Biologia
https://doi.org/10.1007/s11756-024-01622-8

ORIGINAL ARTICLE

DNA barcoding and genetic distances in three genera of Naididae (Annelida: Clitellata)

Sanjib Chakma[1] · Svante Martinsson[2] · Mohammed Ibrahim Naveed[1]

Received: 3 April 2023 / Accepted: 30 January 2024

Abstract
Identification of organism groups such as Oligochaetes to species level is difficult to achieve using morphological examination alone. This is especially true in species where sexual reproduction is rare, such as the subfamily Naidedae. Here, variation in the barcoding marker COI of three genera of freshwater clitellates, *Dero*, *Nais*, and *Pristina* (Clitellata: Naididae) was studied, using publicly available sequences together with newly generated ones, with the aim to test for the presence of global barcode gaps. In total 17 sequences of *Dero* spp., 168 of *Nais* spp., and 22 sequences of *Pristina* spp. were included in the analyses. The uncorrected pairwise distances in the *Dero* dataset ranges from 0.00 to 0.18, in the *Nais* dataset they range from 0.00 to 0.21, and in the *Pristina* dataset they range from 0.00 to 0.36, with a large gap between 0.21 and 0.35. No global barcoding gap was found in any of the datasets. In all three genera clusters including more than one species and/or species found in more than one cluster were found, indicating taxonomical problems, such as cryptic species and misidentified sequences.

Keywords *Dero* · *Nais* · *Pristina* · COI · India · Oligochaeta

Introduction

There are many organism groups that are difficult to identify to species level through morphologically examination, and where incorrect identification are common. Oligochaetes (Annelida: Clitellata) are one of such group especially in their immature stage. Aquatic oligochaetes, including the family Naididae, have traditionally been considered a taxonomically difficult group (e.g., Timm 2009). Species identification in many groups of oligochaetes is based on differences in their reproductive organs, which makes it hard to identify immature specimens. There are also several groups, e.g., the subfamily Naidinae (Clitellata: Naididae), where sexual reproduction is rare. There are several cases, within Naididae, where distinguishing between species is difficult e.g., *Pristina foreli* (Piguet, 1906) and *Pristina aequiseta* Bourne, 1891 are only distinguishable due to the presence of giant chaetae in VI segment of *P. aequiseta*, in such situations DNA-barcoding is a good option to reach a correct identification.

DNA-barcoding is the matching of short standardized genetic markers for identification. For animals, the mitochondrial gene cytochrome c oxidase I (COI) is the standard marker (Hebert et al. 2003a). Early on, a fixed threshold for genetic distances was often used to determine if sequences were conspecific or belonged to separate species (e.g., Hebert et al. 2003b; Smith et al. 2005; Waugh 2007). However, this has been shown to be problematic, and if thresholds should be used, they should be generated from the data (see discussion in Collins and Cruickshank 2013). The main assumption for using a global threshold is the existence of a global barcoding gap, where the intraspecific, i.e., within species, are lower than the interspecific, i.e., between species, distances across the whole dataset (Collins and Cruickshank 2013). This does not exist in many groups, including Annelida (Kvist 2016), there are many instances with intraspecific distances > 10% e.g., the earthworms *Lumbricus terrestris* Linnaeus, 1758 and *Eisenia fetida*

✉ Svante Martinsson
svante880@gmail.com

[1] P.G & Research Department of Zoology, The New College, Chennai, India

[2] Stenkullen, Sweden

Published online: 05 March 2024

(Savigny, 1826) (Martinsson and Erséus 2017, 2018) and interspecific distances < 5% e.g., between species of *Branchiodrilus* Michaelsen, 1900 (Martin et al. 2018). Further consideration for reliably using DNA-barcoding for identifying specimens is the reference library used, this consists of already identified sequences to which the new sequences are compared. Both having enough sequences of each species, as well has having representation across its range is important to cover the genetic variation (Papadopoulou et al. 2008; Luo et al. 2015; Zhang et al. 2019). DNA-barcoding may be particularly useful for easy identification of species which are rare, fragile, or small and are difficult to identify morphologically, and therefore more likely to be misidentified or left unidentified if only morphology is used. DNA-barcoding can further reveal taxonomically significant geographic variation and cryptic species (Bucklin et al. 2011). The application of DNA barcoding can help to distinguish species with similar morphologies as well as resolve the status of subspecies and morphologically cryptic species in addition to being used for routine identification of specimens. In addition to that, DNA barcodes provide potentially useful information for groups that are already well studied, and such taxa do not constitute the majority of biodiversity or those in most need of research attention (Rubinoff 2006).

