



The taxonomy's new clothes: a little more about the DNA-based taxonomy

LEANDRO M. SANTOS^{1,2} & LUIZ R. R. FARIA^{1,3}

1- *Laboratório de Biologia Comparada de Hymenoptera, Departamento de Zoologia, Universidade Federal do Paraná. Caixa Postal 19.020, 81531-980 Curitiba, PR, Brazil.*

2- *Programa de Pós-Graduação em Entomologia, Departamento de Zoologia, Universidade Federal do Paraná. CNPq fellowship. megalopta@gmail.com*

3- *Current Address: Departamento de Ciências Agrárias e Biológicas, Centro Universitário Norte do Espírito Santo, Universidade Federal do Espírito Santo (CEUNES/UFES). Rodovia BR-101 Norte, Km 60, Bairro Litorâneo, 29.932-540, São Mateus, ES, Brazil. nunofariajr@gmail.com*

The use of DNA sequences, including DNA barcoding, as a taxonomical tool has been happening for some time (Tautz *et al.*, 2003; Hajibabaei *et al.*, 2007; Packer *et al.*, 2009). However, the description of new species based solely on DNA sequences is a new idea (Cook *et al.*, 2010) and a new practice (Brower, 2010). Our aim is not to further polarize the war between advocates of strictly molecular or strictly morphological systematics (following, *e.g.*, Pires & Marinoni 2010). The objectives here are (i) to present some arguments regarding the perils of the proposition of a model (theory) for solely DNA-based descriptions (Cook *et al.* 2010) and the actual publication (practice) of such descriptions (Brower 2010), (ii) to discuss some reasons why we believe that adopting strictly DNA taxonomy for species description, setting aside everything we have learned from classic taxonomy, may not be the best alternative and (iii) to present the point of view about these matters of a PhD candidate and a recently graduated PhD working with taxonomy in a developing country.

Critical theory of DNA-taxonomy

Cook *et al.* (2010:322) advocated there would be no compelling evidence to exclude DNA-only descriptions. In terms of nomenclatural rules, DNA-based descriptions are not in contradiction with the ICZN criteria for name availability (since these descriptions comply with publication adequacies). Nevertheless, the main question regards the significance of these taxa described. In accordance with Ebach & Carvalho (2010), we believe that small uncontextualized differences in sequences of DNA cannot necessarily define distinct taxa. As suggested some time ago by Darwin himself, “*No one supposes that all the individuals of the same species are cast in the same actual mould*” (Darwin 1872:34). A species is a hypothesis of relationship and its proposition relies on the study of a large number of specimens (exception for rare species, species where few individuals are collected or fossils) of the same or from different localities that allow a distinction between intraspecific and interspecific variation. We believe this is one of the most distinctive aspects of morphological based taxonomy in relation to molecular practice that usually covers data from one or few individuals as representative of a given species.

Cook *et al.* (2010) also claim that the classic taxonomic methodology is a virtual black box, that “*taxonomy remains as a ‘closed’ profession*” and that the “*descriptions are traditionally the territory of taxonomists*”. But the molecular approach for species definition does not constitute, in any way, a clearer methodology than the morphological one. The claim that a sequence is not influenced by subjective assessments ignores the difficulty of aligning sequences of different length, distinguishing paralogs from orthologs (exception for DNA barcoding). Selecting appropriate genes for any particular taxonomic study (Lipscomb *et al.* 2003) in addition to the choice of models and analyses for species delimitation can also generate different results (Brower 2006). In molecular taxonomy speech, it seems that the ease of identifying the character states (ATCG)—and consequently the ease of saying ‘this A is different from that T’—obscures the difficulty of establishing homologies among them. Furthermore, it is obvious that species description is the territory of taxonomists, as the structure of proteins regards biochemists or the study of a dinosaur fossil is the territory of paleontologists. Each field of biological sciences has its peculiarities and characteristics and the suitable training for practicing it. Cook *et al.* (2010:323) also claim that DNA sequencing is cheap and easy. But, in our opinion, the matter does not regard the economic viability of DNA-based taxonomy in relation to classic taxonomy, nor the price ratio between DNA sequencing and the purchase of stereomicroscopes. These apparent facilities provided by molecular databases may lead to a harmful robotization of the taxonomical science, where the questions exist just to be quickly solved. The use of DNA sequences for describing new species should be used as an auxiliary tool and treated under a phylogeographic context or through a revision study of taxa, preferably with a large number of individuals.

