

OPINION

Citation of taxonomic publications: the why, when, what and what not

RUDOLF MEIER

Department of Biological Sciences, National University of Singapore, Singapore, Singapore

Species-level research in taxonomy is inadequately cited (Krell, 2000; Agnarsson & Kuntner, 2007; Ebach *et al.*, 2011; Haszprunar, 2011; Wägele *et al.*, 2011). This has been recognized for many decades, but it did not hurt taxonomy until citation counts started to influence funding, careers, and positions (Krell, 2000; Agnarsson & Kuntner, 2007; Bortolus, 2008; Seifert *et al.*, 2008). One seemingly easy fix that has been proposed or even mandated by some journals is to require full citations of original species descriptions (see also Seifert *et al.*, 2008; Haszprunar, 2011). For example, *Organisms, Diversity, and Evolution* ‘will now require that author citations for species names be supported by inclusion of the full, accompanying reference’ (Bininda-Emonds, 2011) and *Zootaxa*’s instructions for authors includes: ‘Author(s) of species name must be provided when the scientific name of any animal species is first mentioned (the year of publication needs not be given; if you give it, then provide a full reference of this in the reference list)’. Other journals encourage or allow such citations (see Wägele *et al.*, 2011).

Here, I argue that citing original descriptions is often misleading and that the proper citation credits should instead go to the literature that was used for specimen identification and the publications that contain the most accurate information on the currently accepted species limits (‘taxon concept’ *sensu* Franz, 2005) of the species under study (see also Agnarsson & Kuntner, 2007). I argue that biological publications need to include identification techniques and taxon concepts in the ‘Methods’ section and that the identifications should be treated as a ‘Result’ in the corresponding section. The information should be succinctly presented so that reviewers and editors can assess compliance.

Why many species descriptions should not be cited

Publications should be cited when they contain significant scientific background information or support an argument in a manuscript. Many original species descriptions – often those from the 18th and 19th centuries – do not meet either threshold.

Correspondence: Rudolf Meier, Lee Kong Chian Natural History Museum and Department of Biological Sciences, National University of Singapore, 14 Science Dr 4, Singapore 117543, Singapore. E-mail: meier@nus.edu.sg

This means that citing is inappropriate. I use Meigen’s description of *Drosophila melanogaster* as an example (Meigen, 1830) because it has been argued elsewhere that he should be cited in the modern *D. melanogaster* literature. Meigen’s description reads: ‘head, scutum, and legs clay-yellow; abdomen black. The halteres are white, the wings colorless – from Austria, Kiel, and Hamburg. 3 mm’ [translated from German; Fig. 1] and was published in a treatise that also contains species descriptions for nine new *Drosophila* species (Table S1) of which four are now synonymized (Brake & Bächli, 2008). Notably, one (*D. fasciata*) is now synonymized with *Drosophila melanogaster*. This means that based on our current understanding of the species, Meigen’s species boundaries for this species were incorrect. Whoever opts to cite Meigen in a recent paper is mis-citing Meigen because his *D. melanogaster* is not the taxon used today. If authors of a recent paper would like to give credit to Meigen, they should also provide the citation to a subsequent paper that contains the currently accepted taxon concept and they should state that the species boundaries of the latter were used. Note that citing such subsequent literature on *D. melanogaster* is also needed because Meigen’s description of the species was too superficial to be useful in the 21st century.

Of course, there are reasons why Meigen’s work remains relevant. The first is that his nomenclatural skills were impeccable and he picked a name (*Drosophila melanogaster*) that had not yet been used (and he followed all nomenclatural conventions of the day). The second is that names are tied to species through name-bearing types. This means that any author reviewing the taxonomic history of *D. melanogaster* will have to consult Meigen’s description and types.

Taxon-delimitation versus naming

The main reason why I am against mechanically citing original species descriptions is that it conflates taxon concepts (species delimitation) and species naming (Franz, 2005; Agnarsson & Kuntner, 2007). Taxon delimitation is hypothesis-driven because species are taxon hypotheses as long as they are based on explicit species concepts (Laamanen *et al.*, 2003; Tan *et al.*, 2010). Naming taxa, however, is not hypothesis-driven, albeit important for more technical reasons. After all, biology needs an

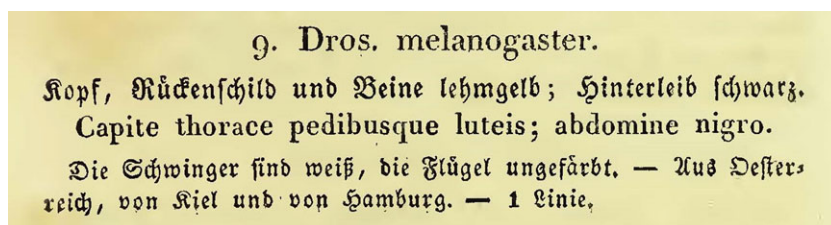


Fig. 1. Original description of *Drosophila melanogaster* by Meigen. [Colour figure can be viewed at wileyonlinelibrary.com].

