

LETTER TO THE EDITOR

Open Access

Response to the Letter to the Editor by Harris



Telleasha L. Greay^{1,2*}, Alireza Zahedi¹, Anna-Sheree Krige¹, Jady M. Owens¹, Robert L. Rees³, Una M. Ryan¹, Charlotte L. Oskam¹ and Peter J. Irwin¹

Abstract

In a letter to the Editor, Harris considers the eight new species of Apicomplexa that were recently identified and named to be invalid on the basis that only molecular characters were provided in the species descriptions. In this response, we counter that the species names are valid as the descriptions have met the requirements of the *International Code of Zoological Nomenclature*; molecular characters can be used to satisfy article 13.1.1 of the code.

Letter to the Editor

We recently described eight novel apicomplexan species in ticks collected from companion animals in Australia: *Babesia lohae*, *Babesia mackerrasorum*, *Hepatozoon banethi*, *Hepatozoon ewingi*, *Theileria apogeana*, *Theileria palmeri*, *Theileria paparinii* and *Theileria worthingtonorum* [1]. The ticks were screened for 18S rDNA (18S) of *Babesia*, *Hepatozoon* and *Theileria* species using PCR and Sanger sequencing. Further molecular characterisation was performed with near full-length 18S sequences obtained (~1400–1700 bp). The genetic distances showed that the 18S sequences of the novel species were sufficiently different from their most closely related named species (1.7–7.9% dissimilarity), which warranted new species classifications, in our opinion [1].

In a Letter to the Editor, Harris [2] suggests that specific characters and valid descriptions were not provided for the novel Australian Apicomplexa described in our study [1], and that the proposed names are invalid. Harris [2] disputes the validity of our approach, stating that naming of these parasites based on genetic distinctiveness and phylogenetic relationships, rather than determination of specific characters, was not consistent with the *International Code of Zoological Nomenclature* (ICZN).

The ICZN does not explicitly rule out the use of DNA sequences alone, which are character-based, to describe a new species and does not specifically state that the description should be morphological. Article 13.1.1 states

that “To be available, every new name published after 1930 ... must be accompanied by a description or definition that states in words characters that are purported to differentiate the taxon” [3].

Traditionally, taxonomists have described organisms using morphological characters, geographical distribution patterns and host specificities. Harris [2] highlights that there is a “Linnean Shortfall” for Apicomplexa, and that molecular studies can help to overcome this issue. Indeed, we are now in an era where it is faster to describe molecular features of microorganisms than morphological characters due to advanced molecular screening approaches. For example, next-generation sequencing enables microorganisms, including viruses, bacteria and protozoans, to be rapidly screened with high accuracy, and millions of sequences from thousands of different species can be generated in a single sequencing assay [4]. The codes and committees governing nomenclature of viral (*International Committee on Taxonomy of Viruses*) and bacterial (*International Code of Nomenclature of Prokaryotes*) microorganisms have largely adopted the use of sequence data to describe novel species, albeit maintaining the outdated requirement of “*Candidatus*” usage for uncultured bacterial species. The criticisms of DNA-sequence-data-only descriptions for protozoans and other eukaryotic taxa have been rebutted in comprehensive reviews and discussions previously [5–7].

The 18S locus is a conservative marker among eukaryotes and due to the unreliability of morphological characters for delimiting most haemoparasites, it is the most widely used locus to delimit haemoparasite species. Importantly, morphological characters are

*Correspondence: telleasha.greay@outlook.com

²Western Australian State Agricultural Biotechnology Centre, Murdoch University, Perth, WA, Australia

Full list of author information is available at the end of the article



indistinguishable among many piroplasm and *Hepatozoon* species, especially in their merozoite, trophozoite and gamont stages, and therefore species can be only reliably delimited using molecular characters (e.g. *18S* sequences). Morphological overlaps in piroplasms can also occur at the genus level. For example, *18S* sequence data were used to reclassify *Babesia equi* as *Theileria equi* [8]. *18S* sequence data have been used to identify and classify several previously unknown haemoprotozoans and other eukaryotic parasites in many other studies including, but not limited to, Persing [9], Quick et al. [10], Thomford et al. [11], Herwaldt et al. [12], Katzer et al. [13], Gubbels et al. [14], Nijhof et al. [15], Moss et al. [16], Jorger & Schrodler [7], Ryan et al. [17], Boscaro et al. [18] and Zatti et al. [19].