The aims of this work are to add new barcode sequences of naidid worms from India, test for the presence of a barcode gap and study the variation in COI of three genera of Naididae, *Dero* Oken, 1815, *Nais* Müller, 1774, and *Pristina* Ehrenberg, 1831 using publicly available sequences from GenBank, together with the newly generated sequences, from India.

Materials and methods

Publicly available COI sequences of the genera *Dero*, *Nais*, and *Pristina* were downloaded from GenBank (Accessed 2022-03-17), sequences shorter than 500 bp were discarded. In addition, for all sequences the species name given in GenBank were used. Three new sequences of *Dero dorsalis* Ferronnière, 1899, *Nais pseudobtusa* Piguet, 1906, and *Pristina foreli* were generated. The newly sequenced specimens were collected from three different locations of Tamil Nadu. *Pristina foreli* was collected from Kamaraj Sagar, Ooty (11°26‘21.2’’N, 76°39‘37.7’’E), *N. pseudobtusa* was collected from Pykara Lake, Ooty (11°27‘53.8’’N, 76°36‘21.2’’E), and *D. dorsalis* was collected from Kunyamuthur Lake, Coimbatore (10°57‘45.4’’N, 76°58‘09.1’’E). The specimens were preserved in absolute alcohol. The identification of specimens was performed according to Brinkhurst and Jamieson (1971), Naidu (2005), and Timm (2009). Isolation of DNA and amplification was performed at BioEdge Solutions, Bangalore, using standard Qiagen Kit method. The standard primers LCO1490 and HCO2198 (Folmer et al. 1994) were used for COI gene amplification. For PCR analysis the following program was followed initial denaturation 95 °C for 2 min, final denaturation 95 °C with 30 cycles 30 s, annealing 50 °C for 30 s and final termination or extension finishes at 60 °C for 4 min. The COI gene obtained of the three nadids was deposited in GenBank *Pristina foreli* (accession no. OL374056), *Nais pseudobtusa* (accession no. OL374070), *Dero dorsalis* (accession no. OL375209).

Separate alignments were produced for each genus, using MAFFT v7.017 (Katoh et al. 2002), as implemented in Geneious Pro v. 7.1, using the auto-algorithm and default settings, the alignments were trimmed to the only cover the standard barcoding region with a length of 658 bp. For the phylogenetic analysis outgroups were added, for the *Dero* dataset *Branchiodrilus cleistochaeta* Dalh, 1957 was added, for the *Nais* dataset *Amphichaeta sannio* Kallstenius, 1892 was added, and for the *Pristina* dataset *Trieminentia corderoi* (Harman, 1969) was added, the outgroups were selected based on earlier studies (Erséus et al. 2010, 2017; Martin et al. 2018). In total 17 sequences of *Dero*, 168 of *Nais*, and 22 sequences of *Pristina* + one outgroup in each alignment were included in the analysis.

Distance analyses

Uncorrected genetic p-distances were calculated for the datasets, without outgroups, in MEGA X (Kumar et al. 2018), using pairwise deletion for missing data, the distances were categorised as either intra-, interspecific or unidentified, based on the identification given to the sequences in GenBank, were a pairwise distance were one or both of the sequences were not identified to species was classified as unidentified, if both were identified as the same species it were classified as intraspecific and the sequences were identified as belonging to two different identified species it was classified as interspecific. Further, for the morphotypes in the *Nais communis/variabilis* complex identified in (Envall et al. 2012), we followed their conclusion of species delimitation, but classified the distances between members of the morphotypes and either *N. communis* Piguet 1906 or *N. variabilis* Piguet 1906 as unidentified. The distances were analysed and summarised in histograms, using R v.3.6.3. (R Core Team 2020) with the packages ggplot2 (Wickham 2016), dplyr (Wickham et al. 2021), and readr (Wickham and Hester 2020).

