



HAL
open science

Challenges of molecular barcode-based identification of earthworm specimens for biodiversity assessment

Andréa Barraux, Shanèze Noël, Virginie Roy, Lise Dupont

► To cite this version:

Andréa Barraux, Shanèze Noël, Virginie Roy, Lise Dupont. Challenges of molecular barcode-based identification of earthworm specimens for biodiversity assessment. *Frontiers in Ecology and Evolution*, 2024, 12, pp.1358984. 10.3389/fevo.2024.1358984 . hal-04639447

HAL Id: hal-04639447

<https://hal.u-pec.fr/hal-04639447>

Submitted on 9 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN ACCESS

EDITED BY
Manqiang Liu,
Lanzhou University, China

REVIEWED BY
Régis Vivien,
Swiss Ecotox Center, Switzerland
Jing Sun,
Chengdu Technological University, China

*CORRESPONDENCE
Lise Dupont
✉ lise.dupont@u-pec.fr

RECEIVED 20 December 2023
ACCEPTED 14 June 2024
PUBLISHED 08 July 2024

CITATION
Barraux A, Noël S, Roy V and Dupont L (2024)
Challenges of molecular barcode-based
identification of earthworm specimens for
biodiversity assessment.
Front. Ecol. Evol. 12:1358984.
doi: 10.3389/fevo.2024.1358984

COPYRIGHT
© 2024 Barraux, Noël, Roy and Dupont. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Challenges of molecular barcode-based identification of earthworm specimens for biodiversity assessment

Andréa Barraux, Shanèze Noël, Virginie Roy and Lise Dupont*

Univ Paris Est Creteil, Sorbonne Université, Université Paris-Cité, CNRS, IRD, INRAE, Institut d'écologie et des sciences de l'environnement de Paris, Créteil, France

Specimen identification at the species level is a critical challenge for understanding community structure and conserving biodiversity. The use of mitochondrial DNA barcodes, in addition to morphology, has proven to be a useful tool for earthworm identification, but it has also raised difficulties. Thus, approaches to delineating molecular operational taxonomic units (MOTUs) from cytochrome c oxidase subunit 1 (COI) or 16S rDNA sequence data often reveal more mitochondrial lineages than morphospecies, raising the question of whether these MOTUs should be used as taxonomic entities in community structure studies. Here, we used a newly acquired dataset of 576 COI barcodes of earthworms from Metropolitan France that were clustered in 36 MOTUs, corresponding to 21 morphospecies. We also incorporated data gathered from the literature to investigate this question. In order to match our MOTUs with already-described mitochondrial lineages, we downloaded reference sequences from the GenBank and BOLD platforms. In light of the difficulties encountered in recovering these sequences, we recommend that any new mitochondrial lineage described in a study be named consistently with previous works. Next, we analyzed the biological, ecological, and molecular data available in the literature on the different mitochondrial lineages that matched our MOTUs in order to determine if there was a consensus for species delimitation. Although the study specimens mainly belong to the Lumbricidae, which is one of the most studied families of earthworms, the data are often missing to determine if the MOTUs correspond to different species. Pending revision of the taxonomy, MOTUs for which mitochondrial divergence has been confirmed by morphological differences should be considered distinct taxonomic entities in community structure studies. In the absence of morphological differences and pending more data, we propose to distinguish for these analyses, within sexually reproducing morphospecies, the MOTUs for which mitochondrial divergence has been confirmed by multilocus nuclear data, while in the case of reproduction by parthenogenesis, ecological differences between MOTUs are necessary to consider them as different taxonomic entities.

KEYWORDS

community ecology, DNA barcoding, earthworm taxonomy, MOTUs, mitochondrial lineages, species assignment

Introduction

Accurate assessment of an area's biodiversity requires properly identifying the different specimens present in an assemblage in order to estimate species richness and relative abundance (Gotelli et al., 2004). For the study of earthworm communities, specimen identification at the species level is a critical challenge because of the lack of external diagnostic characters (Sturzenbaum et al., 2009), the near impossibility of identifying juveniles morphologically (Richard et al., 2010), and the several cases of cryptic variation recently reported (King et al., 2008; Martinsson and Erséus, 2021). In this context, the use of molecular barcodes based on a fragment of the cytochrome oxidase subunit 1 (COI) mitochondrial gene has been proposed for accurately identifying animal specimens (Hebert et al., 2003).

In most studies assessing earthworm biodiversity (i.e., estimating community diversity and describing their spatial distribution) with a DNA barcoding approach, methods proposing *de novo* species partitions from COI sequence datasets are used to delineate molecular operational taxonomic units (MOTUs), which are often considered species proxies (Decaens et al., 2016; Goulpeau et al., 2022; Dupont et al., 2023). In metabarcoding studies aiming at identifying earthworms, another mitochondrial gene, the 16S rDNA gene, is used instead (Bienert et al., 2012; Roy et al., 2022). In these studies, it is expected that the DNA barcoding results corroborate the taxonomic hypothesis established from morphological data (DeSalle et al., 2005) and thus allow species identification. However, the correspondence between MOTUs and described nominal species is particularly tricky to make for earthworms for three main reasons. First, the DNA barcode library is largely incomplete in earthworms (in particular for species from the tropical regions) whereas the main principle of molecular identification is to match the barcode sequence of the unknown sample against a complete barcode reference library (Hebert et al., 2003). As a consequence, DNA barcoding is more often used for species delineation than species identification (Dupont et al., 2023), although this is not the original purpose of this approach (Hebert et al., 2003). Second, several MOTUs may be obtained within a nominal species with no known morphological characters allowing to distinguish them, suggesting that these MOTUs may correspond to cryptic species (Fišer et al., 2018; Marchán et al., 2018). Several complexes of cryptic earthworm species have indeed been revealed over the past 15 years, thanks to molecular studies (King et al., 2008; Martinsson and Erséus, 2017; Martinsson and Erséus, 2021). Third, because different MOTUs obtained from COI sequence data may not be the results of a speciation event, but the result of retention of ancestral polymorphism and incomplete lineage sorting, i.e., the genealogical histories of individual gene loci, may appear misleading or uninformative about the relationships among species or populations because of the retention and stochastic sorting of ancestral polymorphisms (Maddison et al., 2006). For instance, in the *Allolobophora chlorotica* aggregate of species, retention of ancestral polymorphism may explain the lack of reproductive isolation between the divergent mitochondrial (Mt) lineages L2 and L3 (Dupont et al., 2022).

In an integrated taxonomy approach, morphological data are routinely used to estimate the species delimitation hypotheses based on MOTU partition. However, they are not the only characters that can be used. Thus, nuclear molecular data, such as microsatellite genotypes, may be used to test hypotheses of reproductive isolation between MOTUs (Dupont et al., 2016; Dupont et al., 2022). If the genetic clusters delineated using nuclear data correspond to MOTUs obtained using mitochondrial data, it is generally considered that these MOTUs represent effectively different species, whereas if a mitonuclear incongruence is observed, the genetic differentiation between MOTUs is generally explained by the retention of the ancestral polymorphism (Dupont et al., 2016; Martinsson and Erséus, 2017). Moreover, ecological (Leaché et al., 2009) and behavioral (Tan et al., 2010) characters can be an important source of evidence in some cases. For instance, for the earthworm *Lumbricus rubellus*, in which two cryptic lineages (LA and LB) were described in sympatry in the UK (King et al., 2008), Jones et al. (2016) performed behavioral tests that revealed that individuals were significantly more likely to orientate toward the soil conditioned by worms of their own lineage, suggesting the release of lineage-specific chemical signals in the soil that may play a role in the reproductive isolation between these sympatric lineages. Considering that genome size and chromosome numbers are important cytological characters that significantly influence various organismal traits (Leong-Skornickova et al., 2007), data on the ploidy level and chromosome number may also be used as criteria for delineating species in earthworms. For instance, the various ploidy levels observed in *Octodrilus transpadanus* [from $2N$ to $7N$ with $N = 15$ (Garbar et al., 2009)], *Amyntas catenus* [from $2N$ to $4N$ with $N = 56$ (Shen et al., 2011)], or *Aporrectodea rosea* [from $2N$ to $10N$ with $N = 18$ (Briones et al., 2009; Vlasenko et al., 2011)] raise questions about their species status.

