

LETTER

The limitations of DNA-based approaches to species delimitation cannot be overcome by expanded sampling

Luo and Zhu (2022) write that “DNA-based integrative taxonomy” requires researchers to ensure they have sufficient sampling to justify the delimitation of a species. While we agree that sample size is an undeniably important aspect of species delimitation, this is neither unique to DNA-based approaches, nor sufficient on its own to address their limitations. As the authors also note, increasing sample size is not always practical. We would add that nor is it always necessary – many species are rare or narrow endemics (Willis & Undy Yule, 1922) and the majority of new species are adequately described from only a few specimens (Goodwin *et al.*, 2020). It is in this sense that we previously outlined the heuristic nature of species delimitation, not as a selective choice or “species concept”, but as an unavoidable result of the nature of biological research and of species themselves (Wells *et al.*, 2021).

The motivations driving an increase in DNA-based taxonomy appear to be two-fold: (i) a desire to accelerate the taxonomic process, and (ii) the belief that molecular DNA-sequence data somehow provides a superior level of empirical certainty when compared to “traditional” morphological data. It is easy to see how both of these views have taken hold, and while both contain a seductive element of truth, they are also oversimplifications that risk leading taxonomy and systematics down intellectual and practical dead-ends.

Much has been written about the need to accelerate the rate of taxonomic discovery at a time of widespread biodiversity loss and climate change (*e.g.* Godfray, 2002, 2007; Tautz *et al.*, 2002; Wheeler *et al.*, 2004; Wheeler, 2020). The rise of molecular phylogenetics, DNA-barcoding, and relatively cheap and easy sequencing have undoubtedly facilitated more effective taxonomic revisions (Muñoz-Rodríguez *et al.*, 2019). The strength of these techniques lies in their utility for rapidly testing hypotheses of relatedness that have been derived from the comparative analysis of morphological and ecological data. They are not however, a replacement for these comparative analyses, since DNA-sequence data cannot in and of itself provide any evidence of either phenotypic or ecological cohesion (Wells *et al.*, 2022).

In fact, no such universal threshold exists that enables the delimitation of species from a single source of data. Instead, the highly variable nature of biological diversity means that within the same lineage what is a useful character or characters to discover species boundaries can be very different (de Queiroz, 2007). Sometimes in botany for example, leaves can be diagnostic of species boundaries, sometimes pollen, sometimes calyx structure, sometimes ecology and reproductive isolation and so on. Similarly, for zoology and mycology the proxies we use to detect species boundaries are varied and only become apparent after a degree of intensive study.

Quite apart from the fact that thresholds of genetic distances are not therefore an effective universal means of discerning species (Freudenstein *et al.*, 2017; Sukumaran & Knowles, 2017), describing species based on differences in DNA-sequence data will only speed up taxonomy if our aim is simply to create lists of names with no associated information on distribution, phenotype, ecological preferences or conservation status (de Carvalho *et al.*, 2008). The strength and purpose of taxonomy has never been merely identifying and naming things however, but rather *describing* them and their life histories. In Western literature, the description and empirical identification of entities we now call species can be traced at least as far back as Dioscorides, nearly 2000 years ago, and likely much further in Chinese medicinal literature. A major milestone in contemporary taxonomy was Linnaeus’ publication of *Species Plantarum* in 1753 in which both known species with polynomial names and new species being described for the first time were provided with binomial names. This was the birth of the binomial nomenclature still in use today, providing the principles for a standardized set of rules for the formal taxonomic description of the world’s biodiversity. Many of these species are still recognised today. Moreover, a large number of them have been shown to be coincidental to entities recognised by people operating outside an evolutionary taxonomic framework (Mayr, 1963; Diamond, 1966; Bulmer & Tyler, 1968; Berlin, 1973, 1992; Berlin *et al.*, 1974; Bulmer *et al.*, 1975; Majnep & Bulmer, 1977; Atran, 1998; Coyne & Orr, 2004; Ludwig, 2015). In trying to understand the world around them, humans motivated by medicinal, nutritional and shelter needs, or by sheer curiosity, have recognised consistent and important

differences between groups of biological organisms – species – and have named them accordingly. These names serve as ready means of identification and communication, and under the Linnaean system should indicate patterns of relationship, but it is the information attached to those names that provides their true value.