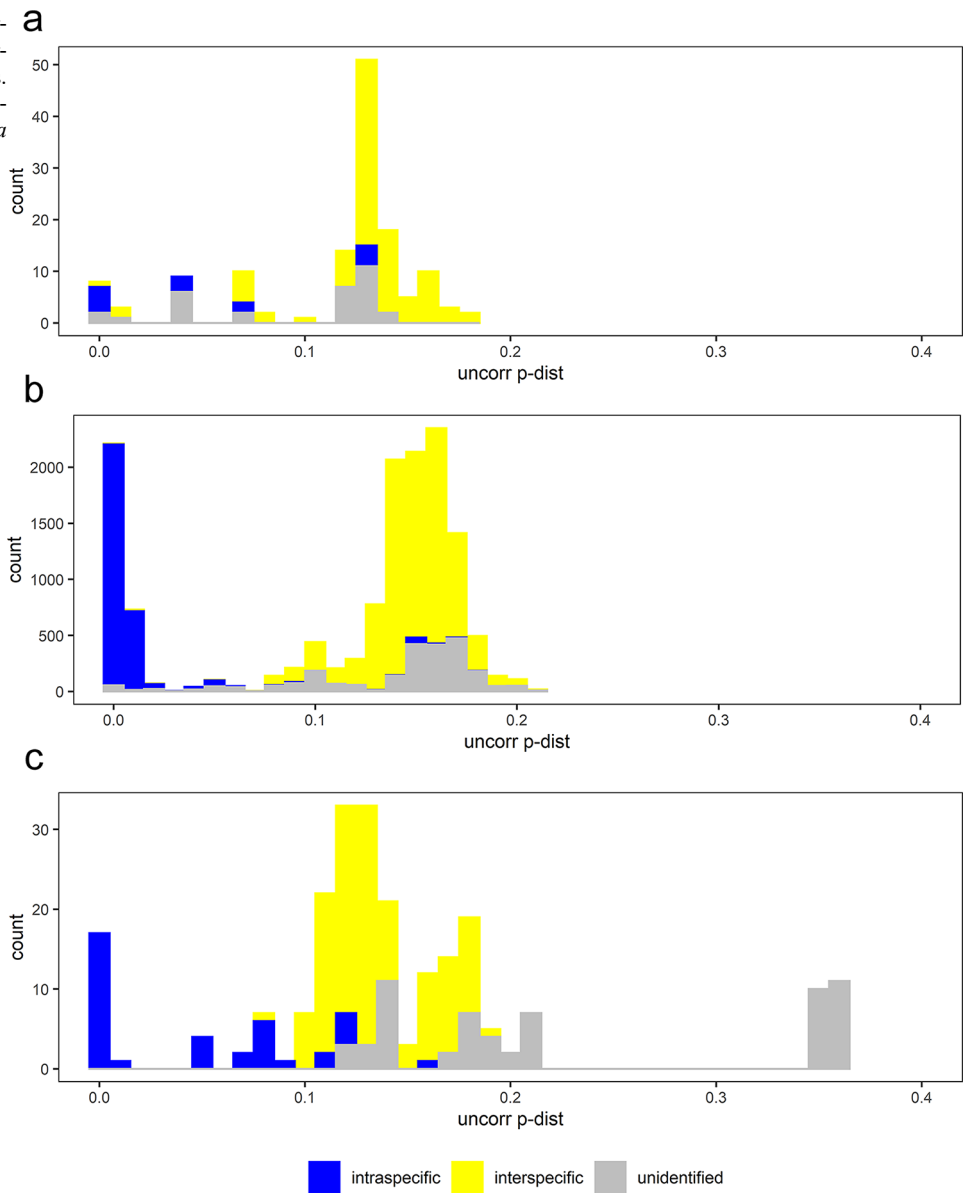

Fig. 1 Histograms of uncorrected pairwise genetic distances in COI. **a** Pairwise distances for *Dero* spp. specimens. **b** Pairwise distances for *Nais* spp. specimens. **c** Pairwise distances for *Pristina* spp. specimens

Phylogenetic estimations

Phylogenies were estimated with Maximum Likelihood using phyML 3.0 (Guindon et al. 2010), Smart Model Selection (Lefort et al. 2017) with Bayesian Information Criterion was used for automatic model selection; and Nearest Neighbour Interchange was used for tree improvement. Branch support was calculated with the SH-like (Shimodaira-Hasegawa test-like) approximative Likelihood Ratio Test (aLRT) (Anisimova and Gascuel 2006). The trees were drawn in FigTree 1.4.2 (Rambaut 2014).

