



A strong backbone for an invertebrate group: anchored phylogenomics improves the resolution of genus-level relationships within the Lumbricidae (Annelida, Crassicitellata)

Daniel F. Marchán^{1,2} · Samuel W. James² · Alan R. Lemmon³ · Emily Moriarty Lemmon⁴ · Marta Novo⁵ · Jorge Domínguez⁶ · Darío J. Díaz Cosín⁵ · Dolores Trigo⁵

Received: 28 February 2022 / Accepted: 20 July 2022

© Gesellschaft für Biologische Systematik 2022

Abstract

Earthworm taxonomy has shown great instability across the twentieth century, as exemplified by the numerous genera system of Lumbricidae. Sanger-sequenced “legacy markers” have allowed to recover genus-level clades within Lumbricidae, but relationships between genera were often recovered as polytomies. Anchored hybrid enrichment (AHE) is a recent source for phylogenomic analyses, providing hundreds of unique orthologous loci from across the genome with several advantages (like informativeness across different taxonomic scales, cost-effectiveness, no necessity of special preservation protocols). In this work, AHE was applied to 31 Lumbricidae genera to validate recent genus-level changes and to provide a solid “backbone” for the systematics of the family. Five hundred ninety-four non-duplicated loci were sequenced, and phylogenomic inference was performed using the concatenated alignments and independent loci. A “legacy marker” dataset was used to study the effect of the inclusion of the AHE phylogenomic tree as constraint. Lumbricidae was recovered as a monophyletic family, with *Diporodrilus pilosus* as the earliest branching taxa, followed by two large clades. The first one comprised the Franco Iberian genera *Castellodrilus*, *Cataladrilus*, *Prosellodrilus*, *Zophoscolex*, *Ethnodrilus*, *Kritodrilus*, *Postandrilus*, *Galiciandrillus*, and *Compostelandrilus*; the Italian *Pietromodeona*; and the Central European *Vindoboscolex*. The second one included the rest of the studied Lumbricidae genera, with relationships between them being well resolved and strongly supported. The phylogenomic backbone improved the topological resolution and support of the legacy marker phylogenetic tree. Thus, AHE proved highly suitable for phylogenomic inference in lumbricid earthworms and closely related taxa, allowing to propose changes to the above-genus-level systematics of the family.

Keywords Earthworms · Phylogenomics · Systematics · Targeted next-generation sequencing

Introduction

Until the advent of molecular phylogenetics, earthworm taxonomy and systematics were hindered by the limited (but steadily increasing) set of useful morphological characters. For the Lumbricidae Rafinesque-Schmaltz 1915 (the most common fauna in the Palearctic), the increased emphasis on characters such as reproductive organs (Michaelsen, 1900), pigmentation, longitudinal musculature, arrangement of setae (Pop, 1941), digestive tract structures, chromosome number (Omodeo, 1956), nephridia (Gates, 1975; Perel, 1979), and the implementation of cladistics principles (Mršić, 1991; Qiu & Bouché, 1998a) gave rise to consecutive genera systems which were not universally accepted by earthworm taxonomists. The first molecular phylogenetic studies showed the potential for clarifying the relationships

✉ Daniel F. Marchán
danief01@ucm.es; danifermch@gmail.com

¹ CEFE, Univ Montpellier, CNRS, EPHE, Montpellier, IRD, France

² Present Address: Maharishi International University, Fairfield, USA

³ Department of Scientific Computing, Florida State University, Tallahassee, FL, USA

⁴ Department of Biological Science, Florida State University, Tallahassee, FL, USA

⁵ Depto. de Biodiversidad, Facultad de Ciencias Biológicas, Universidad Complutense de Madrid, Ecología y Evolución José Antonio Novais, 2, 28040 Madrid, Spain

⁶ Grupo de Ecología Animal (GEA), Universidade de Vigo, 36310 Vigo, Spain



between earthworm genera (James & Davidson, 2012; Pop et al., 2007). The comprehensive study of the Lumbricidae by Domínguez et al. (2015) increased the sampling of genera (to 28) and molecular markers (to eight). This study revealed that at least 10 genera were non-monophyletic as described in morphology-based taxonomy, stressing the need for a thorough systematic revision. Since then, several works have relied on the same set of molecular markers (nowadays referred to as “legacy markers” by some authors) and phylogenetic inference techniques, and have progressively added rare, narrowly restricted species and genera to fill the gaps in the phylogeny of the Lumbricidae (De Sosa et al., 2019; Domínguez et al., 2018; Jiménez et al., 2021; Marchán et al., 2021a, b). These studies demonstrated the suitability of legacy markers to cluster species within genus-level clades and to resolve their phylogenetic relationships below this level. The recovery of species formerly assigned to single genera within phylogenetically unrelated genus-level clades allowed to create several new genera (or to elevate former taxa to genus level) such as *Galiciandrilus* Domínguez et al. (2018), *Compostelandrilus* Domínguez et al. (2018), *Castellodrilus* Jiménez et al. (2021), *Gatesona* Marchán (2021), and *Vindoboscolex* Marchán (2021). However, relationships between some genera (*Dendrobaena* Eisen, 1873; *Octodrilus* Omodeo, 1956; *Octolasion* Örley, 1885; *Helodrilus* Hoffmeister, 1845; *Allolobophora* Eisen, 1873; *Eiseniella* Michaelsen, 1900; and *Lumbricus* Linnaeus, 1758 amongst others) were usually recovered as polytomies, and some well-known species (such as *Aporrectodea rosea* (Savigny, 1826)) behaved as rogue taxa, with no clear phylogenetic position. It could be argued that phylogenetic hypotheses based on a small sample of genes across the genome may fail to represent the true phylogenetic relationships of ancient taxa (Phillips et al., 2019), especially when phenomena such as incomplete lineage sorting are taken into account. However, Anderson et al. (2017) used a phylogenomic approach (transcriptomes) to reconstruct the phylogenetic relationships of a similar set of taxa featured in James and Davidson (2012), finding substantial support for the results of legacy marker analyses. On the other hand, Novo et al. (2016) applied transcriptomics to the family Hormogastridae Michaelsen, 1900, recovering between-genus relationships which were unresolved by legacy markers (Novo et al., 2012).

Anchored hybrid enrichment (AHE) (Lemmon et al., 2012) is a recent source for phylogenomic analyses, providing hundreds of unique orthologous loci from across the genome. Amongst its advantages are its efficiency in non-model species, the high phylogenetically informative content of the loci across different taxonomic scales, the potentially low levels of missing data, rapid data collection, cost-effectiveness (Hamilton et al., 2016), and the lack of need for special preservation protocols for RNA sources.