Obtaining cytological, ecological, and behavioral data is even more critical for species that do not reproduce sexually, for which testing for reproductive isolation is not a solution for delimiting lineages. Indeed, within species that reproduce asexually, every individual is reproductively isolated from every other and thus on its own evolutionary path (Lin et al., 2017). Other criteria than interbreeding are thus necessary to demonstrate that a lineage evolves separately from others and with its own evolutionary tendencies (De Queiroz, 1998). This is the case for the earthworm species that reproduce by parthenogenesis and which are often polyploids (Briones et al., 2009); they are frequent in the Lumbricidae family and have also been reported in the Megascolecidae (Cosin et al., 2011). Moreover, in asexual species, the variation in molecular markers may reflect the accumulation of neutral mutations over historical times rather than the existence of morphologically indiscernible, cryptic species (Škaloud and Rindi, 2013). For the delineation of parthenogenetic species, the concept of barcoding gap is thus irrelevant since it does not have a biological meaning; the clustering of sequences in different MOTUs may instead be the result of insufficient sampling within the taxa or extinction of intermediate clones (Lin et al., 2017). Some studies dealing with species delimitation in asexual taxa have defined species on an ecological basis, considering that if two or more

lineages have been sorted into different niches, then each lineage may be considered to be a distinct species (Lin et al., 2017).

The constant discoveries and updates in systematics, taxonomy, cytogenetics, ecology, and behavior of taxa can involve taxonomic changes that may affect measures of species diversity. Thus, recent results can both reveal unrecognized species and collapse previously delineated groups (Chiarucci et al., 2011). These changes depend on the species concepts adopted and the information and methods used to delimit species (Lessa et al., 2024). The main aim of our study was to evaluate the need for taxonomic changes in earthworms based on the recent splitting of formerly recognized earthworm species into divergent MOTUs using COI sequences and in light of other data gathered from the literature (genetics, ecology, behavior, and cytogenetics data). Here, we propose alternative delineations of taxonomic entities from a newly acquired dataset of 576 COI sequences from earthworms in Metropolitan France. We also provide a set of reference COI and 16S sequences facilitating the assignment of DNA barcodes to previously described Mt lineages and species, and we highlight future necessary directions in earthworm taxonomy and DNA barcoding research.

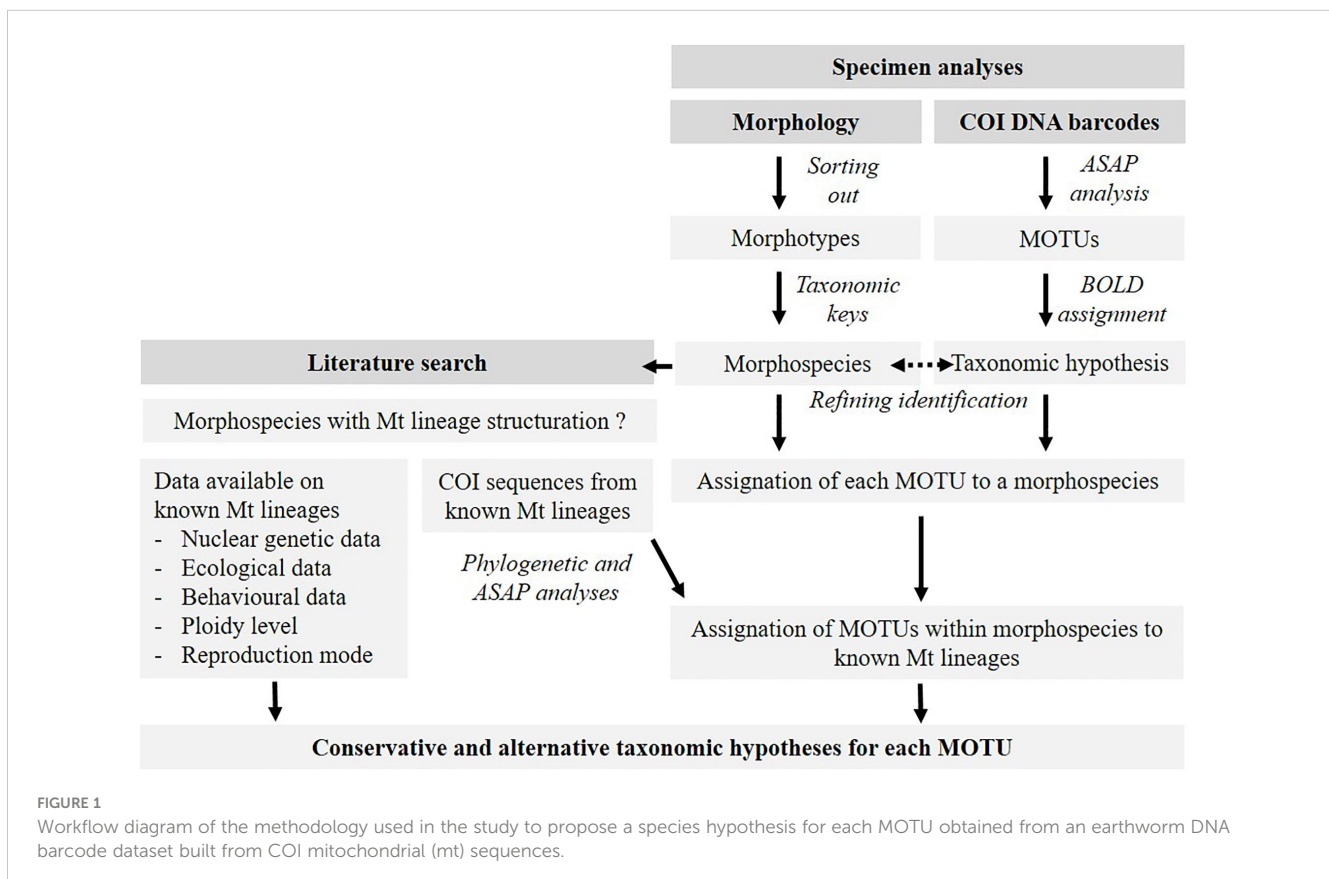
Material and methods

The general principle of the study steps is presented in a workflow diagram in Figure 1 and detailed below.

Specimen analyses

Earthworms were collected in Bretagne, Nouvelle-Aquitaine, and Ile-de-France regions of Metropolitan France (Figure 2) using a solution of Allyl isothiocyanate (Zaborski, 2003) and stored in absolute ethanol. Detailed information about the collection data is given in the public dataset dx.doi.org/10.5883/DS-PLW1 in the Barcode of Life Data Systems (BOLD, <https://www.boldsystems.org/>). After sampling, earthworms were grouped into morphotypes, and a subset of 576 individuals was then selected for this study to ensure it represented the diversity observed in this dataset. Adult earthworms were identified at the morphospecies level using the taxonomic keys of Bouché (1972) and Sims and Gerard (1999). Once the specimens had been grouped into genetically divergent groups (i.e., MOTUs), it was possible to refine the identification of the morphospecies and identify juveniles.

A small piece of cutaneous tissue was collected from each individual for DNA extraction using the NucleoSpin[®] Tissue or NucleoSpin[®] 96 Tissue kits (Macherey–Nagel). The COI gene was amplified using the primer pair described in Folmer et al. (1994). DNA bidirectional sequencing was carried out by the Eurofins Genomics company, and the sequences were manually aligned using the BioEdit program (Hall, 1999). All sequences are available in the BOLD dx.doi.org/10.5883/DS-PLW1 dataset, with GenBank accession numbers described in the Supplementary data (Supplementary Table S1). The identification engine of BOLD was used to check taxonomic assignments.



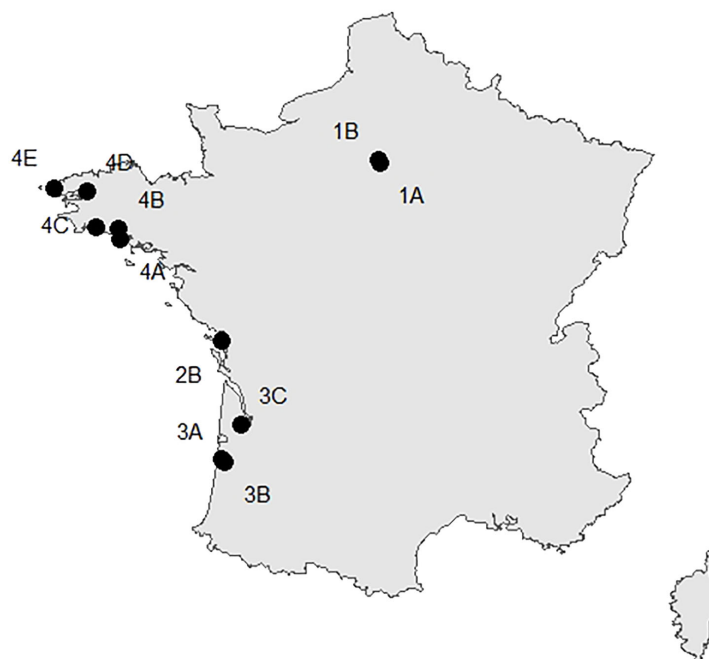


FIGURE 2

Location map of the sites where earthworms were collected in Ile-de-France (1A, B), Nouvelle-Aquitaine (2B, 3A–C), and Bretagne (4A–E).