The alignments used for the trees, the distance files, and R-script are available at https://github.com/Svante-Martinsson/Barcoding_dist_Naididae.

Results

Distance analyses

In the *Dero* dataset there were 136 uncorrected pairwise distances between 17 sequences. The distances ranges from 0.00 to 0.18 (Fig. 1a). In the *Nais* dataset there were 14,028 uncorrected pairwise distances between 168 sequences. The distances ranges from 0.00 to 0.21 (Fig. 1b). In the *Pristina* dataset there were 231 uncorrected pairwise distances between 22 sequences. The distances ranges from 0.00 to 0.36 (Fig. 1c), with a large gap between 0.21 and 0.35, all the distances above the gap are between a sequence from an unidentified specimen from Australia (GenBank Acc.no MW021259) and the

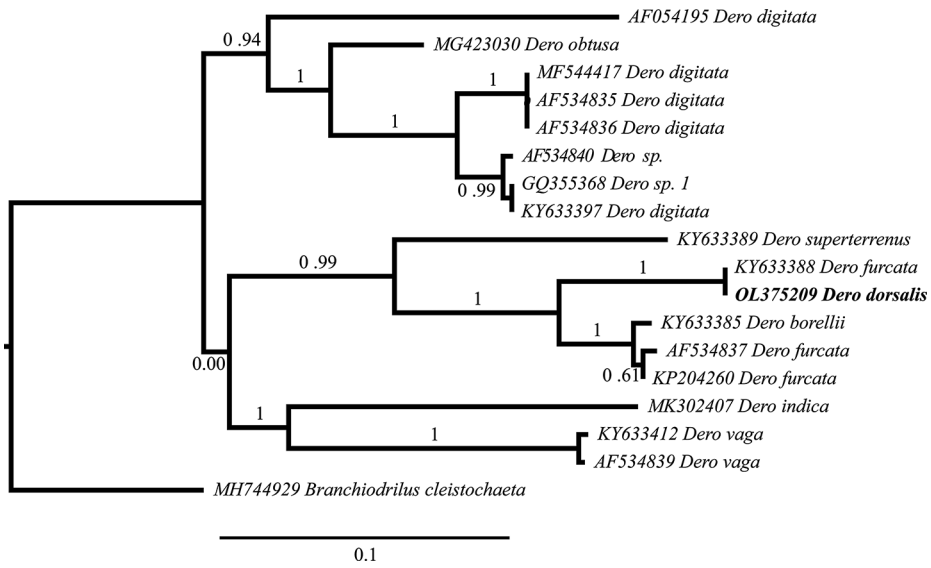

Fig. 2 Phylogeny of *Dero*, based on COI data from GenBank and newly generated sequence (bold). The phylogeny is estimated using ML in phyML. Numbers at branches are SH-like aLRT support values. Scale shows expected number of changes per site

remaining sequences. In all three datasets, the intraspecific distances are generally smaller than the interspecific distances. However, the overlap is large in all datasets.

Phylogenetic analyses

In the tree resulting from the analysis of the *Dero* dataset (Fig. 2) nine distinct lineages of *Dero* are recovered, of them two contained specimens identified as different species, one with *D. borellii* Michaelsen, 1900 and *D. furcata* Oken, 1815, and one with *D. dorsalis* and *D. furcata*. Further, specimens identified as *D. digitata* (Müller, 1774) are found in three distinct lineages, one also including unidentified specimens, two of these lineages are sister-groups.

In the tree resulting from the analysis of the *Nais* dataset (Fig. 3) 19 distinct lineages of *Nais* are recovered, one of these contained a mix of specimens identified as *N. chistinae* Kasprzak, 1973 and *N.variabilis.* Further, specimens identified as *N. communis* are found in five linages, specimen identified as *N. elinguis* Müller, 1774 in two, and specimens identified as *N. variablis* in two lineages.