unambiguous naming system for species so that all information on a species can be accurately filed and retrieved (Patterson *et al.*, 2010). Therefore, authors who would like to give credit for naming a species should also check whether there is a need to cite a subsequent paper that provides an updated taxon concept. Such 'dual' citations were routine in some of the older literature where credit was given to the author who provided the species name and to those authors who updated the taxon concept.

What should be cited?

Specimens used in biological research have to be identified to species before the results are published. In model-based biology, the source of identification is often revealed when authors state the source of cultures (Agnarsson & Kuntner, 2007). Unfortunately, much of the remaining biological literature lacks similar transparency. In many, perhaps most, papers, identification procedures are not mentioned and the utilized resources are not cited (e.g. latest species-level taxonomic literature and/or reference collections). I argue here that this is the ultimate source for the citation crisis in taxonomy. The main problem is that, currently, specimen identification is not treated as part of the 'Methods' (see Bortolus, 2008) and the outcome is not considered a 'Result' (or identification hypothesis; Valdecasas *et al.*, 2014). It is here that a change of policy would be very helpful. Providing tables with GenBank numbers is required for those studies that use molecular data. Let's have a similar requirement for all studies that require the identification of specimens to species. Such studies should include explicit identification and taxon concept statements that are supported by references. For large-scale studies, this information could be presented in a table; species could be in rows and one column each could be dedicated to: (i) identification methods/literature (Bortolus, 2008); (ii) information on which species boundaries were used ('taxon concepts' *sensu* Franz 2005; Kennedy *et al.*, 2005); and (iii) voucher depository (Huber, 1998; Pleijel *et al.*, 2008). When presented as a table, editors and reviewers could easily check whether identification and species boundaries were properly treated and acknowledged. Note that once identifications are covered in the Methods and Results sections, high-quality taxonomic publications will be cited more heavily because comprehensive revisions with good identification keys would get a larger share of citations. Original species descriptions would also continue to be cited but only if they stood the test of time and included valid taxon concepts and/or identification tools.

There are many reasons why specimen identification should be treated as a 'Method' and reported as a 'Result'. The quality and accuracy of most publications are critically dependent on correct species identification (Bortolus, 2008). Any casual treatment of this step devalues publications, as my laboratory experienced when a mistake in an identification key led to the use of an incorrect name in multiple publications (Meier, 1996; Puniamoorthy *et al.*, 2008, 2010; Su *et al.*, 2008; Ang & Meier, 2010). We erroneously identified a species as *Sepsis monostigma* when the correct species name should have been *S. latiforceps*. Our identification was based on an illustrated key which unfortunately contained an error. We used the key correctly to identify a species-level unit, but the name provided in the key was incorrect. If our manuscripts had clearly stated the identification method, the error may have been caught earlier by reviewers who may have been aware of the faulty key. This would also have allowed subsequent users of the publications to track the mistake. Precision with regard to identification methods and taxon concepts would also have helped with avoiding the problems that were created when many of the leeches used in research and medical treatments turned out to belong to multiple species (Bely & Weisblat, 2006; Siddall *et al.*, 2007). Requiring explicit information about identification literature would also flag the use of 'grey' literature that lacks peer review and is nevertheless frequently used in ecological studies (Bortolus, 2008).

Integrative taxonomy and species boundaries

Being explicit about identification methods and species boundaries is particularly important in the 21st century because a large number of species boundaries are revised based on newly available molecular data (Tan *et al.*, 2010; Schwentner *et al.*, 2011; Pante *et al.*, 2015). Many morphologically cryptic species are being discovered (Bickford *et al.*, 2007) which means that 'old' species boundaries are changing and many 'old' species are split into multiple species, i.e. identifications predating the splits now pertain to species groups (Pante *et al.*, 2015). It is therefore important that authors cite the papers that provided the species boundaries and identification keys that were used. Of course, one by-product of properly citing these papers is that proper credit is given to the latest and most up-to-date taxonomic work (see also Wägele *et al.*, 2011) which is usually of much higher quality than many original species descriptions (Agnarsson & Kuntner, 2007; Sangster & Luksenburg, 2015;

Valdecasas *et al.*, 2014). For example, species descriptions of birds published in the 2000s include much more detail (longer descriptions, maps, illustrations), are based on more material (more specimens, species comparisons), and are less likely to be revised soon after description (see Sangster & Luksenburg, 2015) than descriptions from the 1930s. In contrast, mechanically citing original species descriptions would be counterproductive because it ignores the quality differences between early and recent taxonomic work and implies that taxonomy is a static discipline without progress.