From COI sequences, MOTU delimitation was done using the Assemble Species by Automatic Partitioning (ASAP) method based on implementing a hierarchical clustering algorithm (Puillandre et al., 2021), a method that has been proven particularly suitable for OTU delimitation with earthworm DNA barcodes (Goulpeau et al., 2022) and for the Lumbricidae family in particular (Marchán et al., 2023). Pairwise genetic distances (p -distances) were used to build a list of partitions ranked by a composite score computed using the probabilities of groups to be panmictic species and the barcode gap widths. A histogram showing the distribution of the intra- and interspecific p -distances and allowing the visualization of the barcode gap was produced by ASAP, and the MEGA 11 software (Tamura et al., 2021) was used to compute the mean p -distance within MOTUs.

A subset of 127 individuals was chosen (one to 14 individuals per MOTU obtained in the COI ASAP partition) in order to represent the genetic diversity in the dataset. For these specimens, the mitochondrial 16S rDNA gene (about 501 pb) was amplified using 16SAr and 16SBr (Palumbi, 1996) or ewA and ewF (Bienert et al., 2012) primers, and DNA sequencing was carried out by the Eurofins Genomics company, followed by MOTU delimitation using ASAP. For 65 of these specimens (from 33 MOTUs with one to six individuals per MOTU), the nuclear 18S rDNA gene (1,134 pb) was amplified using 18SA and 18SBrev primers and sequenced using 18SA, 18SL, 18SO, and 18SBrev primers (Medlin et al., 1988). Nucleotide sequences were aligned using ClustalW, and a distance-based tree was built using the neighbor-joining method, p -distance, and 1,000 bootstrap replicates, as implemented in the MEGA 11 software. DNA sequences were deposited in GenBank (Supplementary Table S1) and in the BOLD dx.doi.org/10.5883/DS-PLW1 dataset.

Literature search and phylogenetic analysis of reference sequences

In order to serve as reference sequences for species and lineage nomenclature, COI sequences corresponding to previously described complexes of earthworm species were downloaded from the literature (Supplementary Table S2). Using these reference sequences, in addition to a selection of one to three sequences per MOTUs obtained in the first ASAP clustering analysis, a phylogenetic tree was reconstructed using the maximum-likelihood method. The best-fitting model (F81+I+G) was identified with the Bayesian Akaike Criterion (AIC) implemented in jModelTest2 (Guindon et al., 2003; Darriba et al., 2012). Maximum-likelihood (ML) analysis was carried out in the CIPRES Science Gateway web server (RAxML-HPC BlackBox), and 1,000 rapid bootstrap replicates were run. The sequence of an Enchytraeid species was used as an outgroup. Trees were visualized using the iTOL v5 online tool (Letunic and Bork, 2021). In order to check the correspondence of the study MOTUs with the Mt lineages described in the literature, an ASAP clustering analysis was performed on the dataset composed of reference sequences and study sequences.

Moreover, a review of the literature was performed for the study of morphospecies for which mitochondrial structuration was previously published. The different types of data that were recorded were as follows: (i) the known mode of reproduction, (ii) if variable ploidy level was recorded within the morphospecies, (iii) if nuclear genetic data confirmed the genetic differentiation between lineages, (iv) if morphological differences were showed between lineages, and (v) if differences of ecological characteristics were recorded between lineages within a morphospecies. These data were used to propose conservative and alternative hypotheses of species status for each MOTU.

Results

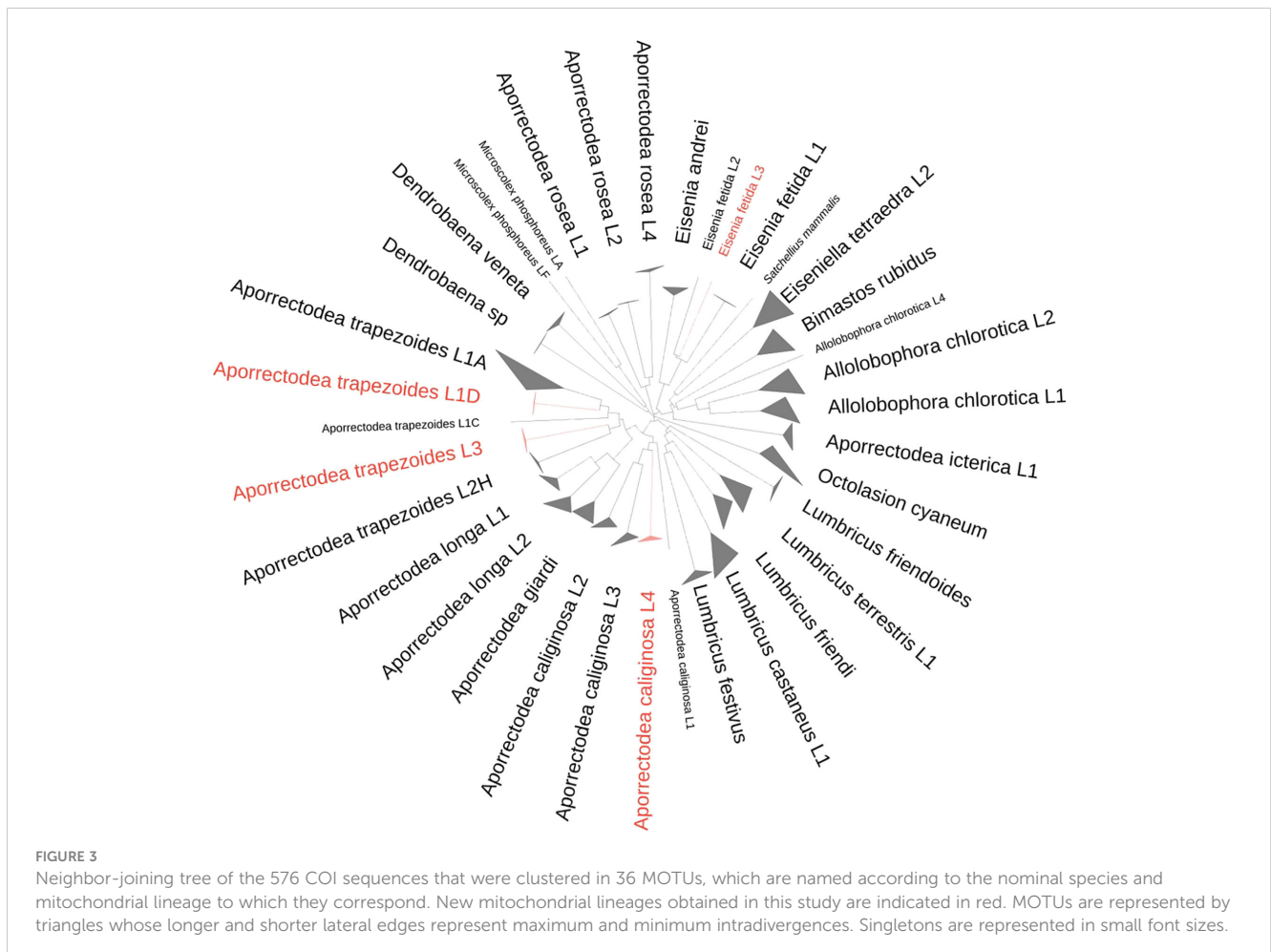
MOTU delimitation

Among the 21 morphotypes that were sorted, 20 nominal species could be morphologically identified. The highest ASAP score of the clustering analysis of the 576 COI sequences of 658 bp (284 informative sites) was obtained for a clustering of the sequences into 36 MOTUs (ASAP score of 6, threshold distance of 6.2%, [Supplementary Figure S1](#); [Figure 3](#)), which matched well with MOTUs and Mt lineages found in the literature. The average p -distance between MOTUs ranged from 7.8% between *Aporrectodea longa* L1 and L2 to 24.9% between *A. longa* L1 and *Eiseniella tetraedra* L2 ([Supplementary Table S3](#)). In several cases, more than one MOTU corresponded to a morphologically identified nominal species: four MOTUs for *Aporrectodea caliginosa* (average pairwise p -distance between MOTUs ranging from 10.6% to 15.1%), two MOTUs for *A. longa* (average p -distance of 7.8%), five MOTUs for *Aporrectodea trapezoides* (average p -distance ranging from 10.9% to 16.3%), three MOTUs for *Aporrectodea rosea* (average p -distance ranging from 13.3% to 13.9%), three MOTUs for *Allolobophora chlorotica* (average p -distance ranging from 12.8% to 16.0%), three MOTUs for *Eisenia fetida* (average p -distance ranging from 11.2% to 12.0%),

and two MOTUs for *Microscolex phosphoreus* (average p -distance of 15.7%). This MOTU delimitation was well recovered by the analysis of the 127 sequences of 16S (187 informative sites) using ASAP analysis. The only COI lineages that were not distinguished using 16S MOTU delimitation, whatever the ASAP score, were two *A. trapezoides* lineages (L1D and L2H) and the two *A. longa* lineages (L1 and L2) ([Supplementary Figure S2](#)). The 65 sequences of 18S were not variable enough despite a 1,134-bp length (24 informative sites), to distinguish most of the morphospecies ([Supplementary Figure S3](#)).