In a similar vein, various scientific disciplines have used species counts and their distribution patterns to try to understand biogeography, phylogenetic history, conservation priorities, latitudinal biodiversity gradients, speciation and biodiversity hotspots. It seems therefore, that the entities recognised in biology as species have proven useful for a variety of purposes over time and by different communities of people using different methods for their identification. By way of analogy, models are developed in science to aid understanding of various phenomena and in the same way species can be viewed as akin to a model of an entity to help us understand the world around us. Models and species share characteristics of approximation and iterative feedback loops that allow necessary finetuning and refinement in light of new data. In this sense species can be viewed as a heuristic, allowing an approach to problem solving that employs practical methods not guaranteed to be optimal or perfect, but sufficient for immediate goals. Attempts to accelerate taxonomy therefore need to keep this utility in mind if they are to be more than an exercise in stamp-collecting or list-making.

The notion that DNA-sequence data can provide an objective level of truth about species delimitation is tied to its association with evolutionary theory. Since DNA is understood to carry the genetic information subject to evolutionary change, understanding the genetics of a species should therefore in some sense mean understanding its evolution. In the evolutionary era taxonomists do not work in a vacuum without recourse to considering phylogenetic relationships, ecological niche and reproductive behaviour, albeit often using proxies to estimate those features. But species are more than evolution. Every species hypothesis is a hypothesis of an entity with shared ecological and evolutionary tendencies and these hypotheses are based on congruence of contingent properties that act as proxies of those evolutionary and ecological tendencies. Wells *et al.* (2021, 2022), as others have before, held the view that species are more than diverging evolutionary lineages, as this view tends to place emphasis on monophyly and molecular systematics to the exclusion of phenotypic and ecological cohesion that are equally as important properties of species. Instead, a balanced view of the contingent properties that are required for species recognition necessarily involves morphological and ecological considerations as well as phylogenetic ones.

Similarly, a fundamental distinction is often made between the species category and species realism. We consider this distinction mistaken, as it reduces the importance of the category to being more or less akin to taxonomic stamp-collecting based on antiquated essentialist and static views of nature, while philosophers, evolutionary biologists and phylogeneticists tackle the important job of wrestling with ontology and epistemology. These disciplines are all reliant on species as a vital unit of research however, requiring them to construct species definitions that are consistent with evolutionary theory. The “species problem” can thus be viewed as the continuous attempt by specific academic disciplines to constrain the category of species in a more limited way to emphasise patterns, processes and philosophical issues in relation to their particular interests.

We argue that the distinction between the species category and species realism has not been a useful one, and furthermore that the species category when properly understood provides an umbrella concept – species as a heuristic – under which the pre-occupations of evolutionary biology, ecology and other disciplines can proceed in their research unhindered by the species problem but informed by its existence. It is in this regard that the species category and how it is empirically determined is one of the most important facets of reconciling the theory and practice of the species problem. By contrast, in simply emphasising the importance of more comprehensive sampling, Luo and Zhu (2022) fail to address the fundamental purpose of species delimitation, which risks obscuring the true limitations of DNA-based taxonomy.

Funding This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) research grant [T001445/1 to R.S.]. T.W. is also funded by the BBSRC as part of the Interdisciplinary Biosciences Doctoral Training Partnership at the University of Oxford.