In the tree from the analysis of the *Pristina* dataset (Fig. 4) 14 distinct linages of *Pristina* are recovered, none of them contained a mix of different species. Further, specimens identified as *P. aequiseta* are found in four lineages, but three of these forms a single clade, and both *P. longiseta* Ehrenberg, 1831 and *P. leidyi* Smith, 1896 are found in two lineages each, the *P. longiseta* forms a clade together with one *P. leidyi*. Due to the large genetic distances between specimen MW021259 and remaining *Pristina* specimens its branch is shortened in Fig. 4.

Discussion

The distances in the tree genera are in general up to about 20%, except for one specimen in *Pristina*, that differ from the remaining specimens with more than 30%, indicating that it may be misidentified, and actually belong to a different genus. In all three genera, there is no clear separation between intra- and interspecific distances, this could be explained, at least to some extent by misidentified specimens and cryptic species. Misidentifications in published sequences is a known problem (e.g., Nilsson et al. 2006; Kunprom and Pramual 2019), and cryptic species are known to be common in annelids as well as in other animal groups (Pfenninger and Schwenk 2007; Nygren 2014; Martinsson and Erséus 2021). There are no clear global barcoding gaps, but for both *Dero* and *Pristina*, the number of sequences is quite low, which make the distribution of pairwise distances less defined. In *Nais*, where the number of sequences is larger, two clear peaks can be seen, one around 0–1% and around 13–18%, but there are distances covering the whole gap between the two peaks. The same pattern with a lack of a global barcoding gap has been observed in Annelida as a whole (Kvist 2016).

In *Dero*, the taxonomical problems regard *D. digitata*, which is found in three separate lineages, and *D. furcata* that are found in two separate lineages. In both lineages there are specimens of other species, one has a specimen identified as *D. borellii* and in the other the newly sequenced *D. dorsalis* from India, most of these problems were already shown in a previous study (Srinivasan et al. 2020). Further, some of the *Dero* species have been included in the genus *Aulophorus*, but this genus is often treated as a synonym or subgenus of *Dero*, and the reciprocal monophyly of *Aulophorus* and *Dero* s.str has not been demonstrated (e.g., Bely and Wray

Fig. 3 Phylogeny of *Nais*, based on COI data from GenBank and newly generated sequence (bold). The phylogeny is estimated using ML in phyML. Numbers at branches are SH-like aLRT support values. Scale shows expected number of changes per site