Assignment of specimens to known mitochondrial lineages

The ML phylogenetic tree was built using reference sequences from the literature ([Supplementary Figure S4](#)), and the ASAP analysis of clustering (ASAP score of 4.5) made it possible to assign the different MOTUs to known Mt lineages already described in the literature. When a study MOTU did not cluster with any of the reference sequences, a new lineage name was created. Thus, in this study, we recorded *Aporrectodea caliginosa* L4, *Aporrectodea trapezoides* L3, *Aporrectodea trapezoides* L1D, and *Eisenia fetida* L3, for which no corresponding published lineages



were found in the literature (in red in [Figure 3](#); [Supplementary Figure S2](#)). Six specimens shared similar morphological characteristics but could not be assigned morphologically to any nominal species. They shared the same COI sequence, which was assigned with 100% similarities to the sequences of *Dendrobaena* cf. *attemsi* and *Dendrobaena* sp. *CEB* uploaded in the BOLD research ID engine. These specimens were thus labeled as *Dendrobaena* sp. in this study, and further work on morphological identification is needed to identify them.

Assignment of specimens to species hypotheses

Of the 36 MOTUs obtained, nine corresponded to a nominal species without any Mt lineage structuration (*Bimastos rubidus*, *Dendrobaena veneta*, *Dendrobaena* sp., *Eisenia andrei*, *Satchellius mammalis*, *Octolasion cyaneum*, *Lumbricus friendi*, *Lumbricus festivus*, and *Aporrectodea giardi*), five corresponded to a nominal species with known lineage structuration in the literature, but only one lineage was recorded in the dataset (*Aporrectodea icterica*, *Eiseniella tetraedra*, *Lumbricus castaneus*, *Lumbricus friendoides*, and *Lumbricus terrestris*), and seven corresponded to a nominal species with known lineage structuration and with several lineages in the dataset (*Aporrectodea caliginosa*, *Aporrectodea longa*, *Aporrectodea trapezoides*, *Aporrectodea rosea*, *Allolobophora chlorotica*, *Eisenia fetida*, and *Microscolex phosphoreus*). On the 12 morphospecies for which Mt lineages were described in the literature, most are bisexual and diploid, with a few exceptions ([Table 1](#)): *A. rosea* presents several reproduction modes and variable ploidy levels; *A. trapezoides* (3N and 4N) and *E. tetraedra* (4N) are parthenogenetic, and *E. fetida* (2N) is capable of selfing. *M. phosphoreus* is bisexual, but the ploidy data were not available ([Table 1](#)). Nuclear molecular markers of different categories (microsatellites, AFLPs, Radseq, and sequences of H3, 28S, and ITS2 genes) were used to investigate the taxonomy of seven of these morphospecies. For three of them (*A. icterica*, *A. longa*, and *L. terrestris*), the divergence between Mt lineages was not confirmed by nuclear data. For the other four, nuclear data confirmed the divergence among some lineages (between L1, L4, and L2/L3 of *A. chlorotica*, between L1 and L2 of *A. trapezoides*, and between LA and LB of *L. rubellus*) but not among all ([Table 1](#)). For only four morphospecies, data other than genetic could be associated with Mt lineages: ecological and behavioral data allow us to distinguish two morphs (L2/L3 versus L1 and L4 Mt lineages) in the *A. chlorotica* morphospecies, morphological and behavioral data allow to distinguish the LA and LB lineages of *L. rubellus*, and some morphological differences could be found *a posteriori* between lineages of the *A. caliginosa* and *M. phosphoreus* morphospecies ([Table 1](#)).

In light of these data, two proposals for grouping specimens into putative species emerged. In the conservative hypothesis, different Mt lineages are considered to be different taxonomic entities only if reproductive isolation is demonstrated. This was rarely the case and was an inapplicable criterion for parthenogenetic species. An alternative hypothesis proposes to consider that certain lineages

could constitute separate taxonomic entities in studies of community structure if the mitochondrial divergence was confirmed by morphological data and/or at least nuclear genetic data in morphospecies reproducing sexually or ecological data in morphospecies reproducing by parthenogenesis ([Table 1](#)).

Discussion

The molecular barcode-based identification of earthworm specimens collected in ecological studies certainly presents advantages but can be difficult to implement. Although obtaining molecular data is relatively easy, accurately grouping specimens within species may be trickier. According to the Unified Species concept ([De Queiroz, 2007](#)), species are ancestor-descendant population-level lineages that are on a unique evolutionary trajectory and evolve independently from other such lineages. This conception of species is distinguished from the practical delimitation of species. Thus, species delimitation consists of raising hypotheses about the existence and independent evolution of population-level lineages, and delimiting species requires the use of a wide range of different methods and criteria that target different aspects of evolutionary independence ([Conix, 2022](#)). Here, we propose a workflow ([Figure 1](#)) to pose species hypotheses using multiple lines of evidence for species delimitation in earthworms, and we highlight the complexity of the evolutionary processes that shape this group of organisms.

Considering the existence of cryptic species and different Mt lineages within species of uncertain status, it appears essential, for accurate earthworm identification, to sequence the specimens studied in addition to the morphological description. However, identification by mitochondrial molecular barcodes is not without caveats. One of the first difficulties encountered comes from the availability of sequences used in previous papers and the use of the GenBank and BOLD libraries of DNA barcodes. We noticed that some sequences are not available in the public databases, or the information given in the papers is not sufficient to retrieve easily the sequences. For instance, some authors only give a range of accession numbers, which causes an important loss of time for users trying to retrieve sequences for each specimen individually. It is essential to allow easy consultation and use of the sequences mentioned in the articles. In addition to being incomplete for earthworms, these libraries show mistakes in the attribution of a species name to a sequence for a variety of taxa ([Meiklejohn et al., 2019](#)).

Most of the time, the Mt lineage is not specified in the specimen data of the sequence deposited in GenBank or BOLD, although it is done in some studies. For instance, in our study, lineages for *Eiseniella tetraedra* could not be annotated based on a recent study that identified eight lineages within two clades ([De Sosa et al., 2023](#)) using COI, 16S, 28S, 12S, and ND1 genes. Unfortunately, the different GenBank accession numbers for these sequences could not be linked to the described lineages, making it impossible for us to use them in our reference sequences. The lineage of our *E. tetraedra* specimens was thus annotated according to [Porco et al. \(2018\)](#). [Porco et al. \(2018\)](#) already highlighted the necessity to establish a consistent annotation of the described

TABLE 1 Characteristics of the study morphospecies, for which several mitochondrial (Mt) lineages were described in the literature.

