Crimson Publishers
Wings to the Research

Mini Review

Frequent Misclassification by mtDNA Barcoding as Revealed by nuDNA and/or Testable Analysis of its Expression Products

Bernhard Seifert*

Senckenberg Museum of Natural History Görlitz, Germany

ISSN: 2770-6745

***Corresponding author:** Bernhard Seifert, Senckenberg Museum of Natural History Görlitz, D-02826 Görlitz, Am Museum 1, Germany

Submission: January 25, 2024
Published: February 08, 2024

Volume 4 - Issue 4

How to cite this article: Bernhard Seifert*. Frequent Misclassification by mtDNA Barcoding as Revealed by nuDNA and/or Testable Analysis of its Expression Products. Biodiversity Online J. 4(4). BOJ. 000592. 2024.
DOI: 10.31031/BOJ.2024.04.000592

Copyright@ Bernhard Seifert. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract

Species identities are best indicated by analyses of nuclear DNA which is the material representing the working points of evolution. Additional good indicators of species identities are those expression products of nuclear DNA which are least modified by environmental influence and, as a consequence, show the highest correlation with the genetical core information. Among expression products such as morphology, products of metabolism and ethological or ecological traits, morphology is rated here as indicator with the highest correlation. The use of morphology as most important accessory indicator is furthermore favored by the leading position it played in species descriptions over 280 years of taxonomic research. Focusing on the example of ants, the paper considers 13 studies with parallel application of mtDNA barcoding, analysis of nuclear DNA and application of Numeric Morphology-Based Alpha-Taxonomy (NUMOBAT). Selected were only studies based on sufficiently high within-species numbers of samples. With nuclear DNA and NUMOBAT used as objective and testable control systems, the average classification error of mtDNA barcoding per sample or individual was 16.8% over 10 genera with 66 species with the extremes ranging from 0 to 32%. Ancient hybridization is considered a much more likely cause for mtDNA mismatches in ants than incomplete lineage sorting.

Keywords: Species classification; Ants; mtDNA barcoding; Nuclear DNA; Morphology; Based numeric; Alpha-taxonomy

Introduction

The year 2003 was the starting point for a determined campaign that was to have a strong and long-lasting impact on alpha-taxonomy. Hebert [1] presented the term “DNA barcoding” and praised it as a ready-to-use silver bullet for reliable species identification. They claimed that all the 10–25 million animal species on earth could be quickly recognized by large-scale screening of a mitochondrial DNA reference gene with comparably low costs. This idea of a turbo-taxonomic approach received an enormous echo ranging from top-ranking science journals such as Nature [2] to popular media such as The Times [3]. Massive counter-evidence for disagreement of mtDNA classifications with other indicators of species identity was presented by a meta-analysis of 584 studies of 526 eumetazoan genera already in the same year by Funk & Omland [4] who detected mtDNA paraphyly or polyphyly in 23% of 2319 assayed species. This and numerous follow-up papers revealing mismatches between species identities and mtDNA indication in the years since then could not stop the development of the global Consortium for the Barcoding of Life (CBOL) and its some 50 national offshoots. The wide application of mtDNA barcoding continued even after Ross [5] presented another similarly broad meta-analysis confirming the conclusions of Funk & Omland [4]. Caused by much improvement in analysis of nuclear DNA (nuDNA), both regarding methodology and costs, we currently observe an avertion of several biodiversity students from mtDNA barcoding but the general tenacity to adhere to this method remains astonishing. A frequent answer

in personal talks with convinced barcoders was that, though not denying occasional occurrence of paraphylies in many species, they deemed the frequency of misclassification on individual level to be negligibly low. Yet, are the frequencies really such low? Figuring out the true performance of Hebert's barcoding requires to fulfill three preconditions. Precondition 1 is that the alternative methods checking the classification by mtDNA are testable. Testability, or falsification and verification of classification hypotheses is only possible when they are based, in one or the other way, on numeric analyses but not on subjective idiosyncratic decision. Finding real frequencies furthermore requires sufficiently high within-species numbers of samples (precondition 2) and not juxtaposition of alternative classification systems in singletons as seen in the trees frequently published. Most important, or essential, for realistic assessment of barcoding performance is the that the controlling classification systems have the highest likelihood to indicate "true" species identities (precondition 3). In a paper introducing the Gene and Gene Expression (GAGE) species concept, [6] wrote Species are separable clusters that have passed a threshold of evolutionary divergence and are exclusively defined by nuclear DNA sequences and/or their expression products. Nuclear DNA sequences and their expression products are different character systems but have a highly correlated indicative function. Character systems with the least risk of epigenetic or ontogenetic modification have superior indicative value when conflicts between character systems of integrative studies arise.