The *18S* locus evolves more slowly than other barcoding loci, such as the cytochrome *c* oxidase subunit 1 (*cox1*) gene, therefore less variability is observed in *18S* compared with more rapidly evolving genes and can underestimate species diversity in some organisms [20]. The fact that a significant level of variation was observed at the slowly evolving *18S* locus between the novel and known species in our study only strengthens our conclusion that the species are distinct.

It is certainly true that description of additional genetic markers, genomes and morphological features offers a more complete understanding of the characters of a species compared to a single genetic marker. However, as morphology is not always reliable and a comprehensive genetic database of piroplasms and *Hepatozoon* sequences at loci other than the *18S* locus is not available, then it is appropriate to describe the species based on *18S* sequences. There are no requirements under the ICZN about how much or what type of information should be obtained to describe a new taxon, and neither should there be; the types of biological data that scientists can analyse has expanded over the last century, and any type of data that sufficiently delineates novel species can and should be used to describe new organisms. *18S* is the only genetic marker that has been sequenced for *Hepatozoon* species at present. Therefore, *H. banethi* and *H. ewingi* sequence data from other loci would not have assisted us to determine their species status, as there are no other sequences for comparison from other *Hepatozoon* species present in GenBank®. As for the piroplasms, only the *18S* gene has been sequenced for all the closest congeners of the new piroplasms described in our study, with one exception; the heat-shock protein gene (*hsp*) has been sequenced for the closest relative to *B. mackerrasorum*, *Babesia macropus* [21]. For all the novel species described in our study, the *18S* sequences were compared to *18S* sequences in GenBank, and the genetic distances between the novel apicomplexans and their closest

relatives (described to date) were greater than the genetic distances between the next two most closely related species.

If naming novel microorganisms is constrained by a requirement to describe morphological characteristics, the taxonomic lag (the “Linnean Shortfall”) will be further exacerbated. Taxonomy should enable us to communicate about organisms without confusion, yet there is a propensity to deposit sequence data for novel organisms in GenBank with the “sp.” abbreviation, likely due to the view that morphological descriptions of type-specimens submitted to museums are needed to name species. As a result of this: (i) there are overwhelming numbers of sequences in GenBank that have not been assigned to a species, which makes species identification of similar sequences a cumbersome task (there are currently ~3000 *Babesia*, >2000 *Hepatozoon* and >2000 *Theileria* *18S* sequences on GenBank, and 25% (average length 832 bp), 49% (average length 699 bp) and 27% (average length 927 bp) have no species assigned, respectively); (ii) the ability to communicate about unique and novel “sp.’s” is impeded as there is no requirement for informative isolate names to be provided that are unique to all others; and (iii) there is likely a considerable underestimation of biodiversity on the planet for which data presently exists.

In all cases in our publication, each description was accompanied by GenBank accession numbers of the *18S* sequences, which allows full subsequent comparisons of new and existing sequence data to our sequence data. The text “see above” in the Diagnosis sections of the descriptions refers to the results sections of the paper, where the species differentiations were described and stated in words. The genetic distances were linked to defined characters, the *18S* sequences. The interspecific genetic distances of the *18S* sequences for each novel species that delimit them from other species are stated in the sections “Novel *Babesia* species”, “Novel *Hepatozoon* species”, “Novel *Theileria* species” and “Genetic distance estimates”, where we also refer to summaries of the BLAST results and pairwise genetic distances in Table 5, Additional file 2 and Additional file 4 [1].