Critical practice of DNA-taxonomy

Brower's (2010) description of ten new *Astraptus* species illustrates the practical hazards of describing species based solely on molecular characters. Perhaps these new species may represent, at some level, just an inflation of available names of these taxa (see Brower, 2010) and not an increase in knowledge about the diversity of the genus. Only a morphological taxonomic study will be able to compare the name-bearing specimens of the recently described taxa with the previously described species. Considering that the access to molecular information of most of name-bearing specimens (responsible for the stability of taxonomic knowledge) is precarious nowadays, should we really proceed with solely DNA-based species descriptions? The only practical example of this attempt (Brower 2010) does not suggest so. The contribution of Brower (2010) seemed to create more problems than solutions when describing new species without comparing them to name-bearing specimens.

Future problems of nomenclature

The impossibility of immediate comparison among the new DNA species and most name-bearing specimens seems to us the most serious problem of strict molecular descriptions. Is it possible to determine if a new species proposed only by DNA sequences corresponds, in fact, to a synonymous species described in the 18th or 19th century? Will a simple collection of individuals in the type-locality (and sequencing them) solve the problem? But how to act in cases where closely related species are also sympatric or when the type locality is imprecise or unknown? Besides, the development of new techniques of DNA extraction and amplification to access molecular data from specimens collected long ago is limited (*e.g.*, Hundsdoerfer & Kitching 2010). The opposite problem may also take place, since the comparison between a given specimen and a species described solely based on DNA data would necessarily rely on the sequencing of this specimen. We strongly agree with Seberg *et al.* (2003): matching existing Linnaean names with DNA sequences is fraught with hazards. Fixing neotypes for known species in cases of unavailable genetic information from the original types was regarded as a feasible remedy (see Tautz *et al.* 2003). The first problem of this proposal regards the nomenclatural troubles that unbridled neotype designation would cause, particularly on historically neglected groups. The Code only allows the designation of neotypes in cases where name-bearing type specimens are lost or destroyed (article 75.3.4) besides seven other adequacies of fixation of neotypes (see article 75, ICZN 1999).

And even if the Code was respected and the formal proposition of neotypes does not occur, it wouldn't be appropriate if sequences of one or few individuals morphologically compared with name-bearing type specimens come to be assumed as representative of a given species. In this case, such individuals would work, in fact, as informal type specimens, since all subsequent molecular comparisons would involve them.

Is the molecular data better than others?

Since the proposition of a new species constitutes a hypothesis, it is most useful to consider the largest possible number and kinds of characters (morphological, geographical, behavioral, DNA sequences and others). In this sense, we do not agree with Cook *et al.* (2010), when the authors state that "taxonomy is restricted to morphological data". Along the history of taxonomy, new techniques, methodologies and data sources were incorporated to the taxonomist's toolbox. We follow Carvalho *et al.* (2007) view that web-initiatives and technology are necessary as long as they contribute to enhance the existing taxonomic enterprise, but do not aim at replacing it. DNA-taxonomy is unequivocally useful in particular situations such as immature-adults and male-female associations. DNA-taxonomy is just another technique for species identification and delimitation and its utility as a global identification system remains undisputed (Prendini 2005; Silva-Brandão *et al.* 2009). We endorse the claim by Seberg *et al.* (2003) that it would be a good idea to include DNA sequences in the diagnosis of taxa. This practice could even be recommended in a future edition of The Code. But mandatory introduction of DNA sequences into taxonomy seems to us a retrograde step (see Seberg *et al.* 2003).

Impediment to taxonomy

The main question in relation to DNA taxonomy in developing and underdeveloped countries refers to sociological limitations of DNA taxonomy (see Prendini 2005). An expensive and centralized DNA-based taxonomy might exclude many taxonomists who have limited access to sequencing technology (*e.g.*, Seberg *et al.* 2003) and funds for collecting expeditions. It would be a cruel irony if third-world taxonomists, much of them living in countries with enormous biodiversity and scarce science funding, found themselves placed apart from taxonomic science. Other sociological limitation regards the depositary of described species and the access to them. In a centralized and nowadays still expensive DNA-based descriptions paradigm, most of the taxonomic work would be carried out, at first, in developed countries with wide access to species sequencing. In this case, it is reasonable to believe that the deposition of type-specimens in countries different from where they were collected, as in past centuries, will be even more common.

Conclusions

We would like to reiterate that our purpose is not to reject the inclusion of DNA sequences in taxonomy, or even set fire in the existing cold war among defenders of strictly molecular or morphological based taxonomy. We only dread that the proposal of Cook *et al.* (2010)—and the subsequent contribution of Brower (2010)—give impetus to the description

of meaningless species by people untrained in taxonomy and presenting descriptions devoid of context in the classification of a given group. Therefore, as a complementary source of information for diagnosis and species descriptions, the proposal of Cook *et al.* (2010) would constitute a step forward. Like any other source of taxonomic data (including morphology), DNA sequences must be contextualized under a framework of training and expertise. Molecular information itself will certainly not solve all the problems of taxonomy.