In this study, AHE was implemented as a new phylogenomic tool for earthworm systematics, with the following objectives: (i) to test the suitability of the approach for this animal group, (ii) to support the recent genus-level changes proposed for the systematics of Lumbricidae by molecular phylogenetic analyses based on legacy markers, and (iii) to produce a solid “backbone” for the phylogeny of Lumbricidae consisting of well-supported relationships between the main genera.

Materials and methods

Taxon sampling

Representatives of 31 Lumbricidae genera were chosen (Fig. 1, Supplementary File 1) and obtained from the UCMLT collection of the Universidad Complutense de Madrid, the collection of the Animal Ecology Group of the University of Vigo (GEA), and North Carolina State University Museum (NCSM); for the genera *Proselodrilus* Bouché (1972), *Cernosvitovia* Omodeo (1956), *Vindoboscolex*, *Compostelandrilus*, *Gatesona*, and *Aporrectodea* Orley (1885), two or three (for *Aporrectodea*) representatives were chosen in order to test internal relationships.

Representatives of the closest families were chosen as outgroups: *Lutodrilus* sp. (Lutodrilidae McMahan, 1976), *Criodrilus lacuum* (Hoffmeister, 1845) (Criodrilidae Vejdovsky, 1884), *Ailoscolex lacteospumosus* Bouché, 1969, *Hemigastrodrilus monicae* Bouché, 1970, and *Vignysa teres* Duges, 1828 (Hormogastridae).

Library construction and sequencing

Total genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) from ventral integument samples of approximately 5 × 5 mm. Library preparation and AHE enrichment were performed at the Center for Anchored Phylogenomics (www.anchoredphylogeny.com) following Lemmon et al. (2012) and Prum et al. (2015). In brief, extracted DNA was sonicated to a fragment size of 150–500 using a Covaris ultrasonicator. Dual-indexed libraries (8 bp per side) were constructed using a Beckman-Coulter FXp Liquid handling robot and were then enriched using the Agilent SureSelect XP kit recently developed for Annelids (Phillips et al., 2019; Taheri et al., 2018). This kit targets ~594 loci totalling 100 kbp. Libraries were pooled into groups of ~16 prior to enrichment. The enriched libraries were sequenced at the Translational Lab at the FSU College of Medicine on an Illumina NovaSeq6000 sequencer with a PE150bp protocol. A total of 45 Gb of raw sequence data were collected.

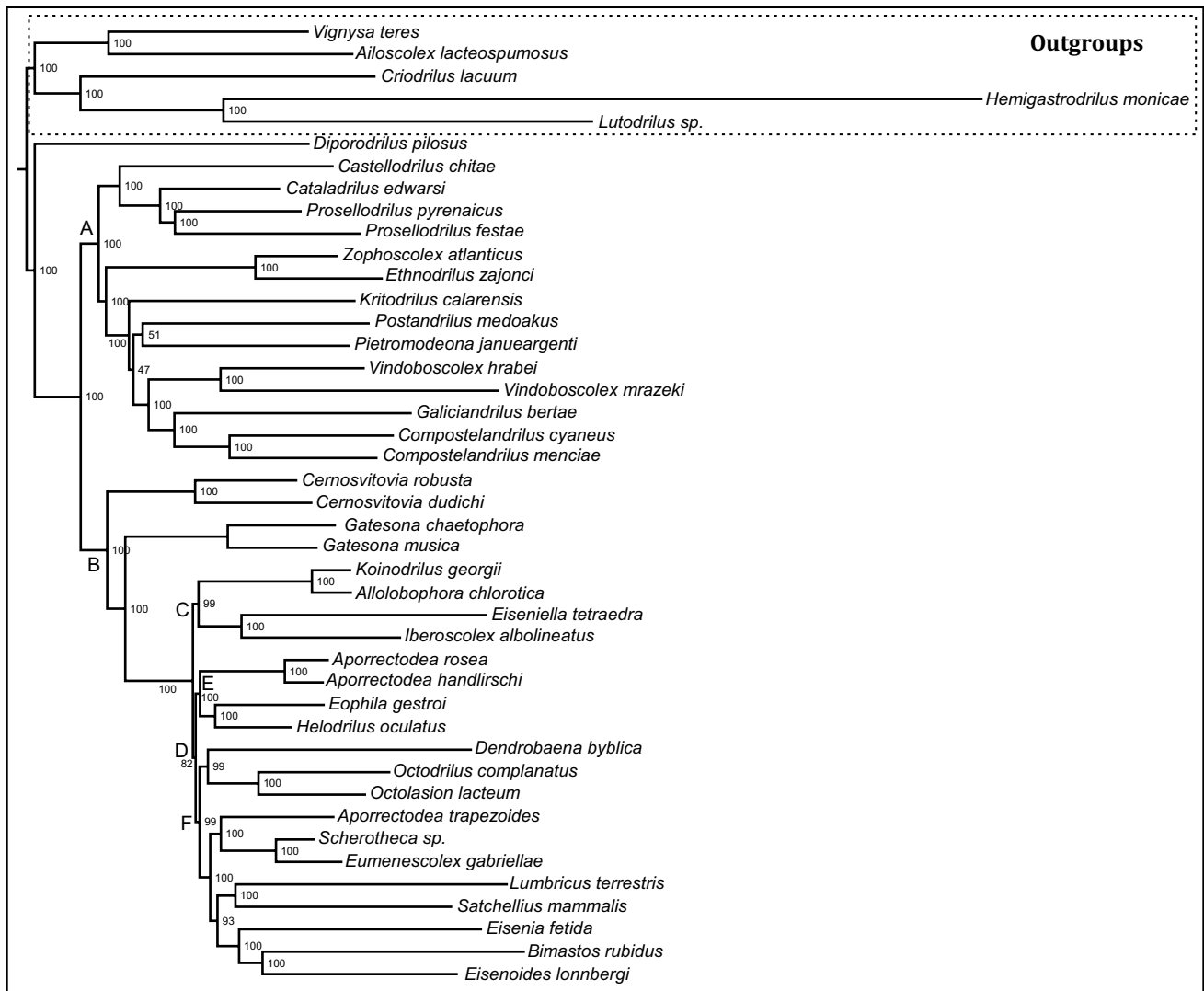


Fig. 1 Best-scoring maximum likelihood tree obtained from the concatenated anchored hybrid enrichment loci. Bootstrap values are shown beside the nodes. Clades A–F are referred to in the main text

Sequence processing

Overlapping sequence read pairs were merged following Rokyta et al. (2012). This process corrects sequencing errors and trims adapters. The resulting reads were assembled using the quasi-de novo assembler described by Hamilton et al. (2016). With this process, reads were mapped to loci using three divergent references representing Clitellata (from *Dendrobaena* sp., *Helobdella* sp., and *Mesenchytraeus* sp.). Assembly clusters containing few reads (fewer than 242) were removed for downstream analysis. For each targeted locus, the orthology of the retained consensus sequences was based on pairwise sequence similarity (see Hamilton et al., 2016 for details). Sequences in each orthologous set were aligned using MAFFT v7.023b (Katoh & Standley, 2013). The resulting alignments were trimmed and masked following Hamilton

et al. (2016), with MINGOODSITES = 14 and MISSINGALLOWED = 0.5 settings.