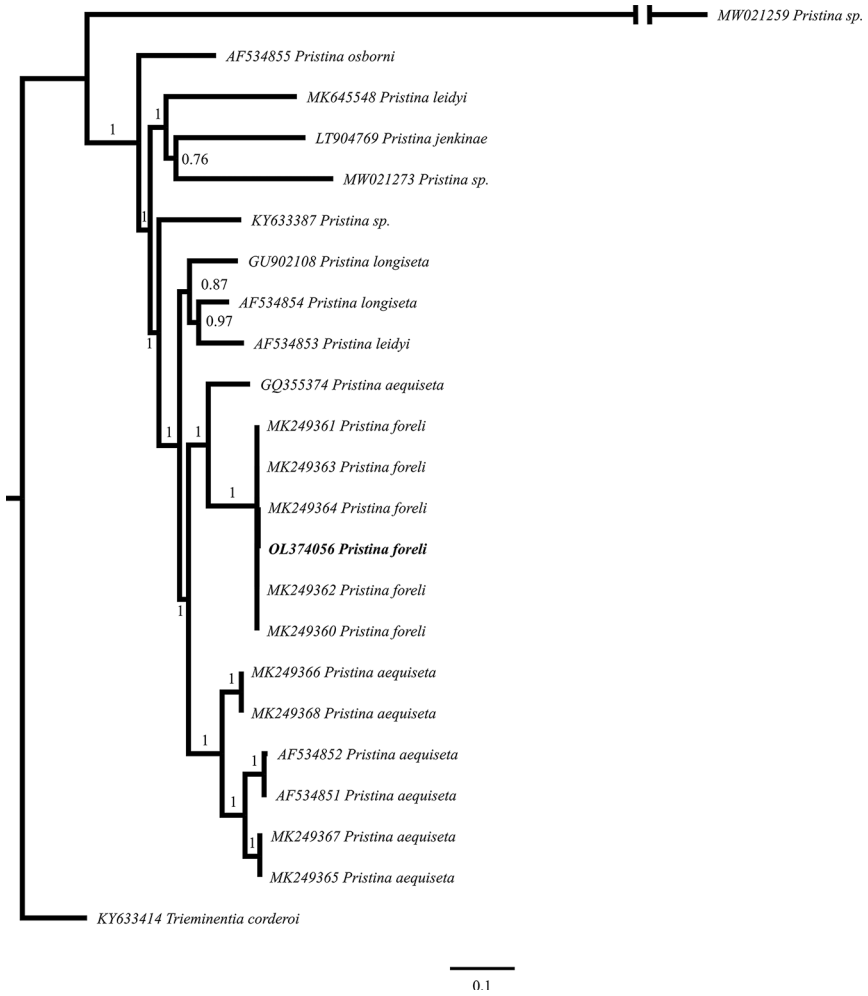

Fig. 4 Phylogeny of *Pristina*, based on COI data from GenBank and newly generated sequence (bold). The phylogeny is estimated using ML in phyML. Numbers at branches are SH-like aLRT support values. Scale shows expected number of changes per site, note that the branch for MW021259 *Pristina* sp. is shorten

2004; Erséus et al. 2017; Srinivasan et al. 2020), therefore we treat *Aulophorus* as a junior synonym in this paper.

In *Nais* the problems mainly concern the *N. communis/variabilis* complex, which were studied in detail by Envall et al. (2012), members of this complex are found in 10 lineages, two pairs of lineages were shown to belong to the same species by (Envall et al. 2012). The newly sequenced *N. pseudobtusa* from India is found together with a conspecific.

In *Pristina*, potential cryptic species are found in *P. aequiseta*, *P. longiseta*, and *P. leidyi*, which all are found in more than one lineage. The newly sequences *P. foreli* is found together with conspecifics. Morphological similarity between *P. foreli* and *P. aequiseta* are limited to the presence of giant ventral chaetae in *P. aequiseta* in VI segment (Jelinek and Rhaesa 2018). Our analysis supports that *P. foreli* and *P. aequiseta* are closely related.

Even though it has been shown that COI, and other mitochondrial markers, alone often overestimate the numbers of species if used alone, and therefore cannot be used on its own to delimit species (e.g., King et al. 2008; Dasmahapatra et al. 2010; Achurra and Erséus 2013; Martinsson et al. 2013; Martinsson et al. 2017; Martinsson et al. 2020), it still useful for identify potential cryptic species and other taxonomical problems, that require further studies. Here, the use of additional markers will be required for future studies to separate the cryptic species in *Dero*, *Nais* and *Pristina*.

In this study, we have shown that there is no global barcoding gap in any of the three studied genera. We also identify taxonomical problems in the threes studied genera, which needs further study. Lastly, we have added additional barcodes from India, increasing sampling form the area. We hope that this study is only a beginning towards creating a barcoding library for Indian freshwater clitellates.

Acknowledgements We thank The New College Management, Principal and Head, Dept of Zoology for providing the facilities for carrying out this work. We are grateful to BioEdge Solutions, Bangalore for their support in DNA barcoding. We thank Dr. Kerryn Elliott for her constructive comments and suggestions. We thank all the teaching, non-teaching staff and research scholars in the Department of Zoology, The New College, Chennai for their help and encouragement.

Author contributions All authors contributed to the study conception and design. Specimen collection, handling and identification were performed by SC, and MIN. Analyses were performed by SM, all authors contributed to the draft and read and approved the final manuscript.

Funding No funding was obtained for the study.

Declarations

Ethical approval No ethical approval was needed for the study.

Informed consent No humans were studied, and no informed consent was needed.

Conflict of interest The authors report that there are no competing interests to declare.