Are species names hypotheses?

This has been repeatedly argued (e.g. ‘Basically, species names are hypotheses of evolutionary, genealogical relationships, which may change when tested against new data...’; Sluys, 2013), but such statements conflate taxon concept and species name. Species-level taxon concepts are hypotheses that can change (Dominguez & Wheeler, 1997; Franz, 2005; Kennedy *et al.*, 2005; Valdecasas *et al.*, 2014), but species names are firmly tied to types and are not testable. Let me elaborate. If we assume that species are not man-made constructs, then the biological world has always been inhabited by millions of species. At some point, *Homo sapiens* evolved and, a bit later, taxonomists started to delimit species based on a variety of techniques that yielded data on those specimens that were collected for scientific study. When a new species was discovered, a species or a series of specimens was explicitly designated or implicitly fixed as a name-bearing type and the name was attached to the specimen(s). Placing multiple specimens into the same species, circumscribing the species, and deciding that the species is new to science are hypothesis-driven activities. However, placing a type label on a specimen is not. Indeed, if names were hypotheses, they should be testable but the whole point of types and nomenclature is that names stay attached to types regardless of how much new evidence is collected. Species boundaries may continue to evolve, but determining the correct species name is a mechanical process: if none of the specimens belonging to a delimited species-level unit is a type, the species is new and has to be described. If one of the specimens in a newly delimited species-level unit is a type, the species’ name can be settled based on label information. If multiple specimens in a delimited species-level unit are types, nomenclature sorts out which name has priority, i.e. the naming process is deterministic and not hypothesis-driven.

There is no doubt that the current citation practices in biology are hurting taxonomy. I argue here that the main reason is that species identification is not considered part of the Methods and species lists are not considered part of the Results. Once journals require formal statements concerning how species identifications were obtained, the chronic under-citation of the relevant taxonomic literature will disappear. Such a requirement will hopefully also encourage authors to reflect more on this crucial step of any study, while mechanically citing species description will only generate more work for undergraduate research assistants.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/syen.12215

Table S1. Species described or identified as *Drosophila* in Meigen’s 1830 publication. Validity of taxon concepts according to Brake & Bächli, 2008.

Acknowledgements

I would like to thank Peter Cranston, Nate Hardy, Shaun Winterton and two anonymous reviewers for providing comments on the manuscript.

References

- Agnarsson, I. & Kuntner, M. (2007) Taxonomy in a changing world: seeking solutions for a science in crisis. *Systematic Biology*, **56**, 531–539.
- Ang, Y.C. & Meier, R. (2010) Five additions to the list of Sepsidae (Diptera) for Vietnam: *Perochaeta cuirassa* sp. n., *Perochaeta lobo* sp. n., *Sepsis spura* sp. n., *Sepsis sepsi* Ozerov, 2003 and *Sepsis monostigma* Thompson, 1869. *Zookeys*, **70**, 41–56.
- Bely, A.E. & Weisblat, D.A. (2006) Lessons from leeches: a call for DNA barcoding in the lab. *Evolution & Development*, **8**, 491–501.
- Bickford, D., Lohman, D.J., Sodhi, N.S. *et al.* (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148–155.
- Bininda-Emonds, O. (2011) Supporting species in ODE: explaining and citing. *Organisms Diversity & Evolution*, **11**, 1–2.
- Bortolus, A. (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A Journal of the Human Environment*, **37**, 114–118.
- Brake, I. & Bächli, G. (2008) *Drosophila Meigen 1830. Systema Dipterorum, Version 1.5. 9 records* [WWW document]. URL <http://www.diptera.org/> [accessed on 1 April 2016].
- Dominguez, E. & Wheeler, Q.D. (1997) Forum – taxonomic stability is ignorance. *Cladistics*, **13**, 367–372.
- Ebach, M.C., Valdecasas, A.G. & Wheeler, Q.D. (2011) Impediments to taxonomy and users of taxonomy: accessibility and impact evaluation. *Cladistics*, **27**, 550–557.
- Franz, N.M. (2005) On the lack of good scientific reasons for the growing phylogeny/classification gap. *Cladistics*, **21**, 495–500.
- Haszprunar, G. (2011) Species delimitations – not ‘only descriptive’. *Organisms Diversity & Evolution*, **11**, 249–252.
- Huber, J.T. (1998) The importance of voucher specimens, with practical guidelines for preserving specimens of the major invertebrate phyla for identification. *Journal of Natural History*, **32**, 367–385.
- Kennedy, J.B., Kukla, R. & Paterson, T. (2005) Scientific names are mmambiguous as identifiers for biological taxa: their context and definition are required for accurate data integration. *Data Integration in the Life Sciences: Second International Workshop, DILS 2005, San Diego, CA, USA, July 20–22, 2005, Proceedings* (ed. by B. Ludäscher and L. Raschid), pp. 80–95. Springer, Berlin and Heidelberg.
- Krell, F.T. (2000) Impact factors aren’t relevant to taxonomy. *Nature*, **405**, 507–508.
- Laamanen, T.R., Petersen, F.T. & Meier, R. (2003) Kelp flies and species concepts – the case of *Coelopa frigida* (Fabricius, 1805)