Phylogenomic analyses

Phylogenetic trees were first estimated using a maximum likelihood (ML) approach, as implemented in RAXML v2.2.3 (Stamatakis, 2006). In addition to a single phylogeny estimated from concatenated alignments (with the GTR + G model of sequence evolution partitioned by locus), gene-specific trees were estimated from locus-specific alignments (with a single GTR + G model assumed). One hundred bootstrap replicates were performed for each ML analysis. Those trees were used as input in ASTRAL v5.7 (Mirarab & Warnow, 2015) for species tree estimation.

In order to test the effect of incorporating the phylogenomic-derived phylogenetic relationships into taxa-dense legacy marker analysis, Bayesian inference was performed in MrBayes 3.1.2 (as implemented in the CIPRES Science Gateway V. 3.3). The dataset analyzed was obtained from Domínguez et al. (2015), Pérez-Losada et al. (2015), Domínguez et al. (2018), De Sosa et al. (2019); Jiménez et al. (2021), and Marchán et al. (2021a, b) and consisted of the nuclear marker 28S rRNA and the mitochondrial 16S rRNA, NADH dehydrogenase (*ND1*), 12S rRNA, and *COI*. The best fitting evolutionary model for each partition was selected with jModelTest v. 2.1.3 (Darriba et al., 2012) by applying the Akaike information criterion (AIC, Akaike, 1973), and Bayesian information criterion (BIC, Schwarz, 1978). GTR+I+G was selected as the best-fitting evolutionary model for COI, 28S, and ND1; GTR+G was selected for 12S; and HKY+I+G was selected for 16S. The best tree obtained from the concatenated ML analyses was implemented as a partial constraint (Slater, 2013), which constrains the topology of some taxa (those shared between datasets) while allowing others to vary freely. An unconstrained analysis was also performed to compare the obtained topologies and support values. Parameters were set to 50 million generations and sampled every 5000th generation (10,000 trees). Two independent runs each with four chains were performed and 20% of the trees were discarded as burn-in. The remaining trees were combined and summarized on a 50% majority-rule consensus tree.

Results

Locus assembly and alignment

Using the AHE approach, a large majority of the targeted loci were recovered for most of the samples (at least 90% of the loci were recovered for > 97% of the samples, and consensus sequences included at least 500 bp for > 90% of the loci). The trimming and masking procedure produced 662 alignments (in total 230,748 bp) with 17% missing characters. Note that the number of loci recovered exceeded the target number (594) because of gene duplications, which were resolved during the orthology assessment.

Species tree estimation

The best-scoring maximum likelihood tree obtained from the concatenated loci recovered a monophyletic Lumbricidae including *Diporodrilus pilosus* Bouché, 1972 as the most basally branching taxon (Fig. 1). Two strongly supported clades were recovered within the Lumbricidae. The first clade (A) included the Franco Iberian genera *Castellodrilus*, *Cataladrilus* Qiu and Bouché (1998a, b, c), *Prosello-drilus*, *Zophoscolex* Qiu & Bouché (1998a, b, c), *Ethnodrilus*

Bouché (1972), *Kritodrilus* Bouché (1972), *Postandrilus* Qiu and Bouché (1998a, b, c), *Galiciandrilus*, and *Compostelandrilus*; the Italian *Pietromodeona* Qiu and Bouché (1998a, b, c) and the Central European *Vindoboscolex*. Within this clade, *Castellodrilus* appeared closely related to *Cataladrilus* and *Prosello-drilus* (including *Prosello-drilus festae* Rosa, 1892), *Zophoscolex* and *Ethnodrilus* were recovered as closely related sister taxa, and *Vindoboscolex* (*Vi. hrabei* Černosvitov, 1935 and *Vi. mrazeki* Černosvitov, 1935) appeared closely related to *Galiciandrilus* and *Compostelandrilus*.

The second clade (B) included *Cernosvitovia* Omodeo (1956) and *Gatesona* as the earliest branching clades. The relationships between the other Lumbricidae genera were well resolved and strongly supported. The representatives of *Aporrectodea* were recovered in two separate clades, with *Aporrectodea trapezoides* Duges, 1828 (type species of the genus) closely related to *Scherotheca* Bouché (1972) and *Aporrectodea rosea* and *Aporrectodea handlirschi* (Rosa, 1897) as sister taxa. Some sister relationships already recovered in previous analyses were supported, such as *Eiseniella* and *Iberoscolex* Qiu and Bouché (1998a, b, c), *Dendrobaena* and *Octodrilus-Octolasion*, and *Eisenia* Malm (1877), *Bimastos* Moore (1893) and *Eisenoides* Gates (1969). However, other novel relationships were discovered, such as those between *Koinodrilus georgii* (Michaelsen, 1890) and *Allolobophora chlorotica* (Savigny, 1826), *Scherotheca* and *Eumenescolex* Qiu and Bouché (1998a, b, c), and *Lumbricus* and *Satchellius* Gates (1975).

The species tree obtained from the individual gene trees in ASTRAL was mostly congruent (Suppl. File 2), recovering the same clades and sister taxa relationships. The only exceptions were the placement of *Diporodrilus* Bouché, 1970 as a sister taxon of *Criodrilus* Hoffmeister, 1845 plus Hormogastridae, and the position of *Satchellius* closer to *Aporrectodea* plus *Scherotheca-Eumenescolex* than to *Lumbricus*. Support values were generally similar, but lower in some cases.

As expected, the addition of the phylogenomic backbone to the Bayesian phylogenetic inference of the legacy marker dataset changed the topology of the resulting tree to reflect the improved between-genus resolution of the former (Fig. 2). These changes were less noticeable for the earliest branching clades, as they were already well supported and bifurcating in the unconstrained analysis. Within clade B (see above), several polytomies disappeared, as dichotomous relationships were enforced between several representatives of the included genera. Interestingly, some relationships not directly enforced by the constraint (such as non-*byblica* *Dendrobaena* + *Octodrilus-Octolasion*, or *Helodrilus* + *Eophila* + *Proctodrilus* + Eastern *Helodrilus*) were now recovered.