In other words, "true" species identities are best indicated when classification methods focus on the working points of evolution and this is nuDNA and those expression products of nuDNA least modified by environmental influence and thus most strongly correlating with nuDNA. These correlations are highest in protein sequences and morphology, considerable in behavioral traits and products of metabolism, and weak in ecological traits [7,8]. Another argument to favor morphology among the expression products is the leading role it played in the history of taxonomy since Linnaeus. Condition 1 (testability) and condition 3 (indication close to "true" species identities) are best fulfilled by an approach or working philosophy named by Seifert, Numeric Morphology-Based-Alpha-Taxonomy (NUMOBAT). Using the example of ants (*Hymenoptera*: *Formicidae*), I present here the results of 13 studies being in agreement with the three preconditions outlined above. Such hard checks, either nuDNA- or NUMOBAT-based (or both in combination) are still very rare and I have the impression that the advocates mtDNA barcoding mentally displace such studies. The account below lists up the percentage of disagreeing classifications by mtDNA barcoding with the checking systems given in square brackets. Note that "zoogeography" was added as indicator in a case of a strictly parapatric distribution. The error was

0% in *Plagiolepis taurica* and *pyrenaica* [9] [nuDNA and NUMOBAT],

1% in *Temnothorax nylanderi* and *crassispinus* [10] [NUMOBAT, zoogeography],

6% in 4 species of the *Tapinoma nigerrimum* group [11] [NUMOBAT],

7% in 5 species of the *Cardiocondyla nuda* group [12] [NUMOBAT],

15% in *Formica pratensis* and *lugubris* [13] [NUMOBAT],

16% in *Cardiocondyla latifrons* and *micropila*. (Heinze pers. comm) [NUMOBAT],

17% in 10 species of *Tetramorium* [14] [NUMOBAT, nuDNA],

19% in 17 species of Neotropical *Linepithema* [15] [nuDNA],

20% in 8 species of *Serviformica* (Purcell pers. comm) [nuDNA],

21% in 6 species of the *Cataglyphis albicans* group [16] [nuDNA],

23% in 3 species of African *Cataglyphis* [17] [nuDNA, NUMOBAT],

24% in 3 species of Tibetan *Myrmica* [8] [NUMOBAT, nuDNA],

32% in 2 species of *Colobopsis* [18] [NUMOBAT].

Averaging these data, the result is sobering: the mean estimated misclassification per individual within 13 studies over 10 genera and 66 species is 16.8%. Wrong indications of Hebert's barcoding in the same range are supposed by subjective assessment of morphology for the genera *Anochoetus* and *Odontomachus* [19] and *Solenopsis* [20]. Furthermore, there is introgression of mtDNA into nuDNA-defined lineages in socially hybridogenetic ants [21]. In very broad study in Finland, Beresford [22] showed massive bidirectional introgression of heterospecific mtDNA into the populations of *Formica polyctena* and *aquilonia*. Looking at the data in the list above, there is clear trend that the lowest classification errors by mtDNA barcoding occur in species with parapatric zoogeography (the *Plagiolepis* and *Temnothorax* cases) or reduced frequency of outcrossing due to high frequency of intranidal mating (the *Tapinoma nigerrimum* and *Cardiocondyla nuda* group examples). In contrast, higher errors are more frequent in species groups performing extranidal mating or normal nuptial flights and having at least partially sympatric geographic ranges. These data support the idea that ancient hybridization with subsequent introgression of "wrong" matrilines is the most frequent source for mtDNA barcoding errors in ants whereas incomplete lineage sorting during species splitting is rarer and, in the cases reported here, probably responsible for the situation in the parapatric *Colobopsis* species.

Rererences

1. Hebert PDN, Cywinska A, Ball SL, Waard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270: 313-321.
2. Blaxter M (2003) Counting angels with DNA. Nature 421: 122-124.
3. Hendersen M (2005) Is it a bird? Is it a plant? Hang on, I'll check its barcode. The Times.
4. Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology and Evolutionary Systematics 34: 397-423.

