLD identification note	Top Blast hit	Sequence name	Query cover	Ident	Blast note
KX243281	16S	Clarias_gariepinus_odooba_6	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	97	99	Top three hits from source publication, fourth hit reported
KX243282	16S	Clarias_gariepinus_odooba_7	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	94	99	Top six hits from source publication, seventh hit reported
KX243283	16S	Clarias_gariepinus_odooba_8	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	94	99	Top four hits from source publication, fifth hit reported
KX243284	16S	Clarias_gariepinus_odooba_9	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	94	99	Top six hits from source publication, seventh hit reported
KX243285	16S	Clarias_gariepinus_odooba_10	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	84	99	Top four hits from source publication, fifth hit reported
KX243286	16S	Clarias_gariepinus_odooba_11	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	93	Top two hits from source publication, third hit reported
KX243287	16S	Clarias_gariepinus_asejire_12	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	51	95	Top six hits from source publication, seventh hit reported
KX243288	16S	Clarias_gariepinus_asejire_13	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	94	99	Top four hits from source publication, fifth hit reported
KX243289	16S	Clarias_gariepinus_asejire_14	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	94	99	Top five hits from source publication, sixth hit reported
KX243290	16S	Clarias_gariepinus_asejire_15	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	84	99	Top four hits from source publication, fifth hit reported
KX243291	16S	Clarias_gariepinus_asejire_16	NA		KT001082.1	Clarias gariepinus mitochondrion, complete genome	92	93	Top two hits from source publication, third hit reported

F1000Research 2016, 5:2654 Last updated: 17 MAY 2019

Dataset 2. Phylogenetic tree

http://dx.doi.org/10.5256/f1000research.9829.d141384

(RAxML tree, NEWICK format)

Data availability

F1000Research: Dataset 1. Aligned COI sequence data, 10.5256/f1000research.9829.d141383[22]

F1000Research: Dataset 2. Phylogenetic tree, 10.5256/f1000research.9829.d141384[23]

Author contributions

MS, JM, IvR, MZK, and DS conceived the study and outlined major points. IvR and JM analyzed the data and wrote initial drafts of the manuscript. All authors were involved in the revision of the draft manuscript and have agreed to the final content.

Competing interests

No competing interests were disclosed.

Grant information

The Endless Forms Research Group (Naturalis) budget footed the bill for the journal club drinks at Meneer Jansen in Leiden. IvR is supported by the 'Nederlandse organisatie voor Wetenschappelijk Onderzoek' (NWO Open Programme 824.14.014).

References

1. Schindel DE, Miller SE: **DNA barcoding a useful tool for taxonomists.** *Nature.* 2005; **435**(7038): 17.
PubMed Abstract | Publisher Full Text
2. Miller SE: **DNA barcoding and the renaissance of taxonomy.** *Proc Natl Acad Sci U S A.* 2007; **104**(12): 4775–4776.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Becker S, Hanner R, Steinke D: **Five years of FISH-BOL: brief status report.** *Mitochondrial DNA.* 2011; **22**(Suppl 1): 3–9.
PubMed Abstract | Publisher Full Text
4. Moritz C, Cicero C: **DNA barcoding: promise and pitfalls.** *PLoS Biol.* 2004; **2**(10): e354.
PubMed Abstract | Publisher Full Text | Free Full Text
5. Meier R: **DNA sequences in taxonomy opportunities and challenges.** In: Wheeler QD, editor. *The New Taxonomy.* CRC Press. 2008; 95–127.
Publisher Full Text
6. Ebach MC, de Carvalho MR: **Anti-intellectualism in the DNA barcoding enterprise.** *Zoologica (Curitiba).* 2010; **27**(2): 165–178.
Publisher Full Text
7. Lipscomb D, Platnick N, Wheeler Q: **The intellectual content of taxonomy: A comment on DNA taxonomy.** *Trends Ecol Evol.* 2003; **18**(2): 65–66.
Publisher Full Text
8. Rubinoff D, Holland BS: **Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference.** *Syst Biol.* 2005; **54**(6): 952–961.
PubMed Abstract | Publisher Full Text
9. Rubinoff D, Cameron S, Will K: **A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification.** *J Hered.* 2006; **97**(6): 581–594.
PubMed Abstract | Publisher Full Text
10. Will KW, Rubinoff D: **Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification.** *Cladistics.* 2004; **20**(1): 47–55.
Publisher Full Text
11. Hebert PD, Cywinska A, Ball SL, *et al.*: **Biological identifications through DNA barcodes.** *Proc Biol Sci.* 2003; **270**(1512): 313–321.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Falade MO, Opeme AJ, Benson O: **DNA barcoding of *Clarias gariepinus*, *Coptodon zillii* and *Sarotherodon melanotheron* from Southwestern Nigeria [version 1; referees: 1 approved].** *F1000 Res.* 2016; **5**: 1268.
Publisher Full Text
13. Ratnasingham S, Hebert PD: **BOLD: The Barcode of Life Data System (http://www.barcodinglife.org).** *Mol Ecol Notes.* 2007; **7**(3): 355–364.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Pertsemlidis A, Fondon JW 3rd: **Having a BLAST with bioinformatics (and avoiding BLASTphemy).** *Genome Biol.* 2001; **2**(10): REVIEWS2002.
PubMed Abstract | Free Full Text
15. Ratnasingham S, Hebert PD: **A DNA-based registry for all animal species: the barcode index number (BIN) system.** *PLoS One.* 2013; **8**(7): e66213.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Mutanen M, Kivelä SM, Vos RA, *et al.*: **Species-Level Para- and Polyphyly in DNA Barcode Gene Trees: Strong Operational Bias in European Lepidoptera.** *Syst Biol.* 2016; **65**(6): 1024–1040. pii syw044.
PubMed Abstract | Publisher Full Text | Free Full Text
17. Katoh K, Misawa K, Kuma K, *et al.*: **MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.** *Nucleic Acids Res.* 2002; **30**(14): 3059–3066.
PubMed Abstract | Publisher Full Text | Free Full Text
18. Maddison WP, Maddison DR: **Mesquite: a modular system for evolutionary analysis.** Version 3.10. 2016.
Reference Source
19. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics.* 2014; **30**(9): 1312–1313.
PubMed Abstract | Publisher Full Text | Free Full Text
20. Miller MA, Pfeiffer W, Schwartz T: **Creating the CIPRES Science Gateway for inference of large phylogenetic trees.** In: *Proceedings of the Gateway Computing Environments Workshop (GCE).* New Orleans, LA. 2010; 1–8.
Publisher Full Text
21. Hochberg ME, Chase JM, Gotelli NJ, *et al.*: **The tragedy of the reviewer commons.** *Ecol Lett.* 2009; **12**(1): 2–4.
PubMed Abstract | Publisher Full Text
22. Miller JA, van Riemsdijk I, Khalik MZ, *et al.*: **Dataset 1 in: Comment on Falade *et al.* (2016) DNA-barcoding of *Clarias gariepinus, Coptedon zillii* and *Sarotherodon melanotheron* from Southwestern Nigeria.** *F1000Research.* 2016.
Data Source
23. Miller JA, van Riemsdijk I, Khalik MZ, *et al.*: **Dataset 2 in: Comment on Falade *et al.* (2016) DNA-barcoding of *Clarias gariepinus, Coptedon zillii* and *Sarotherodon melanotheron* from Southwestern Nigeria.** *F1000Research.* 2016.
Data Source