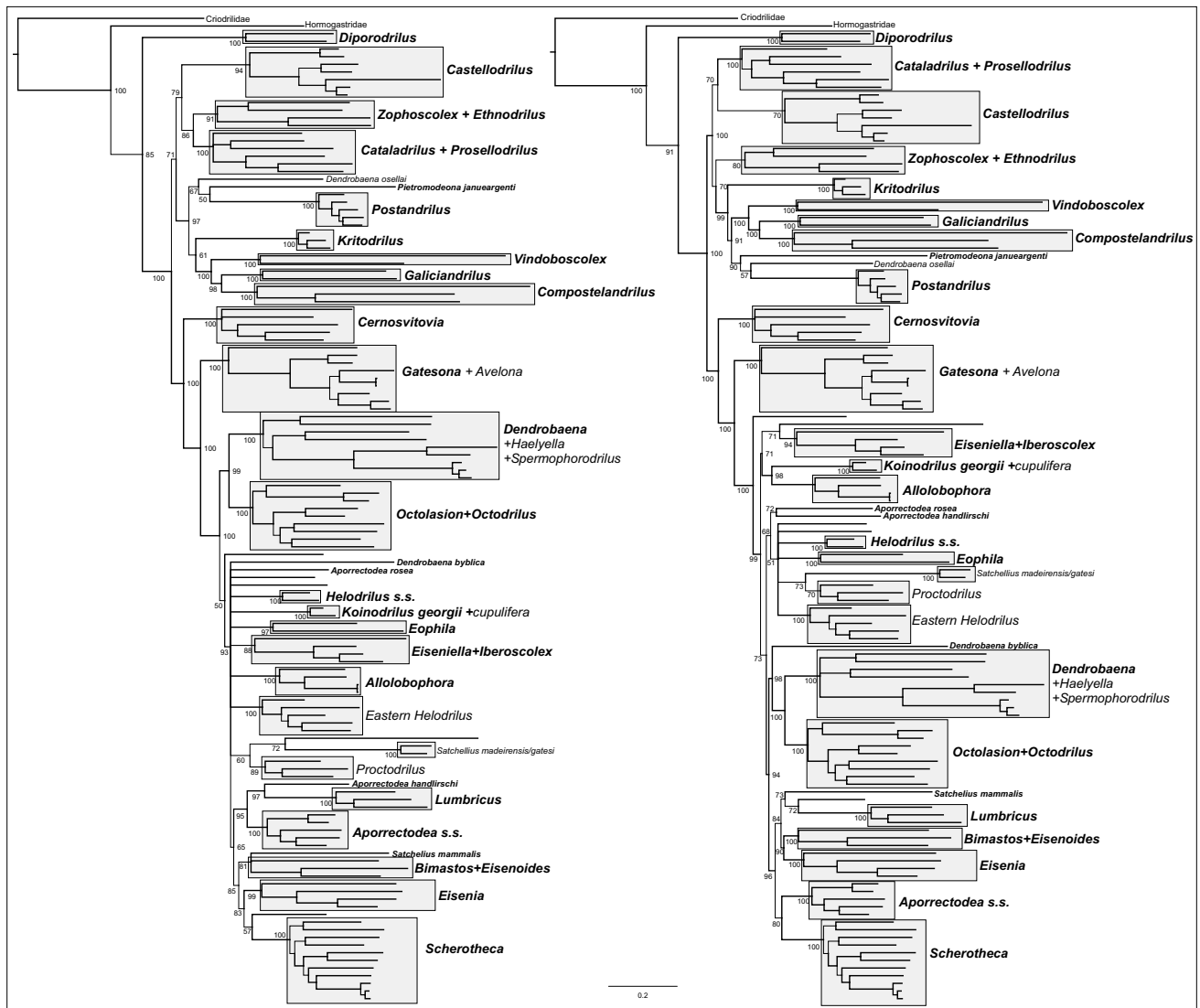


Fig. 2 Left. Bayesian inference of the phylogenetic tree based on the concatenated sequences of the nuclear marker 28S rRNA and the mitochondrial 16S rRNA, NADH dehydrogenase (*ND1*), 12S rRNA, and *COI*. Right. Phylogenetic tree based on the same analysis but imple-

menting the best scoring tree obtained from the concatenated maximum likelihood analysis of the anchored hybrid enrichment loci as a constraint. Genera and genus-level clades are indicated by gray squares. Posterior probability values are shown beside the nodes

Discussion

Methodological performance

Based on the large number of sequenced loci, the strong coverage across all of the studied taxa and loci, and the robust phylogenetic results obtained, AHE appears to be a highly suitable sequencing approach for phylogenomic inference in lumbricid earthworms and closely related taxa. AHE has previously been applied to the earthworm genus *Pontoscolex* Schmarda, 1861 (Taheri et al., 2018), although at a much narrower systematic scale. In the current study, the loci number was similar (594 vs 609) and

significantly larger than in AHE studies of other clitellates (301 loci, Phillips et al., 2019).

Phylogenetic inference methods based on the AHE dataset showed enhanced resolution of the deep nodes of the family Lumbricidae, revealing well-supported sister relationships between genera in cases in which Sanger-based multigene phylogenetic trees recovered polytomies, as well as placing with confidence elusive “rogue” taxa. This is consistent with the findings of Phillips et al. (2019) who showcased the utility of AHE for resolving deep phylogenetic relationships relative to individual or concatenated legacy markers. Although individual AHE loci appeared to generate different degrees of noise (losing phylogenetic

information), this appeared to be cancelled out by high levels of signal in the total dataset.

Interestingly, phylogenetic relationships between the earliest branching taxa (mostly within clade A) were congruent with the topologies recovered by “legacy” multigene phylogenetic trees. On the one hand, this confirms most of the recently adopted systematic decisions based on 3–7 mitochondrial and nuclear datasets (De Sosa et al., 2019; Domínguez et al., 2018; Jiménez et al., 2021; Marchán et al., 2021a, b), thus reinforcing their usefulness for resolving relationships across a large portion of the Lumbricidae tree. On the other hand, the difference in performance of traditional markers in a problematic crown group vs the rest of the phylogenetic tree suggests some different underlying evolutionary patterns. Polytomies have frequently been associated with rapid cladogenetic events, saturation of the phylogenetic signal, and incomplete lineage sorting (Janko et al., 2011; Mendes et al., 2016; Stanley et al., 2011). The short branch lengths observed in clades C–F (Fig. 1) are also compatible with those phenomena. Incomplete lineage sorting (ILS) could be prevalent amongst the “legacy” marker dataset, but it does not appear to be widespread in the AHE dataset: very few incongruences between the concatenated and the ASTRAL species tree were observed. The most important conflict was recovered at the very root of the tree (the position of *Diporodrilus* in relation to Lumbricidae and outgroup families) and not in the crown group. This particular incongruence could be explained by ILS during a rapid early divergence of the aforementioned taxa, as observed in Chen et al. (2020).

The hypothesis of rapid cladogenetic events (either at the root or at the crown of the tree) could be explored by further research implementing divergence time estimation to this AHE dataset. Although calibration presents several challenges in clitellates, different approaches have been applied (Marchán et al., 2021b), and the wealth of available loci from which to select those with the most clock-like behavior (Smith et al., 2018) should provide more consistent time-calibrated phylogenetic trees.