Name	Repro.	Ploidy	Mt lineages	Differences between Mt lineages		Conservative hypothesis	Alternative hypothesis	Study MOTU
				Nuclear data	Other data			
<i>Allolobophora chlorotica</i> Savigny, 1826	Bisexual (Briones et al., 2009)	2N (N = 16) (Briones et al., 2009)	L1, L2, L3, L4, L5 (King et al., 2008), L6, and L7 (Dupont et al., 2011)	No diff. between L2 and L3; diff. between L1, L4, and L2/L3 with hybrids (H3, ITS2, μ sats) (Dupont et al., 2011; Dupont et al., 2016; Dupont et al., 2022)	L2 and L3 = green morph common in wet soil (Lowe and Butt, 2007; Dupont et al., 2011), epigeic behavior (Bouché, 1972) L1 and L4 = pink morph common in dry soils (Lowe and Butt, 2007; Dupont et al., 2011) Repro. isolation between morphs (Lowe and Butt, 2008) (Dupont et al., 2022)	L2/L3 is a diff. sp	L2/L3, L1, and L4 are diff. sp	L1, L2, and L4
<i>Aporrectodea caliginosa</i> Savigny, 1826	Bisexual (Briones et al., 2009)	2N (N = 18) (Briones et al., 2009)	L1, L2, and L3 (Perez-Losada et al., 2009; Porco et al., 2018)		Morpho. diff. between L1, L2, and L3 (Shekhovtsov et al., 2021)	One sp	L1, L2, and L3 are diff. sp	L1, L2, L3, and L4 ^a
<i>Aporrectodea icterica</i> Savigny, 1826	Bisexual (Torres-Leguizamon et al., 2012)	2N (N = 16) (Briones et al., 2009)	L1 and L2 (Torres-Leguizamon et al., 2012; Torres-Leguizamon et al., 2014; Porco et al., 2018)	No diff. (μ sats) (Torres-Leguizamon et al., 2012; Torres-Leguizamon et al., 2014)		One sp	One sp	L1
<i>Aporrectodea longa</i> Ude, 1885	Bisexual (Briones et al., 2009)	2N (N = 18) (Briones et al., 2009)	L1 and L2 (Martinsson et al., 2015)	No diff. (ITS2) (Martinsson et al., 2015)		One sp	One sp	L1 and L2
<i>Aporrectodea trapezoides</i> Dugés, 1828	Partheno.	3N, 4N (N = 18) (Briones et al., 2009)	L1 (A to C) and L2 (D to H) (Fernandez et al., 2011; Fernández et al., 2012)	Diff. between L1 and L2 (28S, H3) (Fernández et al., 2012)	Diff. in ploidy could not be linked to Mt lineages (Perez-Losada et al., 2009)	One sp	Insufficient data	L1A, L1C, L1D ^a , L2H, and L3 ^a
<i>Aporrectodea rosea</i> Savigny, 1826	Bisexual, partheno (Briones et al., 2009), Mixte (Terhivuo and Saura, 2006)	2N, 3N, 4N, 5N, 6N, 8N, and 10N (N = 18) (Briones et al., 2009)	Eurosiberian (L1, L2, L3, L4, L5, and L6) and Mediterranean lineages (Fernandez et al., 2016; Shekhovtsov et al., 2020)	28S and H3 data exist (Fernandez et al., 2016; Shekhovtsov et al., 2020) but congruence with COI lineages not tested	Diff. in ploidy (Vlasenko et al., 2011), repro. (Briones et al., 2009), and morpho. (Terhivuo and Saura 2006) could not be linked to Mt lineages	One sp	Insufficient data	L1, L2, and L4
<i>Eisenia fetida</i> Savigny, 1826	Bisexual selfing (Briones et al., 2009)	2N (N = 11) (Briones et al., 2009)	L1 and L2 (Perez-Losada et al., 2005; Römbke et al., 2016)			One sp	Insufficient data	L1, L2, and L3 ^a
<i>Eiseniella tetraedra</i> Savigny, 1826	Partheno (De Sosa et al., 2023)	4N (De Sosa et al., 2023)	L1 and L2 (Porco et al., 2018); 8 lineages (De Sosa et al., 2023)	No diff. (28S) (De Sosa et al., 2023)		One sp	Insufficient data	L2
<i>Lumbricus castaneus</i> Savigny, 1826	Bisexual (Briones et al., 2009; Dupont et al., 2019)	2N (N = 18) (Briones et al., 2009)	L1, L2, and L3 (Porco et al., 2018)			One sp	Insufficient data	L1

(Continued)

TABLE 1 Continued

Name	Repro.	Ploidy	Mt lineages		Differences between Mt lineages		Conservative hypothesis	Alternative hypothesis	Study MOTU
			Mt lineages		Nuclear data	Other data			
<i>Lumbricus rubellus</i> Hoffmeister, 1843	Bisexual (Briones et al., 2009)	2N (N = 18) (Briones et al., 2009)	L1, L2, and L3 (Porco et al., 2018); LA, LB, LC, LD, LE, LG, LH, LJ, LK, and LM (Donnelly et al., 2014; Giska et al., 2015; Martinsson and Erséus, 2017); L2 = LA = L. <i>friendoides</i> (Briones et al., 2022)	Diff. between LA and LB (usats) (Donnelly et al., 2013), AFLPs (Kille et al., 2013), and RADseq (Anderson et al., 2017). No diff. between LA, LC, LD, and LE (Giska et al., 2015)	Diff. between LA and LB in morpho (Donnelly et al., 2014) and behav (Jones et al., 2016)	<i>L. friendoides</i> a diff. sp Others = one sp	<i>L. friendoides</i> and LB are diff. sp Insufficient data for other lineages	<i>L. friendoides</i>	
<i>Lumbricus terrestris</i> Linnaeus, 1758	Bisexual (Briones et al., 2009)	2N (N = 18) (Briones et al., 2009)	L1 and L2 (Martinsson and Erséus, 2017)	No diff. (H3) (Martinsson and Erséus, 2017)		One sp	Insufficient data	L1	
<i>Microsclex phosphoreus</i> (Dugés, 1837)	Bisexual (Bouché, 1972)	No data	LA, LB, LC, LE, and LF (Rota et al., 2018)		Morpho. diff. between LA, LC, LD, and LF (Rota et al., 2018)	One sp	LA, LC, LD, and LF are diff. sp	LA and LF	

The reproductive mode (bisexual, parthenogenesis, selfing), the ploidy level with the haploid number of chromosomes (N), the identified Mt lineages, the nuclear molecular markers used to study the lineages (H3, 28S, or ITS2 genes; just for microsatellites, AFLPs, Radseq), and whether concordance with Mt data was obtained, as well as the known morphological (morpho), ecological, behavioral (behav.), or molecular differences (diff.) between Mt lineages are indicated when known. The conservative hypothesis of species (sp) delimitation as well as the alternative hypothesis that could be considered in light of the current literature data but which may evolve with further data and the found MOTUs in the present study are also indicated. *Lineages newly found (present study).

earthworm lineages. In a context of uncertain and changing taxonomy, this type of annotation facilitates the correspondence between the results obtained about these lineages in different studies as well as the updating of the data (i.e., attribution of the correct results to the correct lineage or species in the case of a taxonomic revision). Here, our analyses made it possible to match the lineage “nomenclatures” of different studies. For instance, the LA and L2 Mt lineages previously found within *L. rubellus* are the same (Table 1; Supplementary Figure S4) and have now been described as *Lumbricus friendoides* species (Briones et al., 2022). According to the work of Giska et al. (2015), we can therefore assume that the LC, LD, and LE lineages that are not reproductively isolated from LA also belong to the species *L. friendoides*. In light of the difficulties encountered in our study, we recommend building a reliable, practical, and accessible database for earthworms capable of evolving with taxonomic research and whose sequences are associated with both a species and lineage name.

A literature search of biological, ecological, and molecular data available on known Mt lineages is necessary to propose hypotheses of species delimitation for biodiversity assessment and community structure analysis. For instance, we searched in the literature if nuclear data made it possible to confirm differentiation observed with mitochondrial data. Nuclear data can mainly be of two types: sequence data and genotype data, which will not provide the same degree of information. Genotype data, obtained from multiple codominant (i.e., the heterozygotes can be distinguished from homozygotes) molecular markers such as microsatellites or SNPs, allow for testing for reproductive isolation between Mt lineages using a population genetic approach (Twyford and Ennos, 2012). If the MOTUS obtained using mitochondrial data are not confirmed by results obtained from genotyping, it is thus considered that these MOTUs do not correspond to different species. For instance, Torres-Leguizamon et al. (2012) showed that microsatellite data did not allow to distinguish the two Mt lineages observed in *Aporrectodea icterica*. It was thus proposed that the structuration in two COI lineages in this species is the reflection of a past history of divergence in allopatry, which was not long enough to lead to reproductive isolation, followed by secondary contact with interbreeding (Torres-Leguizamon et al., 2014). A similar result was obtained for the L2 and L3 lineages in *Allolobophora chlorotica*, for which microsatellite data and parentage analysis revealed a lack of reproductive isolation despite a genetic divergence at the mitochondrial level (Dupont et al., 2016; Dupont et al., 2022).

In other cases, genetic differentiation between Mt lineages was confirmed using microsatellite data, and this differentiation was often also confirmed by another kind of data (e.g., morphological, ecological, behavioral), allowing us to consider that these MOTUs are different species. This is, for instance, the case of the L1 and L2/L3 lineages in *A. chlorotica*. Microsatellite data and parentage analysis revealed reproductive isolation between these lineages (Dupont et al., 2016; Dupont et al., 2022). This differentiation was confirmed by pigmentation data since L1 corresponds to a pink morph and L2/L3 to a green morph; by ecological data since the pink morph is more frequent in dry soil and the green morph in wet soil (Lowe and Butt, 2007); and by behavioral data with the green morph, which has a more epigeic behavior (Bouché, 1972).

Similarly, in *L. rubellus*, many data have been accumulated for the LA and LB lineages. It was shown that they have nuclear (microsatellites, AFLPs, RADseq [Donnelly et al., 2013; Kille et al., 2013; Anderson et al., 2017]), morphological (Donnelly et al., 2014), and behavioral differences (Jones et al., 2016). This led to the elevation of LA to species rank, now named *L. friendoides* (Briones et al., 2022). This taxonomic revision then questions the name to assign to other existing lineages, particularly LC, LD, and LE, which seem not reproductively isolated from LA (Giska et al., 2015) but whose isolation or nonisolation with LB has not been demonstrated.