References

Achurra A, Erséus C (2013) DNA barcoding and species delimitation: the *Stylodrilus heringianus* case (Annelida: Clitellata: Lumbriculidae). Invertebr Syst 27:118–128. https://doi.org/10.1071/Is12049

Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst Biol 55:539–552. https://doi.org/10.1080/10635150600755453

Bely AE, Wray GA (2004) Molecular phylogeny of naidid worms (Annelida: Clitellata) based on cytochrome oxidase I. Mol Phylogenet Evol 30:50–63. https://doi.org/10.1016/s1055-7903(03)00180-5

Brinkhurst RO, Jamieson BGM (1971) Aquatic Oligochaeta of the world. Oliver and Boyd, Edinburgh

Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. Annu Rev Mar Sci 3:471–508. https://doi.org/10.1146/annurev-marine-120308-080950

Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. Mol Ecol Resour 13:969–975. https://doi.org/10.1111/1755-0998.12046

Dahl IO (1957) Results from the Danish expedition to the French Cameroons, 1949–1950. XXII Oligochaeta BIFAN 19:1154–1172

Dasmahapatra KK, Elias M, Hill RI, Hoffman JI, Mallet J (2010) Mitochondrial DNA barcoding detects some species that are real, and some that are not. Mol Ecol Resour 10:264–273. https://doi.org/10.1111/j.1755-0998.2009.02763.x

Ehrenberg CG (1828–1831) Phytozoa turbellaria Africana et Asiatica in Phytozoorum Tabula IV Et V delineata. In: Hemprich FG, Ehrenberg CG (eds) Symbolae physicae, seu icones et descriptiones corporum naturalium novorum aut minus cognitorum quae ex itineribus per libyam, Aegyptium, Nubiam, Dongalum, Syriam, Arabiam, et habessiniam, pars zoologica II, animalia evertebrata exclusis insecti. Officina Academica, Berlin, pp 53–67 + Plates IV-V

Envall I, Gustavsson LM, Erséus C (2012) Genetic and chaetal variation in *Nais* worms (Annelida, Clitellata, Naididae). Zool J Linn Soc 165:495–520. https://doi.org/10.1111/j.1096-3642.2012.00828.x

Erséus C, Envall I, Marchese M, Gustavsson L (2010) The systematic position of Opistocystidae (Annelida, Clitellata) revealed by DNA data. Mol Phylogenet Evol 54:309–313. https://doi.org/10.1016/j.ympev.2009.09.037

Erséus C, Envall I, De Wit P, Gustavsson LM (2017) Molecular data reveal a tropical freshwater origin of Naidinae (Annelida, Clitellata, Naididae). Mol Phylogenet Evol 115:115–127. https://doi.org/10.1016/j.ympev.2017.07.016

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010

Harman WJ (1969) Revision of the family Opistocystidae (Oligochaeta). Trans Am Microsc Soc 88:472–478

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. Proc R Soc Biol Sci Ser B 270:313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert PD, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Biol Sci 270 Suppl 1S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Jelinek H, Rhaesa AS (2018) Taxonomic investigations on Naididae: *Pristina aequiseta* Bourne, 1891 and *Pristina forel*i (Piguet, 1906) (Annelida, Oligochaeta, Naididae) are distinct species. Verh Naturwiss Ver Hamburg 51:53–71

Kallstenius E (1892) Eine Neue Art Der Oligochaetengattung *Amphichaeta*. Biol Fören Förh 4:42–55

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier trasform. Nucl Acids Res 30:3059–3066

King RA, Tibble AL, Symondson WO (2008) Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. Mol Ecol 17:4684–4698. https://doi.org/10.1111/j.1365-294X.2008.03931.x

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Kunprom C, Pramual P (2019) DNA barcoding of fruit flies (Diptera: Tephritidae) in Thailand: ambiguity, misidentification and cryptic diversity. Mitochondrial DNA DNA Mapp Seq Anal 30:861–873. https://doi.org/10.1080/24701394.2019.1693550

Kvist S (2016) Does a global DNA barcoding gap exist in Annelida? Mitochondrial DNA DNA Mapp Seq Anal 27:2241–2252. https://doi.org/10.3109/19401736.2014.984166

Lefort V, Longueville J-E, Gascuel O (2017) SMS: Smart Model Selection in PhyML. Mol Biol Evol 34:2422–2424. https://doi.org/10.1093/molbev/msx149

Luo A, Lan H, Ling C, Zhang A, Shi L, Ho SY, Zhu C (2015) A simulation study of sample size for DNA barcoding. Ecol Evol 5:5869–5879. https://doi.org/10.1002/ece3.1846