At present, there is no standardised format among scientific journals for sequence data descriptions in articles that describe novel species. In the study by Jorger & Schrodler [7], novel species descriptions were accompanied with tables of the nucleotide compositions of the reference DNA sequences, which is not a unique feature of a species, hence GenBank accession numbers of the reference sequences were cited. The nucleotide positions and the genetic distances to other described species are what make the sequences unique. The review by Cook et al. [6] provided a different example of a DNA-sequence-data-only description, with the nucleotide polymorphisms

outlined. We have applied this approach to our sequences in Additional file 1: Figure S1 to illustrate why this format was unsuitable for our multiple long sequences, and would also be unsuitable for large datasets (especially genomes). Providing the accession number for the sequences in GenBank and a description of the genetic distances was more concise in our case, and DNA sequences deposited in GenBank convey the same information as sequences in the text of a paper and are freely accessible.

In all cases, we also deposited the type-material (bisected ticks and DNA extracts derived from the other half of the tick) in recognised museum collections at the Australian Centre for Wildlife Genomics at the Australian Museum, the Queensland Museum and the Tasmanian Museum and Art Gallery for future reference.

Additional file

Additional file 1: Figure S1. Nucleotide sequence alignments generated by the MUSCLE alignment tool [22] in Geneious v10.2.2 (<http://www.geneious.com>, [23]) of the longest reference 18S sequences for novel *Babesia*, *Hepatozoon* and *Theileria* species described in [1] compared with 18S sequences from their closest named relatives. Nucleotide polymorphisms are highlighted (adenine (A): red; guanine (G): yellow; thymine (T): green; cytosine (C): blue). Gaps are represented by dashes. Each line is labelled with the species and corresponding GenBank® accession number in parentheses. Nucleotide base positions are numbered above the sequences. **A** *Babesia lohiae* (MG593272). **B** *Babesia mackerrasorum* (MG593271). **C** *Hepatozoon banethi* (MG758137). **D** *Hepatozoon ewingi* (MG593275). **E** *Theileria apogeana* (MG758116). **F** *Theileria palmeri* (MG758113). **G** *Theileria paparinii* (MG758115). **H** *Theileria worthingtonorum* (MG758114).

Authors' contributions

TLG wrote the manuscript. AZ, A-SK, JMO, RLR, UMR, CLO and PJI contributed to the preparation of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the anonymous reviewers of this manuscript for their assistance.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The data supporting the conclusions of this article are included within our previously published article, Greay et al. [1]. The sequences were submitted in the GenBank database under the accession numbers MG062865, MG571580–MG571582, MG593271–MG593276, MG758109–MG758121 and MG758124–MG758138. Type-material was submitted to the Australian Museum, Sydney, Australia, under the accession numbers KS.128102 and KS.128103, the Queensland Museum, Brisbane, Australia, under the accession numbers QMS108579, A015180 and A015181 and the Tasmanian Museum and Art Gallery, Hobart, Australia, under the accession numbers K463–K4640.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

Not applicable.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹ Vector and Waterborne Pathogens Research Group, College of Science, Health, Engineering and Education, Perth, WA, Australia. ² Western Australian State Agricultural Biotechnology Centre, Murdoch University, Perth, WA, Australia. ³ Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Melbourne, VIC, Australia.