Sometimes, mitochondrial data are completed by sequences of nuclear markers such as H3, ITS2, 18S, and 28S in order to use a phylogenetic approach for investigating reproductive isolation. However, studies using this kind of nuclear data are largely constrained by the phylogenetic resolution of each locus and the low number of loci that can be successfully sequenced for several different species. Comparisons of poorly resolved gene trees using markers with limited sequence divergence between species are likely to be uninformative in tracing the reticulate history of species (Linder and Rieseberg, 2004). In the case where a nuclear gene does not confirm the differentiation between the MOTUs, as in our study with the 18S rRNA gene, no conclusion can be drawn because this absence of differentiation can be explained by a weak polymorphism of the nuclear gene, which leads to similar sequences even between different species. Thus, this type of result should not put an end to investigations about species delimitation. This is, for instance, the case for *Lumbricus terrestris*, a species that has long been confused with *Lumbricus herculeus*, but the two species could finally be distinguished on the basis of morphological, mitochondrial (James et al., 2010), and nuclear divergence (Martinsson and Erséus, 2017). Within *L. terrestris*, two COI lineages are known, and one of them, *L. terrestris* L2 is closer to *Lumbricus festivus* than to *L. terrestris* L1 according to the COI gene but not the H3 gene (Richard et al., 2010; Martinsson and Erséus, 2017). These two lineages are therefore still considered a single species, but further data should be acquired to confirm this. This situation was also found within the *Aporrectodea longa* species, in which a mitonuclear discordance between COI and ITS2 genes has been shown (Martinsson et al., 2015).

In some cases, nuclear data are lacking, but other kinds of data allow us to propose a species hypothesis. For instance, *Microscolex phosphoreus* has six COI lineages, annotated from A to F (Rota et al., 2018). The A and F lineages are present in this study with more than 15% *p*-distance for COI. Morphological differences were shown between LA, LC, LD, and LF (Rota et al., 2018), thus suggesting that they may be different species. Rota et al. (2018) indeed found that chaetal intervals and spermathecal structure were good discriminating characters to distinguish these Mt lineages. A species name should thus be attributed to them. The case is different for *A. caliginosa*, for which three lineages (L1 to L3) were previously described (Porco et al., 2018), and we add an L4 lineage in our study. Morphological differences exist between L1, L2, and L3, but no character may be used to distinguish the lineages with sufficient accuracy (Shekhovtsov et al., 2021). Shekhovtsov et al. (2021)

indeed showed that L1 was significantly smaller than L2 and L3, with fewer segments and smaller clitellum, and that L2 was the biggest of all three and also tended toward the earlier start of the clitellum, by about three-fourths of a segment compared to L3. In the absence of clear discriminating characters, we propose to consider these different lineages as different entities in the less conservative species hypothesis, although the species delimitation should be confirmed by other kind of data (e.g., parentage analysis and ecological data).

Lastly, the existence of parthenogenetic species with mixed reproduction (i.e., amphimixis and parthenogenesis) represents a difficult issue for earthworm taxonomy. For these, the interbreeding criterion cannot be used. Although other criteria may be used instead, ambiguous results are often obtained in these species, which makes decision-making difficult. One of the most emblematic morphospecies of those parthenogenetic Lumbricidae species is *Aporrectodea trapezoides*, which was originally part of *A. caliginosa* complex (Perez-Losada et al., 2009). This morphospecies is composed of two major lineages, and within each lineage, several clones were described: A to C in L1 and D to H in L2 (Fernandez et al., 2011). Using COI, COII, 28S, and H3 genes, Fernández et al. (2012) showed that L1 is monophyletic while L2 is paraphyletic, highlighting the need for a taxonomic revision for *A. trapezoides*. Similar difficulties were encountered for the *Aporrectodea rosea* morphospecies, for which up to seven COI lineages were described (Fernandez et al., 2016; Shekhovtsov et al., 2020). In addition, differences in ploidy (Vlasenko et al., 2011), reproductive mode (Briones et al., 2009), and morphology (Terhivuo and Saura, 2006) have been described in this morphospecies. Unfortunately, these data are impossible to match with the described Mt lineages, forcing us to retain the single-species conservative hypothesis pending further work.

Conclusion

For biodiversity assessment or studies of earthworm communities based on COI DNA barcoding data, three solutions for delimiting taxonomic entities are possible. According to the most conservative hypothesis, different Mt lineages are considered different taxonomic entities only if reproductive isolation has been demonstrated. Due to limited data for most investigated Mt lineages in this study, and with some species reproducing via parthenogenesis, the most conservative species hypothesis necessitates considering each earthworm morphospecies as a unique species. Exceptions include the cases of *Allolobophora chlorotica* and *Lumbricus friendoides*. Another solution is to use all Mt lineages as taxonomic entities, but this is not satisfactory since divergent Mt lineages may show a lack of reproductive isolation, which has not yet been tested or verified by scientists. At last, we propose an alternative solution, which is to consider distinguishing Mt lineages as different entities, pending a formal description of the species, when the COI divergence is congruent with morphological characters and/or at least multilocus nuclear data for sexually reproducing species or ecological data for

parthenogenetic species. In order to highlight possible ecological differences between Mt lineages, alternative species delimitations could be tested in studies of the structure of earthworm communities. However, for this to be the case, future studies on these morphospecies must make the effort to barcode the samples and assign them to already-known lineages. This approach to lineage annotation is a way to ensure the correspondence and continuity of the data obtained. It only makes sense if research continues to focus on the differences between these Mt lineages in order to confirm the need to revise the existing taxonomy. In particular, more phylogenetic, phylogenomic, and population genetic studies focusing on the nuclear genome are needed to confirm sexual species delimitation that is still ambiguous, while studies aiming at highlighting ecological differences are needed for parthenogenetic species. Once the Mt lineage is considered a distinct species, a species name should then be assigned.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: Barcode of Life Data Systems (BOLD, <https://www.boldsystems.org/>), [dx.doi.org/10.5883/DS-PLW1](https://doi.org/10.5883/DS-PLW1). NCBI GenBank accession numbers are indicated in the [Supplementary Material](#).

Ethics statement

Ethical approval was not required for the study involving animals in accordance with the local legislation and institutional requirements because research on invertebrates (earthworms).

Author contributions

AB: Investigation, Formal analysis, Writing – original draft. SN: Investigation, Writing – review & editing. VR: Supervision, Writing – review & editing. LD: Funding acquisition, Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

References

- Anderson, C., Cunha, L., Sechi, P., Kille, P., and Spurgeon, D. (2017). Genetic variation in populations of the earthworm, *Lumbricus rubellus*, across contaminated mine sites. *BMC Genet.* 18, 97. doi: 10.1186/s12863-017-0557-8
- Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J. J., et al. (2012). Tracking earthworm communities from soil DNA. *Mol. Ecol.* 21, 2017–30. doi: 10.1111/j.1365-294X.2011.05407.x
- Bouché, M. B. (1972). “Lombriciens de France,” in *Ecologie et systématique* (INRA, Paris).
- Briones, M. J. I., García-Souto, D., Galindo, J., Morán, P., Keith, A., and Schmidt, O. (2022). Molecular data confirms the existence of distinct lineages within *Lumbricus friendi* (Cognetti 1904) and related “friends.” *Eur. J. Soil Biol.* 108, 103382. doi: 10.1016/j.ejsobi.2021.103382
- Briones, M. J. I., Moran, P., and Posada, D. (2009). Are the sexual, somatic and genetic characters enough to solve nomenclatural problems in lumbricid taxonomy? *Soil Biol. Biochem.* 41, 2257–71. doi: 10.1016/j.soilbio.2009.07.008
- Chiarucci, A., Bacaro, G., and Scheiner, S. M. (2011). Old and new challenges in using species diversity for assessing biodiversity. *Phil. Trans. R. Soc. B* 366, 2426–37. doi: 10.1098/rstb.2011.0065
- Conix, S. (2022). Consensus and a unified species paradigm: reality or idle hope. *Philosophy Theory Pract. Biol.* 14, 8. doi: 10.3998/ptpbio.2102
- Cosin, D. J. D., Novo, M., and Fernandez, R. (2011). *Reproduction of earthworms: sexual selection and parthenogenesis* (Dans: Biology of earthworms. Berlin Heidelberg: Springer-Verlag), 69–86.
- Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772–772. doi: 10.1038/nmeth.2109
- Decaens, T., Porco, D., James, S. W., Brown, G. G., Chassany, V., Dubs, F., et al. (2016). DNA barcoding reveals diversity patterns of earthworm communities in remote tropical forests of French Guiana. *Soil Biol. Biochem.* 92, 171–83. doi: 10.1016/j.soilbio.2015.10.009

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the French National Research Agency (ANR) as a part of the project PLATWORM No. ANR-21-C02-0016.