- and *C. nebularum* Aldrich, 1929 (Diptera : Coelopidae). *Journal of Zoological Systematics and Evolutionary Research*, **41**, 127–136.
- Meier, R. (1996) Larval morphology of the Sepsidae (Diptera: Sciomyzoidea), with a cladistic analysis using adult and larval characters. *Bulletin of the American Museum of Natural History*, **228**, 3–147.
- Meigen, J.W. (1830) *Systematische Beschreibung der bekannten europaischen zweifluegeligen Insekten*. Schulzische Buchhandlung, Hamm.
- Pante, E., Puillandre, N., Viricel, A. *et al.* (2015) Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, **24**, 525–544.
- Patterson, D.J., Cooper, J., Kirk, P.M., Pyle, R.L. & Remsen, D.P. (2010) Names are key to the big new biology. *Trends in Ecology & Evolution*, **25**, 686–691.
- Pleijel, F., Jondelius, U., Norlinder, E. *et al.* (2008) Phylogenies without roots? A plea for the use of vouchers in molecular phylogenetic studies. *Molecular Phylogenetics and Evolution*, **48**, 369–371.
- Puniamoorthy, N., Su, K.F.Y. & Meier, R. (2008) Bending for love: losses and gains of sexual dimorphisms are strictly correlated with changes in the mounting position of sepsid flies (Sepsidae : Diptera). *BMC Evolutionary Biology*, **8**, 155.
- Puniamoorthy, N., Kotrba, M. & Meier, R. (2010) Unlocking the "Black box": internal female genitalia in Sepsidae (Diptera) evolve fast and are species-specific. *BMC Evolutionary Biology*, **10**, 275.
- Sangster, G. & Luksenburg, J.A. (2015) Declining rates of species described per taxonomist: slowdown of progress or a side-effect of improved quality in taxonomy? *Systematic Biology*, **64**, 144–151.
- Schwentner, M., Timms, B.V. & Richter, S. (2011) An integrative approach to species delineation incorporating different species concepts: a case study of *Limnadoropsis* (Branchiopoda: Spinicaudata). *Biological Journal of the Linnean Society*, **104**, 575–599.
- Seifert, K., Crous, P. & Frisvad, J. (2008) Correcting the impact factors of taxonomic journals by appropriate citation of taxonomy (ACT). *Persoonia*, **20**, 105.
- Siddall, M.E., Trontelj, P., Utevsky, S.Y., Nkamany, M. & Macdonald, K.S. (2007) Diverse molecular data demonstrate that commercially available medicinal leeches are not *Hirudo medicinalis*. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **274**, 1481–1487.
- Sluys, R. (2013) The unappreciated, fundamentally analytical nature of taxonomy and the implications for the inventory of biodiversity. *Biodiversity and Conservation*, **22**, 1095–1105.
- Su, K.F.Y., Kuttu, S.N. & Meier, R. (2008) Morphology versus molecules: the phylogenetic relationships of Sepsidae (Diptera: Cyclorhapha) based on morphology and DNA sequence data from ten genes. *Cladistics*, **24**, 902–916.
- Tan, D.S.H., Ang, Y.C., Lim, G.S., Ismail, M.R.B. & Meier, R. (2010) From 'cryptic species' to integrative taxonomy: an iterative process involving DNA sequences, morphology, and behaviour leads to the resurrection of *Sepsis pyrrhosoma* (Sepsidae: Diptera). *Zoologica Scripta*, **39**, 51–61.
- Valdecasas, A.G., Peláez, M.L. & Wheeler, Q.D. (2014) What's in a (biological) name? The wrath of Lord Rutherford. *Cladistics*, **30**, 215–223.
- Wägele, H., Klussmann-Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A. & Wägele, J.W. (2011) The taxonomist – an endangered race. A practical proposal for its survival. *Frontiers in Zoology*, **8**, 25.

Accepted 12 September 2016

First published online 21 October 2016