Received: 5 December 2018 Accepted: 8 April 2019

Published online: 24 April 2019

References

1. Greay TL, Zahedi A, Krige A-S, Owens JM, Rees RL, Ryan UM, et al. Endemic, exotic and novel apicomplexan parasites detected during a national study of ticks from companion animals in Australia. *Parasit Vectors*. 2018;11:197.
2. Harris JD. New species need characters: comments on recently described apicomplexan parasites from Australia. *Parasit Vectors*. 2019;12:172. <https://doi.org/10.1186/s13071-019-3424-9>.
3. ICZN. International Commission on Zoological Nomenclature: amendment of articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. *Bull Zool Nomencl.* 2012;69:161–9.
4. Greay TL, Gofton AW, Paparini A, Ryan UM, Oskam CL, Irwin PJ. Recent insights into the tick microbiome gained through next-generation sequencing. *Parasit Vectors*. 2018;11:12.
5. Adl SM, Leander BS, Simpson AGB, Archibald JM, Anderson OR, Bass D, et al. Diversity, nomenclature, and taxonomy of protists. *Syst Biol.* 2007;56:684–9.
6. Cook LG, Edwards RD, Crisp MD, Hardy NB. Need morphology always be required for new species descriptions? *Invertebr Syst.* 2010;24:322–6.
7. Jorger KM, Schrod M. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front Zool.* 2013;10:59.
8. Laha R, Das M, Sen A. Morphology, epidemiology, and phylogeny of *Babesia*: an overview. *Trop Parasitol.* 2015;5:94–100.
9. Persing DH. Diagnostic molecular microbiology. Current challenges and future directions. *Diagn Microbiol Infect Dis.* 1993;16:159–63.
10. Quick RE, Herwaldt BL, Thomford JW, Garnett ME, Eberhard ML, Wilson M, et al. Babesiosis in Washington State: a new species of *Babesia*? *Ann Intern Med.* 1993;119:284–90.
11. Thomford JW, Conrad PA, Telford SR 3rd, Mathiesen D, Bowman BH, Spielman A, et al. Cultivation and phylogenetic characterization of a newly recognized human pathogenic protozoan. *J Infect Dis.* 1994;169:1050–6.
12. Herwaldt BL, Kjemtrup AM, Conrad PA, Barnes RC, Wilson M, McCarthy MG, et al. Transfusion-transmitted babesiosis in Washington State: first reported case caused by a WA1-type parasite. *J Infect Dis.* 1997;175:1259–62.
13. Katzer F, McKellar S, Kirvar E, Shiels B. Phylogenetic analysis of *Theileria* and *Babesia equi* in relation to the establishment of parasite populations within novel host species and the development of diagnostic tests. *Mol Biochem Parasitol.* 1998;95:33–44.
14. Gubbels MJ, Hong Y, van der Weide M, Qi B, Nijman IJ, Guangyuan L, et al. Molecular characterisation of the *Theileria buffeli/orientalis* group. *Int J Parasitol.* 2000;30:943–52.
15. Nijhof AM, Penzhorn BL, Lynen G, Mollel JO, Morkel P, Bekker CPJ, et al. *Babesia bicornis* sp. nov. and *Theileria bicornis* sp. nov.: tick-borne parasites associated with mortality in the black rhinoceros (*Diceros bicornis*). *J Clin Microbiol.* 2003;41:2249–54.
16. Moss JA, Xiao J, Dungan CF, Reece KS. Description of *Perkinsus beihaiensis* n. sp., a new *Perkinsus* sp. parasite in oysters of southern China. *J Eukaryot Microbiol.* 2008;55:117–30.

17. Ryan U, Papparini A, Tong K, Yang R, Gibson-Kueh S, O'Hara A, et al. *Cryptosporidium huwi* n. sp. (Apicomplexa: Eimeriidae) from the guppy (*Poecilia reticulata*). *Exp Parasitol*. 2015;150:31–5.
18. Boscaro V, James ER, Fiorito R, Hehenberger E, Karnkowska A, del Campo J, et al. Molecular characterization and phylogeny of four new species of the genus *Trichonympha* (Parabasalia, Trichonymphea) from lower termite hindguts. *Int J Syst Evol Microbiol*. 2017;67:3570–5.
19. Zatti SA, Atkinson SD, Maia AAM, Bartholomew JL, Adriano EA. Novel *Henneguya* spp. (Cnidaria: Myxozoa) from cichlid fish in the Amazon basin cluster by geographic origin. *Parasitol Res*. 2018;117:849–59.
20. Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D. The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proc Natl Acad Sci USA*. 2012;109:16208–12.
21. Donahoe SL, Peacock CS, Choo AYL, Cook RW, O'Donoghue P, Crameri S, et al. A retrospective study of *Babesia macropus* associated with morbidity and mortality in eastern grey kangaroos (*Macropus giganteus*) and agile wallabies (*Macropus agilis*). *Int J Parasitol Parasites Wildl*. 2015;4:268–76.
22. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7.
23. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