ACKNOWLEDGMENTS

LMS acknowledges CNPq for a doctoral scholarship; LRRF acknowledges CNPq for a postdoctoral scholarship (PROTAX) received during part of the development of this contribution; we acknowledge E.A.B. Almeida and C.J.B. de Carvalho for revising and making valuable comments on the manuscript. O.H.H. Mielke and A.M. Sakakibara for clarifying some doubts we had on the interpretation of the Code. E.D.G. Soares, G.A.R. Melo, A.J.C. Aguiar, F.F. Salles, R.S. Bérnils for the relevant discussions that highly improved this manuscript and M. Sette for reviewing the style and grammar of the manuscript.

REFERENCES

- Brower, A.V.Z. (2006) Problems with DNA barcodes for species delimitation: 'ten species' of *Astraptes fulgerator* reassessed (Lepidoptera: Hesperidae). *Systematics and Biodiversity*, 4, 127–132.
- Brower, A.V.Z. (2010) Alleviating the taxonomic impediment of DNA barcoding and setting a bad precedent: names for ten species of 'Astraptes fulgerator' (Lepidoptera: Hesperidae: Eudaminae) with DNA-based diagnoses. *Systematics and Biodiversity*, 8, 485–491.
- Carvalho, M.R., Bockmann, F.A., Amorim, D.S., Brandão, C.R.F., de Vivo, M., Figueiredo, J.L., Britski, H.A., de Pinna, M.C.C., Menezes, N.A., Marques, F.P.L., Papavero, N., Cancellato, E.M., Crisci, J.V., McEachran, J.D., Schelly, R.C., Lundberg, J.G., Gill, A.C., Britz, R., Wheeler, Q.D., Stiassny, M.L.J., Parenti, L.R., Page, L.M., Wheeler, W.C., Faivovich, J., Vari, R.P., Grande, L., Humphries, C.J., DeSalle, R., Ebach, M.C. & Nelson, G.J. (2007) Taxonomic impediment or impediment to taxonomy? A commentary on systematics and the cybertaxonomic-automation paradigm. *Evolutionary Biology*, 34, 140–143.
- Cook, L.G., Edwards, R.D., Crisp, M.D. & Hardy, N.B. (2010) Need morphology always be required for new species descriptions? *Invertebrate Systematics*, 24, 322–326.
- Darwin, C. (1872) *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. Sixth Edition. John Murray, London, 458pp.
- Ebach, M.C. & Carvalho, M.R. (2010) Anti-intellectualism in the DNA barcoding enterprise. *Zoologia*, 27, 165–178.
- Hajibabaei, M., Singer, G.A.C., Hebert, P.D.N. & Hickey, D.A. (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23, 167–172.
- Hundsdoerfer, A.K. & Kitching, I.J. (2010) A method for improving DNA yield from century-plus old specimens of large Lepidoptera while minimizing damage to external and internal abdominal characters. *Arthropod Systematics & Phylogeny*, 68, 151–155.
- International Commission on Zoological Nomenclature. (1999) *International Code of Zoological Nomenclature*. Fourth Edition. The International Trust for Zoological Nomenclature, Padova, XXIX + 306 pp.
- Lipscomb, D., Platnick, N. & Wheeler, Q. (2003) The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology and Evolution*, 18, 65–66.
- Packer, L., Gibbs, J., Sheffield, C. & Hanner, R. (2009) DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources*, 9: 42–50.
- Pires, A.C. & Marinoni, L. (2010) DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica*, 10, 339–346.
- Prendini, L. (2005) Comment on "Identifying spiders through DNA barcodes". *Canadian Journal of Zoology*, 83, 498–504.
- Seberg, O., Humphries, C.J., Knapp, S., Stevenson, D.W., Petersen, G., Scharff, N. & Andersen, N.M. (2003) Shortcuts in systematics? A commentary on DNA-based taxonomy. *Trends in Ecology & Evolution*, 18, 63–65.
- Silva-Brandão, K.L., Lyra, M.L. & Freitas, A.V.L. (2009) Barcoding Lepidoptera: current situation and perspectives on the usefulness of a contentious technique. *Neotropical Entomology*, 38, 441–451.
- Tautz, D., Arctander, P., Minelli, A., Thomas, R.H. & Vogler, A.P. (2003) A plea for DNA taxonomy. *Trends in Ecology & Evolution*, 18, 70–74.