Acknowledgments

The authors acknowledge the contributions of Agnès Gigon, Thomas Lerch, and Yoan Fourcade for earthworm sampling, as well as Flavie Lamendin and Kaur Gaganpreet for laboratory work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2024.1358984/full#supplementary-material>

- De Queiroz, K. (1998). *The general lineage concept of species, species criteria, and the process of speciation* (Dans: Endless forms: species and speciation. Oxford University Press), 57–75.
- De Queiroz, K. (2007). Species concepts and species delimitation. *Syst. Biol.* 56, 879–86. doi: 10.1080/1063150701701083
- DeSalle, R., Egan, M. G., and Siddall, M. (2005). The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Phil. Trans. R. Soc. B* 360, 1905–16. doi: 10.1098/rstb.2005.1722
- De Sosa, I., Marchán, D. F., Novo, M., Almodóvar, A., and Díaz Cosin, D. J. (2023). Phylogeography of a riparian earthworm shows environmental factors influence genetic structure. *J. Biogeography* 50, 156–68. doi: 10.1111/jbi.14518
- Donnelly, R. K., Harper, G. L., Morgan, A. J., Orozco-Terwengel, P., Pinto-Juma, G. A., and Bruford, M. W. (2013). Nuclear DNA recapitulates the cryptic mitochondrial lineages of *Lumbricus rubellus* and suggests the existence of cryptic species in an ecotoxicological soil sentinel. *Biol. J. Linn. Soc. Lond* 110, 780–95. doi: 10.1111/bij.2013.110.issue-4
- Donnelly, R. K., Harper, G. L., Morgan, A. J., Pinto-Juma, G. A., and Bruford, M. W. (2014). Mitochondrial DNA and morphological variation in the sentinel earthworm species *Lumbricus rubellus*. *Eur. J. Soil Biol.* 64, 23–9. doi: 10.1016/j.ejsobi.2014.07.002
- Dupont, L., Audusseau, H., Porco, D., and Butt, K. R. (2022). Mitonuclear discordance and patterns of reproductive isolation in a complex of simultaneously hermaphroditic species, the *Allolobophora chlorotica* case study. *J. Evolutionary Biol.* 35, 831–43. doi: 10.1111/jeb.14017
- Dupont, L., Brunet, C.-M., Fourcade, Y., James, S., Gabriac, Q., and Coulis, M. (2023). Recording earthworm diversity on the tropical island of Martinique using DNA barcoding unveiled endemic species in bromeliad plants. *Soil Biol. Biochem.* 182, 109038. doi: 10.1016/j.soilbio.2023.109038
- Dupont, L., Lazrek, F., Porco, D., King, R. A., Rougerie, R., Symondson, W. O. C., et al. (2011). New insight into the genetic structure of the *Allolobophora chlorotica* aggregate in Europe using microsatellite and mitochondrial data. *Pedobiologia* 54, 217–24. doi: 10.1016/j.pedobi.2011.03.004
- Dupont, L., Pauwels, M., Dume, C., Deschins, V., Audusseau, H., Gigon, A., et al. (2019). Genetic variation of the epigeic earthworm *Lumbricus castaneus* populations in urban soils of the Paris region (France) revealed using eight newly developed microsatellite markers. *Appl. Soil Ecol.* 135, 33–7. doi: 10.1016/j.apsoil.2018.11.004
- Dupont, L., Porco, D., Symondson, W. O. C., and Roy, V. (2016). Hybridization relics complicate barcode-based identification of species in earthworms. *Mol. Ecol. Resour* 16, 883–94. doi: 10.1111/1755-0998.12517
- Fernandez, R., Almodovar, A., Novo, M., Gutierrez, M., and Cosin, D. J. D. (2011). A vagrant clone in a peregrine species: Phylogeography, high clonal diversity and geographical distribution in the earthworm *Aporrectodea trapezoides* (Duges, 1828). *Soil Biol. Biochem.* 43, 2085–93. doi: 10.1016/j.soilbio.2011.06.007
- Fernández, R., Almodóvar, A., Novo, M., Simancas, B., and Díaz Cosin, D. J. (2012). Adding complexity to the complex: New insights into the phylogeny, diversification and origin of parthenogenesis in the Aporrectodea caliginosa species complex (Oligochaeta, Lumbricidae). *Mol. Phylogenet. Evol.* 64, 368–79. doi: 10.1016/j.ympev.2012.04.011
- Fernandez, R., Novo, M., Marchan, D. F., and Cosin, D. J. D. (2016). Diversification patterns in cosmopolitan earthworms: similar mode but different tempo. *Mol. Phylogenet. Evol.* 94, 701–8. doi: 10.1016/j.ympev.2015.07.017
- Fišer, C., Robinson, C. T., and Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Mol. Ecol.* 27, 613–35. doi: 10.1111/mec.14486
- Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294–9.
- Garbar, A., Onyschuk, I., and Mezhzherin, S. (2009). Polyploid races, genetic structure and morphological features of the earthworm *Octodrilus transpadanus* (Rosa, 1884) (Oligochaeta: Lumbricidae) in the Ukraine. *CCG* 3, 131–41. doi: 10.3897/compctogen.v3i2.20
- Giska, I., Sechi, P., and Babik, W. (2015). Deeply divergent sympatric mitochondrial lineages of the earthworm *Lumbricus rubellus* are not reproductively isolated. *BMC Evol. Biol.* 15, 217. doi: 10.1186/s12862-015-0488-9
- Gotelli, N. J., Godfray, H. C. J., and Knapp, S. (2004). A taxonomic wish-list for community ecology. *Phil. Trans. R. Soc. Lond B* 359, 585–97. doi: 10.1098/rstb.2003.1443
- Goulpeau, A., Penel, B., Maggia, M.-E., Marchán, D. F., Steinke, D., Hedde, M., et al. (2022). OTU delimitation with earthworm DNA barcodes: A comparison of methods. *Diversity*. 14, 866. doi: 10.3390/d14100866
- Guindon, S., Gascuel, O., and Rannala, B. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biol.* 52, 696–704. doi: 10.1080/10635150390235520
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp Ser.* 41, 95–8.
- Hebert, P. D. N., Cywinska, A., Ball, S. L., and Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. London Ser. B: Biol. Sci.* 270, 313–21. doi: 10.1098/rspb.2002.2218
- James, S. W., Porco, D., Decaens, T., Richard, B., Rougerie, R., and Erseus, C. (2010). DNA Barcoding reveals cryptic diversity in *Lumbricus terrestris* L. 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). *PLoS One* 5, e15629. doi: 10.1371/journal.pone.0015629
- Jones, G. L., Wills, A., Morgan, A. J., Thomas, R. J., Kille, P., and Novo, M. (2016). The worm has turned: Behavioural drivers of reproductive isolation between cryptic lineages. *Soil Biol. Biochem.* 98, 11–7. doi: 10.1016/j.soilbio.2016.03.015
- Kille, P., Andre, J., Anderson, C., Ang, H. N., Bruford, M. W., Bundy, J. G., et al. (2013). DNA sequence variation and methylation in an arsenic tolerant earthworm population. *Soil Biol. Biochem.* 57, 524–32. doi: 10.1016/j.soilbio.2012.10.014
- King, R. A., Tibble, A. L., and Symondson, O. C. (2008). Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. *Mol. Ecol.* 17, 4694–8. doi: 10.1111/j.1365-294X.2008.03931.x
- Leaché, A. D., Koo, M. S., Spencer, C. L., Papenfuss, T. J., Fisher, R. N., and McGuire, J. A. (2009). Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proc. Natl. Acad. Sci. U.S.A.* 106, 12418–23. doi: 10.1073/pnas.0906380106
- Leong-Skornickova, J., Sida, O., Jarolimova, V., Sabu, M., Fer, T., Travnicek, P., et al. (2007). Chromosome numbers and genome size variation in Indian species of curcuma (Zingiberaceae). *Ann. Botany.* 100, 505–26. doi: 10.1093/aob/mcm144
- Lessa, T., Stropp, J., Hortal, J., and Ladle, R. J. (2024). How taxonomic change influences forecasts of the Linnean shortfall (and what we can do about it)? *J. Biogeography* jbi.14829, 1–9. doi: 10.1111/jbi.14829
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–6. doi: 10.1093/nar/gkab301
- Lin, Y.-P., Edwards, R. D., Kondo, T., Semple, T. L., Cook, L. G., and Arthofer, W. (2017). Species delimitation in asexual insects of economic importance: The case of black scale (*Parasaissetia nigra*), a cosmopolitan parthenogenetic pest scale insect. *PLoS One* 12, e0175889. doi: 10.1371/journal.pone.0175889
- Linder, C. R., and Rieseberg, L. H. (2004). Reconstructing patterns of reticulate evolution in plants. *Am. J. Bot.* 91, 1700–8. doi: 10.3732/ajb.91.10.1700
- Lowe, C. N., and Butt, K. R. (2007). Life-cycle traits of the dimorphic earthworm species *Allolobophora chlorotica* (Savigny, 1826) under controlled laboratory conditions. *Biol. Fertility Soils* 43, 495–9. doi: 10.1007/s00374-006-0154-x
- Lowe, C. N., and Butt, K. R. (2008). *Allolobophora chlorotica* (Savigny, 1826): Evidence for classification as two separate species. *Pedobiologia* 52, 81–4. doi: 10.1016/j.pedobi.2008.04.001
- Maddison, W. P., Knowles, L. L., and Collins, T. (2006). Inferring phylogeny despite incomplete lineage sorting. *Systematic Biol.* 55, 21–30. doi: 10.1080/10635150500354928. directeur.
- Marchán, D. F., Cosin, D. J. D., and Novo, M. (2018). Why are we blind to cryptic species? Lessons from the eyeless. *Eur. J. Soil Biol.* 86, 49–51. doi: 10.1016/j.ejsobi.2018.03.004
- Marchán, D. F., Navarro, A. M., Pinadero, S. J., Gerard, S., Hedde, M., Domínguez, J., et al. (2023). Understanding the diversification and functional radiation of *Aporrectodea* (Crassiclitellata, Lumbricidae) through molecular phylogenetics of its endemic species. *Eur. J. Soil Biol.* 119, 103559. doi: 10.1016/j.ejsobi.2023.103559
- Martinsson, S., and Erseus, C. (2017). Cryptic speciation and limited hybridization within Lumbricid earthworms (Clitellata: Lumbricidae). *Mol. Phylogenet. Evol.* 106, 18–27. doi: 10.1016/j.ympev.2016.09.011
- Martinsson, S., and Erseus, C. (2021). Cryptic clitellata: molecular species delimitation of clitellate worms (Annelida): an overview. *Diversity*. 13, 36. doi: 10.3390/d13020036
- Martinsson, S., Rhodén, C., and Erseus, C. (2015). Barcoding gap, but no support for cryptic speciation in the earthworm *Aporrectodea longa* (Clitellata: Lumbricidae). *Mitochondrial DNA Part A* 28, 147–55. doi: 10.3109/19401736.2015.1115487
- Medlin, L., Elwood, H. J., Stickel, S., and Sogin, M. L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71(12), 491–9. doi: 10.1016/0378-1119(88)90066-2
- Meiklejohn, K. A., Damaso, N., Robertson, J. M., and Fugmann, S. D. (2019). Assessment of BOLD and GenBank – Their accuracy and reliability for the identification of biological materials. *PLoS One* 14, e0217084. doi: 10.1371/journal.pone.0217084
- Palumbi, S. R. (1996). “Nucleic acids II: the polymerase chain reaction,” in *Dans: molecular systematics* (Sinauer & Associates Inc, Sunderland, MA, USA), 205–47.
- Perez-Losada, M., Eiroa, J., Mato, S., and Domínguez, J. (2005). Phylogenetic species delimitation of the earthworms *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* Bouche, 1972 (Oligochaeta, Lumbricidae) based on mitochondrial and nuclear DNA sequences. *Pedobiologia* 49, 317–24. doi: 10.1016/j.pedobi.2005.02.004
- Perez-Losada, M., Ricoy, M., Marshall, J. C., and Domínguez, J. (2009). Phylogenetic assessment of the earthworm *Aporrectodea caliginosa* species complex (Oligochaeta: Lumbricidae) based on mitochondrial and nuclear DNA sequences. *Mol. Phylogenet. Evol.* 52, 293–302. doi: 10.1016/j.ympev.2009.04.003
- Porco, D., Chang, C.-H., Dupont, L., James, S., Richard, B., and Decaens, T. (2018). A reference library of DNA barcodes for the earthworms from Upper Normandy: Biodiversity assessment, new records, potential cases of cryptic diversity and ongoing speciation. *Appl. Soil Ecol.* 124, 362–71. doi: 10.1016/j.apsoil.2017.11.001
- Puillandre, N., Brouillet, S., and Achaz, G. (2021). ASAP: assemble species by automatic partitioning. *Mol. Ecol. Resour* 21, 609–20. doi: 10.1111/1755-0998.13281