Improvement in the phylogenetic resolution of Sanger-based phylogenetic trees by the addition of an AHE phylogenomic backbone appears promising as a compromise between accuracy and taxon coverage. There is currently a database including more than 125 Lumbricidae species for at least three of the traditional Sanger molecular markers, and generating phylogenomic datasets for all of them (even with the affordability of AHE) would be unrealistic. However, integrating them with a more robust, phylogenomics-derived constraint may be a more efficient approach. There is still room for improvement, as some deep, genus-level clades still lack AHE representatives. Adding one or more (if high internal divergence is detected) representatives of each of those clades, as well as any elusive, “rogue” taxa

which resists phylogenetic assignment to the clades should be an attainable, pragmatic goal aimed at finally resolving the slippery systematics of lumbricid earthworms.

Systematic implications

The sister taxon relationship and relatively short branches between the representatives of *Gatesona* (*Ga. chaetophora* (Bouché, 1972) and *Ga. musica* (Qiu & Bouché, 1998a, b, c)), recently proposed to belong to the same genus (Marchán et al., 2021b) but formerly separated in the genera *Gatesona* and *Helodrilus* (*Acystodrilus*) Qiu and Bouché (1998a, b, c), support said systematic revision. The same can be stated about the two representatives of *Cernosvitovia* (*Ce. dudichi* Zicsi & Šapkarev, 1982 and *Cernosvitovia robusta* (Rosa, 1895)), which belonged to the genera *Cernosvitovia* and *Serbiona* respectively until their revision by Popovic et al. (2022). Interestingly, the pairs of representatives of *Proselodrilus*, *Vindoboscolex*, and *Compostelandrilus* were also recovered as sister taxa but with comparatively longer branches. The internal divergence of these genera may be more ancient, explaining the greater amount of genetic divergence between their representatives. On the other hand, their placement within the same genera may have been excessively conservative. The proportion of genera with two or more representatives within this dataset is too low to draw conclusions. Further work on time-calibrated trees will confirm whether the reconstructed branch lengths are actually proportional to the estimated divergence time and whether this difference in depth to genus root will be maintained.

Some sister taxon relationships between representatives of different genera had already been recovered in “legacy” marker phylogenetic analyses, such as *Zophoscolex* and *Ethnodrilus*, *Eiseniella* and *Iberoscolex*, *Eophila* Rosa, 1893 and *Helodrilus*, *Dendrobaena* + *Octodrilus* + *Octolasion*, *Eisenia* + *Bimastos* + *Eisenoides* (Csuzdi et al., 2017; de Sosa et al., 2019; Jiménez et al., 2021; Marchán et al., 2021a). However, other noteworthy relationships between sister taxa were unprecedented. *Koinodrilus georgii* and *Allolobophora chlorotica* were recovered as a relatively shallow clade, suggesting they may actually belong to the same genus. This is not inconsistent with some previous classifications: *Koinodrilus georgii* was originally placed in *Allolobophora*, and the differences between the species ascribed to *Koinodrilus* Qiu and Bouché (1998a, b, c) and *Allolobophora* by Qiu and Bouché (1998b, c) are extremely subtle. *Aporrectodea rosea* (sometimes assigned to *Koinodrilus*) and *Aporrectodea handlirschi* (sometimes assigned to *Eiseniella* Omodeo, 1956) were also recovered as sister groups with very short branches. Whatever genus they actually belong to (as the type species of *Aporrectodea* is the unrelated *A. trapezoides*), it is rather obvious that they constitute a single genus. The addition of further related species will be

necessary for systematic revision of this genus-level clade. The close relationship between *Scherotheca* and *Eumenescolex* confirms previous suspicions (Marchán et al., 2020) and could even indicate they belong to the same genus. Interestingly, this clade was recovered as sister to *Aporrectodea*, with which they share most of their native range. Finally, *Lumbricus* and *Satchellius* also have overlapping ranges and somewhat similar morphological adaptations to epigeic (litter-dweller) lifestyles.

The robust resolution of the above-genus phylogenetic relationships of Lumbricidae enables revision of the previous classifications proposed on the basis of morphology alone. As the most recent one, presented by Qiu and Bouché (1998a), is the most complete in taxon coverage and based on the largest set of characters, it should be considered the most relevant for this discussion.

The genus *Diporodrilus* was considered a separate family from the Lumbricidae, probably because of the significant difference in the dorsal pore arrangement (two rows vs one). The alternative classification would be the subdivision of the Lumbricidae into the subfamilies Diporodrilinae Bouché, 1970 and Lumbricinae Rafinesque-Schmaltz, 1915 (encompassing all the other lumbricid genera). Our findings do not show unambiguous support for one or the other: the position of *Diporodrilus* as a sister taxon for the remaining Lumbricidae in the concatenated tree could be interpreted as an independent (but closely related) family or as a subfamily. Furthermore, the ASTRAL species tree topology would suggest that Diporodrilidae constitutes a separate family closer to the outgroups. Even though the Diporodrilinae hypothesis appears more consistent with overall morphology of the taxa (only significant difference being the dorsal pore disposition), the phylogenetic uncertainty advises to leave this systematic conundrum open for the moment.

The subfamilies Postandrilinae Qiu and Bouché (1998a, b, c) and Spermophorodrilinae Omodeo and Rota, 1989, as established by Qiu and Bouché (1998a, b, c), do not appear to be supported by the phylogenomic trees. Different genera assigned to Postandrilinae (*Postandrilus*, *Galiciandrilus*, and *Cernosvitovia*) were recovered as unrelated phylogenetic lineages. Representatives of Spermophorodrilinae were not included in the present phylogenomic analyses, but they were included in the constrained “legacy” marker analysis: it recovered *Spermophorodrilus antiquus* (Černosvitov, 1938) within *Dendrobaena* and unrelated to *Bimastos* (another member of Spermophorodrilinae). Thus, a monophyletic Spermophorodrilinae is not supported.

Qiu and Bouché (1998a) proposed a system of 14 tribes, most of which were recovered as polyphyletic or paraphyletic by our results. Hence, it appears wiser to propose a new tribe system, which reflects the stable, well-supported subdivision of the Lumbricinae genera into two large clades (A and B in Fig. 1). The already established tribes

Prosello-drilini Qiu and Bouché (1998a, b, c) and Lumbricini Qiu and Bouché (1998a, b, c) would be re-defined to include all of the genera most closely related to *Prosello-drilus* and to *Lumbricus* respectively. Such tribes have the upside of being supported by a shared morphological character state: representatives of Prosello-drilini have either calciferous glands dilatations or diverticula in 11 or absence of them, while representatives of Lumbricini have calciferous glands dilatations or diverticula in segment 10 (rarely absent).