Martin P, Martinsson S, Wuillot J, Erséus C (2018) Integrative species delimitation and phylogeny of the branchiate worm *Branchiodrilus* (Clitellata, Naididae). Zool Scr 47:727–742. https://doi.org/10.1111/zsc.12316

Martinsson S, Erséus C (2017) Cryptic speciation and limited hybridization within *Lumbricus* earthworms (Clitellata: Lumbricidae). Mol Phylogenet Evol 106:18–27. https://doi.org/10.1016/j.ympev.2016.09.011

Martinsson S, Erséus C (2018) Hybridisation and species delimitation of scandinavian *Eisenia* spp. (Clitellata: Lumbricidae). Eur J Soil Biol 88:41–47. https://doi.org/10.1016/j.ejsobi.2018.06.003

Martinsson S, Erséus C (2021) Cryptic Clitellata: molecular species delimitation of clitellate worms (Annelida): an overview. Diversity 13:36. https://doi.org/10.3390/d13020036

Martinsson S, Achurra A, Svensson M, Erséus C (2013) Integrative taxonomy of the freshwater worm *Rhyacodrilus Falciformis* s.l.

(Clitellata: Naididae), with the description of a new species. Zool Scr 42:612–622. https://doi.org/10.1111/zsc.12032
Martinsson S, Rhodén C, Erséus C (2017) Barcoding gap, but no support for cryptic speciation in the earthworm *Aporrectodea longa* (Clitellata: Lumbricidae). Mitochondrial DNA 28:147–155. https://doi.org/10.3109/19401736.2015.1115487
Martinsson S, Klinth M, Erséus C (2020) Testing species hypotheses for *Fridericia magna*, an enchytraeid worm (Annelida: Clitellata) with great mitochondrial variation. BMC Evol Biol 20:116. https://doi.org/10.1186/s12862-020-01678-5
Naidu KV (2005) The Fauna of India and the adjacent countries-aquatic Oligochaeta. Zoological Survey of India, Kolkata
Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, Koljalg U (2006) Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS ONE 1:e59. https://doi.org/10.1371/journal.pone.0000059
Nygren A (2014) Cryptic polychaete diversity: a review. Zool Scr 43:172–183. https://doi.org/10.1111/zsc.12044
Papadopoulou A, Bergsten J, Fujisawa T, Monaghan MT, Barraclough TG, Vogler AP (2008) Speciation and DNA barcodes: testing the effects of dispersal on the formation of discrete sequence clusters. Philos Trans R Soc Lond B Biol Sci 363:2987–2996. https://doi.org/10.1098/rstb.2008.0066
Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol Biol 7:121. https://doi.org/10.1186/1471-2148-7-121
R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
Rambaut A (2014) FigTree v1.4.2
Rubinoff D (2006) Utility of mitochondrial DNA barcodes in species conservation. Conserv Biol 20:1026–1033. https://doi.org/10.1111/j.1523-1739.2006.00372.x
Smith MA, Fisher BL, Hebert PD (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos Trans R Soc Lond B Biol Sci 360:1825–1834. https://doi.org/10.1098/rstb.2005.1714
Srinivasan S, Martinsson S, Naveed MI (2020) On the identity and phylogenetic position of *Dero indica* (Clitellata: Naididae). Biologia 75:1685–1689. https://doi.org/10.2478/s11756-020-00422-0
Timm T (2009) A guide to the freshwater Oligochaeta and Polychaeta of Northern and Central Europe. Lauterbornia 66:1–235
Waugh J (2007) DNA barcoding in animal species: progress, potential and pitfalls. BioEssays 29:188–197. https://doi.org/10.1002/bies.20529
Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York
Wickham H, Hester J (2020) readr: Read Rectangular Text Data. R package version 1.4.0
Wickham H, François R, Henry L, Müller K (2021) dplyr: A Grammar of Data Manipulation. R package version 1.0.6
Zhang ZP, Wang XY, Zhang Z, Yao H, Zhang XM, Zhang Y, Zhang BG (2019) The impact of genetic diversity on the accuracy of DNA barcoding to identify species: a study on the genus *Phellodendron*. Ecol Evol 9:10723–10733. https://doi.org/10.1002/ece3.5590