- Richard, B., Decaens, T., Rougerie, R., James, S. W., Porco, D., and Hebert, P. D. (2010). Re-integrating earthworm juveniles into soil biodiversity studies: species identification through DNA barcoding. *Mol. Ecol. Resour* 10, 606–14. doi: 10.1111/j.1755-0998.2009.02822.x
- Römbke, J., Aira, M., Backeljau, T., Breugelmans, K., Domínguez, J., Funke, E., et al. (2016). DNA barcoding of earthworms (*Eisenia fetida/andrei* complex) from 28 ecotoxicological test laboratories. *Appl. Soil Ecol.* 104, 3–11. doi: 10.1016/j.apsoil.2015.02.010
- Rota, E., Martinsson, S., Erséus, C., Petushkov, V. N., Rodionova, N. S., and Omodeo, P. (2018). Green light to an integrative view of *Microscolex phosphoreus* (Dugès, 1837) (Annelida: Clitellata: Acanthodrilidae). *Zootaxa*. 4496, 175–89. doi: 10.11646/zootaxa.4496.1.13
- Roy, V., Ventura, M., Fourcade, Y., Justine, J.-L., Gigon, A., and Dupont, L. (2022). Gut content metabarcoding and citizen science reveal the earthworm prey of the exotic terrestrial flatworm, *Obama nungara*. *Eur. J. Soil Biol.* 113, 103449. doi: 10.1016/j.ejsobi.2022.103449
- Shekhovtsov, S. V., Derzhinsky, Y. E. A., Poluboyarova, T. V., Golovanova, E. V., and Peltek, S. E. (2020). Phylogeography and genetic lineages of *Aporrectodea rosea* (Lumbricidae, Annelida). *Eur. J. Soil Biol.* 99, 103191. doi: 10.1016/j.ejsobi.2020.103191
- Shekhovtsov, S. V., Ermolov, S. A., Poluboyarova, T. V., Kim-Kashmenskaya, M. N., Derzhinsky, Y. A., and Peltek, S. E. (2021). Morphological differences between genetic lineages of the peregrine earthworm: *Aporrectodea caliginosa* (Savigny, 1826). *Acta Zool Acad. Sci. Hung* 67, 235–46. doi: 10.17109/AZH.67.3.235.2021
- Shen, H.-P., Tsai, C.-F., Fang, Y.-P., and Chen, J.-H. (2011). Parthenogenesis, polyploidy and reproductive seasonality in the Taiwanese mountain earthworm *Amyntas catenus* (Oligochaeta, Megascolecidae). *Pedobiologia* 54, 133–9. doi: 10.1016/j.pedobi.2010.12.002
- Sims, R. W., and Gerard, B. M. (1999). *Synopsis of the british fauna (31)- earthworms* (London: The Linnean Society of London and the Estuarine and Brackish-water Sciences association), 1–169.
- Škaloud, P., and Rindi, F. (2013). Ecological differentiation of cryptic species within an asexual protist morphospecies: A case study of filamentous green alga *klebsormidium* (Streptophyta). *J. Eukaryotic Microbiol.* 60, 350–62. doi: 10.1111/jeu.12040
- Sturzenbaum, S. R., Andre, J., Kille, P., and Morgan, A. J. (2009). Earthworm genomes, genes and proteins: the (re)discovery of Darwin's worms. *P R Soc. B* 276, 789–97. doi: 10.1098/rspb.2008.1510
- Tamura, K., Stecher, G., Kumar, S., and Battistuzzi, F. U. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–7. doi: 10.1093/molbev/msab120
- Tan, D. S. H., Ang, Y., Lim, G. S., Ismail, M. R. B., and 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. doi: 10.1111/j.1463-6409.2009.00408.x
- Terhivuo, J., and Saura, A. (2006). Dispersal and clonal diversity of North-European parthenogenetic earthworms. *Biol. Invasions* 8, 1205–18. doi: 10.1007/s10530-006-9015-7
- Torres-Leguizamon, M., Mathieu, J., Decaens, T., and Dupont, L. (2014). Genetic structure of earthworm populations at a regional scale: inferences from mitochondrial and microsatellite molecular markers in *Aporrectodea icterica* (Savigny 1826). *PLoS One* 9, e101597. doi: 10.1371/journal.pone.0101597
- Torres-Leguizamon, M., Mathieu, J., Livet, A., and Dupont, L. (2012). Isolation of polymorphic microsatellite markers in *Aporrectodea icterica* (Savigny 1826). *Soil Biol. Biochem.* 51, 16–9. doi: 10.1016/j.soilbio.2012.03.020
- Twyford, A. D., and Ennos, R. A. (2012). Next-generation hybridization and introgression. *Heredity* 108, 179–89. doi: 10.1038/hdy.2011.68
- Vlasenko, R. P., Mezhzherin, S., Garbar, A., and Kotsuba, I. Y. U. (2011). Polyploid races, genetic structure and morphological features of earthworm *Aporrectodea rosea* (Savigny, 1826) (Oligochaeta, Lumbricidae) in Ukraine. *CCG* 5, 91–103. doi: 10.3897/compctogen.v5i2.968
- Zaborski, E. R. (2003). Allyl isothiocyanate: an alternative chemical expellant for sampling earthworms. *Appl. Soil Ecol.* 22, 87–95. doi: 10.1016/S0929-1393(02)00106-3