The constrained “legacy” marker analysis allowed to recover the phylogenetic position of genera for which no representatives could be added to the AHE dataset (such as *Avelona*, *Proctodrilus*, *Spermophorodrilus*, and *Healyella*), allowing to confidently place them within the redefined Lumbricini. This constitutes an example of the combined approach for the systematic revision of Lumbricidae.

Proposed taxonomic changes

Phylum Annelida Lamarck, 1802.

Class Oligochaeta Grube, 1850/Clitellata Michaelsen, 1919.

Order Crassicitellata Jamieson, 1988.

Family Lumbricidae Rafinesque-Schmaltz, 1815.

Included tribes: Prosello-drilini, Lumbricini.

Tribe Prosello-drilini Qiu & Bouché, 1998a, b, c

Diagnostic characters: Calciferous glands dilatations or diverticula in 11, sometimes absent.

Type genus: *Prosello-drilus*.

Included genera: *Castello-drilus*, *Cataladrilus*, *Prosello-drilus*, *Zophoscolex*, *Ethnodrilus*, *Kritodrilus*, *Postandrilus*, *Pietroomodeona*, *Vindoboscolex*, *Galiciandrilus*, *Compostelandrilus*.

Tribe Lumbricini Qiu & Bouché, 1998a, b, c

Diagnostic characters: Calciferous glands dilatations or diverticula in 10 (rarely absent).

Type genus: *Lumbricus*.

Included genera: *Cernosvitovia*, *Avelona** Qiu & Bouché, 1998a, b, c, *Gatesona*, *Allolobophora*, *Eiseniella*, *Iberoscolex*, *Aporrectodea*, *Eophila*, *Helodrilus*, *Proctodrilus** Zicsi, 1985, *Dendrobaena*, *Octodrilus*, *Octolasion*, *Aporrectodea*, *Scherotheca*, *Eumenescolex*, *Lumbricus*, *Satchellius*, *Eisenia*, *Bimastos*, *Eisenoides*, *Spermophorodrilus**, *Healyella** Omodeo and Rota, 1989.

Remarks: Representatives of genera *Avelona*, *Proctodrilus*, *Spermophorodrilus*, and *Healyella* were not included in the phylogenomic analyses, but their close position to *Gatesona*, *Helodrilus*, and *Dendrobaena* in the constrained “legacy” marker analysis supports their inclusion in this tribe. The genera *Orodrilus* Bouché, 1972, *Microeophila* Omodeo, 1956, *Perelia* Easton, 1983, *Rhiphaeodrilus* Csuzdi and Pavlíček, 2005,

Norealidys Blakemore, 2008 and *Murchieona* Gates, 1978 probably belong to the tribe Lumbricini but should be included in molecular phylogenetic analysis before this can be stated unambiguously.

Conclusions

Anchored hybrid enrichment is a highly suitable phylogenomic approach to resolve the above-genus-level relationships within Lumbricidae and closely related taxa. The addition of the resulting phylogenomic backbone to the phylogenetic inference of a legacy marker dataset improved the topological resolution and support within the most conflictive section of the phylogenetic tree.

The strongly supported relationships recovered in the AHE trees allowed to propose changes to the above-genus-level systematics of the family Lumbricidae and displayed putative underlying evolutionary patterns which differed between its tribes (Proselodrilini and Lumbricini).

The combination of a taxon-rich “legacy markers” dataset with an AHE phylogenomic backbone (which should include representatives from additional genera) appears as a short-term attainable goal for resolving Lumbricid systematics.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13127-022-00570-y>.

Acknowledgements We thank Michelle Kortyna at the Center for Anchored Phylogenomics for assistance with data collection.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by DFM, ARL, and EML. The first draft of the manuscript was written by DFM and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by Xunta de Galicia, Consellería de Cultura, Educación e Ordenación Universitaria, Secretaría Xeral de Universidades under grant ED431B 2019/038, and by subproject Fauna Iberica XII. Oligochaeta, Lumbricoidea: Lumbricidae, Hormogastridae #PGC2018-095851-B-C66 from the Spanish Ministry of Sciences, Innovation and Universities. DFM was funded by a Juan de La Cierva-Formación grant (FJCI-2017-32895) and a Maria Zambrano postdoctoral grant from the Spanish Ministry of Sciences, Innovation and Universities and by a Make Our Planet Great Again postdoctoral grant from Campus France, and MN was supported by the Spanish Government through Juan de la Cierva and Ramon y Cajal Fellowships.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval No approval of research ethics committees was required to accomplish the goals of this study because experimental work was conducted with an unregulated invertebrate species.

Competing interests The authors declare no competing interests.

References

- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255–265.
- Anderson, F. E., Williams, B. W., Horn, K. M., Erséus, C., Halanych, K. M., Santos, S. R., & James, S. W. (2017). Phylogenomic analyses of Crassieclitellata support major Northern and Southern Hemisphere clades and a Pangaeian origin for earthworms. *BMC Evolutionary Biology*, *17*(1), 123.
- Blakemore, R. J. (2008). Replacement of *Reynoldsia* Qiu & Bouche, 1998 (preocc.) with *Norealidys* Blakemore, 2008 (Oligochaeta, Lumbricidae). In R. J. Blakemore (Ed.), *A series of searchable texts on earthworm biodiversity, ecology and systematics from various regions of the world* (3rd ed.).
- Bouché, M. B. (1969). *Ailoscolex lacteosporus*, n. gen., n. sp. Un ver de terre aux caractères morphologiques et biologiques remarquables (Oligochaeta, Ailoscolecidae, nov. fam.). *Revue d'Écologie et Biologie du Sol*, *4*(4), 525–531.
- Bouché, M. B. (1970). Remarques sur quelques Lumbricina de France et conséquences de la découverte des nouveaux taxons *Vignysinae* (Subfam. nov.) et *Diporodrilidae* (Fam. nov.). *Pedobiologia*, *10*(4), 246–256.
- Bouché, M. B. (1972). Lombriciens de France. Écologie et systématique. Institut national de la recherche scientifique (Ed.), 671 pp.
- Černosvitov, L. (1935). Monographie des tschechoslovakischen Lumbriciden. *Archiv pro Prirodovedecký Vyzkum Cech*, *19*, 1–86.
- Černosvitov, L. (1938). Zur Kenntnis der Oligochaetenfauna des Balkans. VI. Oligochaeten aus Griechenland. *Zoologischer Anzeiger*, *CXXIII*, 192–200.
- Chen, Y. C., Li, Z., Zhao, Y. X., Gao, M., Wang, J. Y., Liu, K. W., & Wang, Y. D. (2020). The *Litsea* genome and the evolution of the laurel family. *Nature Communications*, *11*(1), 1–14.
- Csuzdi, C., Chang, C. H., Pavlíček, T., Szederjesi, T., Esopi, D., & Szilávecz, K. (2017). Molecular phylogeny and systematics of native North American lumbricid earthworms (Clitellata: Megadrili). *PLoS ONE*, *12*(8), e0181504.
- Csuzdi, Cs., & Pavlíček, T. (2005). Earthworms from Israel II. Remarks on the genus *Perelia* Easton, 1983 with descriptions of a new genus and two new species. *Acta Zoologica Academiae Scientiarum Hungaricae*, *51*(2), 75–96.
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772–772.
- de Sosa, I., Díaz Cosín, D. J., Csuzdi, C., Paoletti, M. G., & Marchán, D. F. (2019). Placing *Eophila tellinii* (Oligochaeta, Lumbricidae) in a molecular phylogenetic context advances the century-old controversy around the problematic genus. *European Journal of Soil Biology*, *94*, 103114.
- Domínguez, J., Aira, M., Breinholt, J. W., Stojanovic, M., James, S. W., & Pérez-Losada, M. (2015). Underground evolution: New roots for the old tree of lumbricid earthworms. *Molecular Phylogenetics and Evolution*, *83*, 7–19.
- Domínguez, J., Aira, M., Porto, G. P., Díaz Cosín, D. J., & Pérez-Losada, M. (2018). Multigene phylogeny reveals two new isolated and relic earthworm genera (Oligochaeta: Lumbricidae). *Zoological Journal of Linnean Society*, *20*, 1–17.
- Dugès, A. L. (1828). Recherche sur la circulation, la respiration, et la reproduction des Annélides sétigères abranches. *Annales des Sciences Naturelles Paris, 1ère série*, *15*, 284–337.