Tom Wells¹*, Tom Carruthers², Robert W. Scotland¹

¹Department of Plant Sciences, University of Oxford, Oxford, UK

²Royal Botanic Gardens, Kew, Richmond, UK

*Corresponding author, E-mail: tom.wells@plants.ox.ac.uk

References

- Atran, S. 1998. Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21: 547–609. doi: 10.1017/S0140525X98001277
- Berlin, B. 1973. Folk systematics in relation to biological classification and nomenclature. *Annual Review of Ecology and Systematics*, 4: 259–271. doi: 10.1146/annurev.es.04.110173.001355
- Berlin, B. 1992. *Ethnobotanical Classification*. Princeton University Press, Princeton, NJ. 335pp.
- Berlin, B., Breedlove, D.E., Raven, P.H. 1974. *Principles of Tzeltal Plant Classification: An Introduction to the Botanical Ethnography of a Mayan-Speaking, People of Highland, Chiapas*. Academic Press, New York & London. 660pp.
- Bulmer, R.N.H., Menzies, J.I., Parker, F. 1975. Kalam classification of birds and reptiles. *Journal of the Polynesian Society*, 84: 267–308.
- Bulmer, R.N.H., Tyler, N.J. 1968. Kalam classification of frogs. *Journal of the Polynesian Society*, 77: 333–385.
- Coyne, J.A., Orr, A.H. 2004. Ch. 1. Species: Reality and concepts. In: Coyne, J.A., Orr, A.H. (eds). *Speciation*. Sinauer Associates, Sunderland, MA. pp. 9–54. Available from <http://www.joelvelasco.net/teaching/2890/coyneorr04-speciationch1.pdf> (accessed 31 March 2022).
- de Carvalho, M.R., Bockmann, F.A., Amorim, D.S., Brandão, C.R.F. 2008. Systematics must embrace comparative biology and evolution, not speed and automation. *Evolutionary Biology*, 35: 150–157. doi: 10.1007/s11692-008-9018-7
- de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*, 56: 879–886. doi: 10.1080/10635150701701083
- Diamond, J.M. 1966. Zoological classification system of a primitive people. *Science*, 151: 1102–1104. doi: 10.1126/science.151.3714.1102
- Freudenstein, J.V., Broe, M.B., Folk, R.A., Sinn, B.T. 2017. Biodiversity and the species concept - Lineages are not enough. *Systematic Biology*, 66: 644–656. doi: 10.1093/sysbio/syw098
- Godfray, H.C.J. 2002. Challenges for taxonomy. *Nature*, 417: 17–19. doi: 10.1038/417017a
- Godfray, H.C.J. 2007. Linnaeus in the information age. *Nature*, 446: 259–260. doi: 10.1038/446259a
- Goodwin, Z.A., Muñoz-Rodríguez, P., Harris, D.J., Wells, T., Wood, J.R.I., Filer, D., Scotland, R.W. 2020. How long does it take to discover a species? *Systematics and Biodiversity*, 18: 784–793. doi: 10.1080/14772000.2020.1751339
- Ludwig, D. 2015. Indigenous and scientific kinds. *The British Journal for the Philosophy of Science*, 68: 187–212. doi: 10.1093/bjps/axv031
- Luo, A., Zhu, C.D. 2022. DNA-based integrative taxonomy: a plea for larger sample size. *Zoological Systematics*, 47(1): 89–92. doi: 10.11865/zs.2022103
- Majnep, I.S., Bulmer, R.N.H. 1977. *Birds of my Kalam Country*. Auckland/Oxford University Press, Auckland/Oxford. 219pp.
- Mayr, E. 1963. *Animal Species and Evolution*. Belknap Press of Harvard University, Cambridge, MA. 797pp.
- Muñoz-Rodríguez, P., Carruthers, T., Wood, J.R.I., Williams, B.R.M., Weitemier, K., Kronmiller, B., Goodwin, Z., Sumadijaya, A., Anglin, N.L., Filer, D., Harris, D., Rausher, M.D., Kelly, S., Liston, A., Scotland, R.W. 2019. A taxonomic monograph of *Ipomoea* integrated across phylogenetic scales. *Nature Plants*, 5: 1136–1144. doi: 10.1038/s41477-019-0535-4
- Sukumaran, J., Knowles, L.L. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, 114: 1607–1611. doi: 10.1073/pnas.1607921114
- Tautz, D., Arctandert, P., Minelli, A., Thomas, R.H., Vogler, A.P. 2002. DNA points the way ahead in taxonomy. *Nature*, 418: 479. doi: 10.1038/418479a
- Wells, T., Carruthers, T., Muñoz-Rodríguez, P., Sumadijaya, A., Wood, J.R. I., Scotland, R.W. 2021. Species as a heuristic: reconciling theory and practice. *Systematic Biology*: syab087. doi: 10.1093/sysbio/syab087
- Wells, T., Carruthers, T., Scotland, R.W. 2022. Heuristics, species and the analysis of systematic data. *Trends in Plant Science* (in press).
- Wheeler, Q. 2020. A taxonomic renaissance in three acts. *Megataxa*, 1(1): 4–8. doi: 10.11646/megataxa.1.1.2
- Wheeler, Q.D., Raven, P.H., Wilson, E.O. 2004. Taxonomy: Impediment or expedient? *Science*, 303: 285. doi: 10.1126/science.303.5656.285
- Willis, J.C., Undy Yule, G. 1922. Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature*, 109: 177–179. doi: 10.1038/109177a0