Copyright © Bernhard Seifert

5. Ross HA (2014) The incidence of species-level paraphyly in animals: A re-assessment. Mol Phylogenet Evol 76: 10-17.

6. Seifert B (2020) The gene and gene expression (gage) species concept: An universal approach for all eukaryotic organisms. Syst Biol 69(5): 1033-1038.

7. Schlick SBC, Steiner FM, Seifert B, Stauffer C, Christian E, et al. (2010) Integrative taxonomy: A multi-source approach to exploring biodiversity. Annu Rev Entomol 55: 421-438.

8. Seifert B, Schultz R, Ritz MS, Ritz CM (2018) Cryptic species of the Myrmica tibetana complex (*Hymenoptera*: *Formicidae*) revealed by integrative taxonomy. Myrmecological News 27: 1-18.

9. Kirchner P, Seifert B, Kröll J, Florian MS, Schlick SBC, et al. (2023) Phylogenomic inference and demographic model selection suggest peripatric separation of the cryptic steppe ant species Plagiolepis pyrenaica stat. rev. Molecular Ecology 32: 1149-1168.

10. Pusch K, Seifert B, Foitzik S, Heinze J (2006) Distribution and genetic divergence of two parapatric sibling ant species in Central Europe. Biological Journal of the Linnean Society 88: 223-234.

11. Seifert B, Eustacchio D, Kaufmann B, Centorame M, Lorite P, et al. (2017a) Four species within the supercolonial ants of the *Tapinoma nigerrimum* complex revealed by integrative taxonomy (*Hymenoptera*: *Formicidae*). Myrmecological News 24: 123-124.

12. Seifert B, Okita I, Heinze J (2017b) A taxonomic revision of the *Cardiocondyla* nuda group (*Hymenoptera*: *Formicidae*). Zootaxa 4290(2): 324-356.

13. Seifert B, Goropashnaya A (2004) Ideal phenotypes and mismatching haplotypes errors of mtDNA treeing in ants (*Hymenoptera*: *Formicidae*) detected by standardized morphometry. Organisms, Diversity & Evolution 4(4): 295-305.

14. Wagner HC, Arthofer W, Seifert B, Muster C, Steiner FM, et al. (2017) Light at the end of the tunnel: Integrative taxonomy delimits cryptic species in the *Tetramorium caespitum* complex (*Hymenoptera*: *Formicidae*). Myrmecological News 25: 95-129.

15. Wild AL (2008) Evolution of the neotropical ant genus *linepithema*. Systematic Entomology 34: 49-62.

16. Eyer PA, Hefetz A (2018) Cytonuclear incongruences hamper species delimitation in the socially polymorphic desert ants of the *Cataglyphis albicans* group in Israel. Journal of Evolutionary Biology 31(12): 1828-1842.

17. Knaden M, Tinaut A, Cerdá X, Wehner S, Wehner R (2005) Phylogeny of three parapatric species of desert ants, *Cataglyphis bicolor*, C. viaticus, and *C. savignyi*: a comparison of mitochondrial DNA, nuclear DNA, and morphometric data. Zoology 108: 169-177.

18. Schifani E, Gianetti D, Czösz S, Castellucci F, Luchetti A, et al. (2021) Is mimicry a diversification-driver in ants? Biogeography, ecology, ethology, genetics and morphology define a second West-Palaearctic *Colobopsis* species (*Hymenoptera*: *Formicidae*). Zoological Journal of the Linnean Society 20: 1-27.

19. Fisher BL, Smith MA (2008) A revision of Malagasy species of Anochetus Mayr and *Odontomachus Latreille* (*Hymenoptera*: *Formicidae*). Public Library of Science One 3(5): e1787.

20. Shoemaker D, Ahrens ME, Ross KG (2006) Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. Molecular Phylogenetics and Evolution 38(1): 200-215.

21. Darras H, Aron S (2015) Introgression of mitochondrial DNA among lineages in a hybridogenetic ant. Biological Letters 11: 20140971.

22. Beresford J, Elias M, Pluckrose L, Sundström L, Butlin RK, et al. (2017) Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. Mol Ecol 26(15): 4013-4026.

 Copyright © Bernhard Seifert