- Easton, E. G. (1983). A guide to the valid names of Lumbricidae (Oligochaeta). In J. E. Satchell (Ed.), *Earthworm ecology, from Darwin to vermiculture* (pp. 475–485). Chapman and Hall.
- Eisen, G. (1873). Om Skandnaviens Lumbricider. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 30(8), 43–56.
- Gates, G. E. (1969). On two American genera of the earthworm family Lumbricidae. *Journal of Natural History*, 9, 305–307.
- Gates, G. E. (1975). Contributions to a revision of the earthworm family Lumbricidae XII. Enterion mammale Savigny, 1826 and its position in the family. *Megadrilogica*, 2(1), 1–5.
- Gates, G. E. (1978). Contributions to a revision of the Lumbricidae. XXI. The earthworm genus Lumbricus in North America. *Megadrilogica*, 3(6), 81–116.
- Hamilton, C. A., Lemmon, A. R., Lemmon, E. M., & Bond, J. E. (2016). Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. *BMC Evolutionary Biology*, 16(1), 1–20.
- Hoffmeister, W. (1845). Die bis jetzt bekannten Arten aus der Familie der Regenwürmer. Als Grundlage zu einer Monographie dieser Familie, etc.
- James, S. W., & Davidson, S. K. (2012). Molecular phylogeny of earthworms (Annelida: Crassicitellata) based on 28S, 18S and 16S gene sequences. *Invertebrate Systematics*, 26(2), 213–229.
- Jamieson, B. G. M. (1988). On the phylogeny and higher classification of the Oligochaeta. *Cladistics*, 4, 367–410.
- Janko, K., Marshall, C., Musilová, Z., Van Houdt, J., Couloux, A., Cruaud, C., & Lecointre, G. (2011). Multilocus analyses of an Antarctic fish species flock (Teleostei, Notothenioidei, Trematominae): Phylogenetic approach and test of the early-radiation event. *Molecular Phylogenetics and Evolution*, 60(3), 305–316.
- Jiménez, S., Marchán, D. F., Novo, M., Trigo, D., Domínguez, J., & Díaz Cosín, D. J. (2021). Sorry atlanticus, you are not my type. Molecular assessment splits *Zophoscolex* (Lumbricidae, Crassicitellata) into French and Iberian genera. *Zoological Journal of the Linnean Society*.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.
- Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5), 727–744.
- Linnaeus, C. (1758). *Systema naturae* (Vol. 1, No. part 1, p. 532). Holmiae (Laurentii Salvii): Stockholm.
- Malm, A. W. (1877). Om Daggmaskar, Lumbricina. *Öfversigt af Sällskapet Hortikulturens Vänners i Göteborg Förhandlingar*, 1, 34–47.
- Marchán, D. F., Csuzdi, C., Decaëns, T., Szederjesi, T., Pizl, V., & Domínguez, J. (2021a). The disjunct distribution of relict earthworm genera clarifies the early historical biogeography of the Lumbricidae (Crassicitellata, Annelida). *Journal of Zoological Systematics and Evolutionary Research*.
- Marchán, D. F., Csuzdi, C., Decaëns, T., Szederjesi, T., Pizl, V., & Domínguez, J. (2021b). The disjunct distribution of relict earthworm genera clarifies the early historical biogeography of the Lumbricidae (Crassicitellata, Annelida). *Journal of Zoological Systematics and Evolutionary Research*.
- Marchán, D. F., Decaëns, T., Cosin, D. J. D., Hedde, M., Lapied, E., & Domínguez, J. (2020). French Mediterranean islands as a refuge of relic earthworm species: *Cataladrilus porquerollensis* sp. nov. and *Scherotheca porterosana* sp. nov. (Crassicitellata, Lumbricidae). *European Journal of Taxonomy*, (701).
- Marchán, D. F., Jiménez, S., Decaëns, T., & Domínguez, J. (2021c). Systematic revision of *Gatesona* (Crassicitellata, Lumbricidae), an endemic earthworm genus from the Massif Central (France). *PLoS ONE*, 16(9), e0255978.
- McMahan, M. L. (1976). *Biology of a limicolous megadrile (Annelida: Oligochaeta) from Louisiana*. Louisiana State University and Agricultural & Mechanical College.
- Mendes, J., Harris, D. J., Carranza, S., & Salvi, D. (2016). Evaluating the phylogenetic signal limit from mitogenomes, slow evolving nuclear genes, and the concatenation approach. New insights into the Lacertini radiation using fast evolving nuclear genes and species trees. *Molecular Phylogenetics and Evolution*, 100, 254–267.
- Michaelsen, W. (1890). Beschreibung der von Herrn Dr. Franz Stuhlmann im Mündungsgebiet des Sambesi gesammelten Terricolen. *Mitteilungen aus dem Naturhistorischen Museum in Hamburg*, 7, 21–50.
- Michaelsen, W. (1900). *Das Tierreich*. 10, *Oligochaeta*. Berlin: R. Friedländer und Sohn xxix, 575 pp.
- Michaelsen, W. (1919). Über die Beziehungen der Hirudineen zu den Oligochäten. *Jahrbuch der Hamburgischen Wissenschaftlichen Anstalten, Hamburg*, 36, 131–153.
- Mirarab, S., & Warnow, T. (2015). ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, 31, i44–i52.
- Moore, H. J. (1893). Preliminary account of a new genus of Oligochaeta. *Zoologischer Anzeiger*, 16, 333–334.
- Mršić, N. (1991). *Monograph on earthworms (Lumbricidae) of the Balkans I–II*. – Slovenska Akademija Znanosti in Umetnosti, Zazred za Naravoslovne Vede Opera. 31. Ljubljana, pp. 757.
- Novo, M., Almodovar, A., Fernandez, R., Trigo, D., & DÍAZ-COSÍN, D. J., & Giribet, G. (2012). Appearances can be deceptive: Different diversification patterns within a group of Mediterranean earthworms (Oligochaeta, Hormogastridae). *Molecular Ecology*, 21(15), 3776–3793.
- Novo, M., Fernández, R., Andrade, S. C., Marchán, D. F., Cunha, L., & Cosín, D. J. D. (2016). Phylogenomic analyses of a Mediterranean earthworm family (Annelida: Hormogastridae). *Molecular Phylogenetics and Evolution*, 94, 473–478.
- Omodeo, P. (1956). Contributo alla revisione dei Lumbricidae. – *Archivio Zoologico Italiano*, 41, 129–212.
- Omodeo, P., & Rota, E. (1989). Earthworms of Turkey. *Bolletino di Zoologia*, 56, 167–199.
- Örley, L. (1885). A palaearktikus övben élő Terrikolákknak revíziója és elterjedése. *Értekezések a Természettudományok Köréből, Magyar Tudományos Akadémia*, 15(18), 1–34.
- Perel, T. S. (1979). *Range and regularities in the distribution of earthworms of the USSR fauna*. – Nauka, Moscow pp. 272. (in Russian).
- Pérez-Losada, M., Breinholt, J. W., Aira, M., & Domínguez, J. (2015). An updated multilocus phylogeny of the Lumbricidae (Annelida: Clitellata: Oligochaeta) earthworms. *Journal of Phylogenetics and Evolutionary Biology*, 2015.
- Phillips, A. J., Dornburg, A., Zapfe, K. L., Anderson, F. E., James, S. W., Erséus, C., & Williams, B. W. (2019). Phylogenomic analysis of a putative missing link sparks reinterpretation of leech evolution. *Genome Biology and Evolution*, 11(11), 3082–3093.
- Pop, A. A., Cech, G., Wink, M., Csuzdi, C., & Pop, V. V. (2007). Application of 16S, 18S rDNA and COI sequences in the molecular systematics of the earthworm family Lumbricidae (Annelida, Oligochaeta). *European Journal of Soil Biology*, 43, S43–S52.
- Pop, V. (1941). Zur phylogenie und Systematik der Lumbriciden. – *Zoologische Jahrbücher Abteilung für Systematik. Ökologie Und Geographie Der Tiere*, 74, 487–522.
- Popovic, F. J., Stojanović, M. M., Domínguez, J., Sekulić, J. M., Trakić, T. B., & Marchán, D. F. (2022). Molecular analysis of five controversial Balkaic species of the genus *Cernosvitovia* Omodeo, 1956. *Zootaxa*, 5116(3), 351–372.
- Prum, R. O., Berv, J. S., Dornburg, A., Field, D. J., Townsend, J. P., Lemmon, E. C., & Lemmon, A. R. (2015). A fully resolved,