Open Peer Review

Current Peer Review Status: ? ✓

Version 1

Reviewer Report 30 January 2017

https://doi.org/10.5256/f1000research.10597.r19277

✓ **Dora Davies**
Faculty of Natural Sciences, Facultad de ciencias Naturales, Universidad Nacional de Salta, Salta, Argentina, Salta, Argentina
Carolina Davies
Instituto de Patología Experimental (Institute of Experimental Pathology), UNSa-CONICET, República Argentina

The title is appropriate for the content of the article. However, there is a spelling mistake since the correct genus is *Coptodon*, instead of *Coptedon* as the authors wrote it. The abstract concisely summarizes the ideas presented in the article. The authors explain with clarity their points of view about the paper under study. Conclusions are justified on the basis of the analysis performed. The information given is adequate, and discussed with clarity.

Multiple factors converge on decision making, therefore studies that bring together morphology, life cycles, ecology, genetics, and bioinformatics are desirable to improve our comprehension about species, in particular those that come from understudied localities. Falade *et al* (2016) identified fish specimens at species level by morphology, later analyzing genes COI and 16S rRNA with the aim of correlating morphologic and genetic data. Miller *et al* (2017) made an objection to the bioinformatic methodology employed by Falade *et al*., stating that it was “nearly impossible to fail to obtain a taxonomic name for each sequence”. Miller *et al* objected particularly one sequence that produced no hits on BOLD database, a problem also addressed by Falade in their original paper. The absence of genetic sequences in public databases from specimens of remote or understudied areas is a problem that researchers from those areas face quite frequently. Even though Miller *et al* (2017) are correct in addressing the methodology shortness in Falade’s work, it is important to remark that Falade *et al*. made an important contribution in submitting genetic sequences from 3 fish species of the underrepresented country Nigeria to public databases such as GeneBank and BOLD. Hopefully, there will be more interdisciplinary studies on Nigerian fish fauna.

In a more philosophical note, none of the branches of biology can alone answer all the questions, or explain or predict the totality of biological phenomena. In particular, definition of the concept “species” is under discussion even today. Molecular biology and bioinformatics are two of the many tools that are

available to elucidate the boundaries between 2 species. For example, to what extent a similarity percent of nearly 100%, based on the study of certain genes in a certain biological group, can be taken as an indicator that two species are different? That percent seems to be different for different taxa, and also varies depending on the genes under study. Bioinformatic tools to analyze genetic data are improving at a fast pace, but it is still important not to underestimate information about morphology, ecology, life cycles, etc. to complete the picture of each taxa. It is also worth noting that the improvement of bioinformatics tools relies on pre-existing information, and when that information is missing there might be a bit of a problem.

Competing Interests: No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 27 January 2017

https://doi.org/10.5256/f1000research.10597.r19279

?

Shadi Shokralla
Biodiversity Institute of Ontario and Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada

The authors discussed many aspects in Falade *et al*. (2016) article but their explanation didn't convince me with their findings. For example:

- The authors focused on CO1 data and almost ignored 16S data.
- I was expected to see more figures to proof their points.
- "The evidence presented by Falade et al. is not sufficient to determine at least the COI sequence KX231778. The method applied by Falade et al. made it nearly impossible to fail ...." what is the right way in the authors' eyes.
- "The remaining COI sequences did cluster with other GenBank sequences in such a way as to suggest the remaining ....." At which level the clustering parameters were set to? It is a vague expression.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.