- comprehensive phylogeny of birds (Aves) using targeted next generation DNA sequencing. *Nature*, 526, 569–573.
- Qiu, J. P., & Bouché, M. B. (1998a). Révision des taxons supraspécifiques de Lumbricoidea. *Documents Pédozoologiques Et Intégrologiques*, 3, 179–216.
- Qiu, J. P., & Bouché, M. B. (1998b). Contribution à la taxonomie des Avelonini trib. nov. (Oligochaeta: Lumbricidae). Avelona gen. nov., Koinodrilus gen. nov. et Nicodrilus cuendeti sp. nov. *Documents Pédozoologiques Et Intégrologiques*, 4, 109–116.
- Qiu, J. P., & Bouché, M. B. (1998c). Le genre Allolobophora Eisen, 1874 (Oligochaeta: Lumbricidae), ses avatars et sa définition moderne. *Documents Pédozoologiques Et Intégrologiques*, 4, 86–97.
- Rafinesque, C. S. (1815). *Analyse de la nature, ou tableau de l'univers et des corps organisés* (p. 224). Palermo: J. Barravecchia.
- Rokyta, D. R., Lemmon, A. R., Margres, M. J., & Arnow, K. (2012). The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics*, 13, 312.
- Rosa, D. (1893). Revisione dei lumbricidi. *Memoire della Reale Accademia delle Scienze di Torino (Serie 2)*, 43, 399–477.
- Rosa, D. (1895). Nuovi lombrichi dell'Europa orientale. *Bolletino dei Musei di Zoologia ed Anatomia comparata della Reale Università di Torino*, 10(215), 1–8.
- Rosa, D. (1897). Nuovi lombrichi dell'Europa orientale (Seconda serie). *Bolletino dei Musei di Zoologia ed Anatomia Comparata della Reale Università di Torino*, 12(269), 1–5.
- Savigny, J. C. (1826). In G. Cuvier (Ed.), *Analyse des travaux de l'académie royale des sciences pendant l'année 1821, partie physique. Mémoires de l'Academie Royale des Sciences de l'Institut de France*, 5, 176–184.
- Schmarda, L. K. (1861). Neue Turbellarien, Rotatorien und Anneliden. In L. K. Schmarda & W. Engelmann (Ed.), *Neue wirbellose Thiere beobachtet und gesammelt auf einer Reise um die Erde, 1853 bis 1857* (Vol. 1, No. 2, pp. 1–164). Leipzig.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461–464.
- Smith, S. A., Brown, J. W., & Walker, J. F. (2018). So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLoS ONE*, 13(5), e0197433.
- Slater, G. J. (2013). Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution*, 4(8), 734–744.
- Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, 2688–2690.
- Stanley, E. L., Bauer, A. M., Jackman, T. R., Branch, W. R., & Mouton, P. L. F. N. (2011). Between a rock and a hard polytomy: Rapid radiation in the rupicolous girdled lizards (Squamata: Cordylidae). *Molecular Phylogenetics and Evolution*, 58(1), 53–70.
- Taheri, S., James, S., Roy, V., Decaëns, T., Williams, B. W., Anderson, F., & Dupont, L. (2018). Complex taxonomy of the 'brush tail' peregrine earthworm *Pontoscolex corethrurus*. *Molecular Phylogenetics and Evolution*, 124, 60–70.
- Zicsi, A. (1985). Über die Gattung *Helodrilus* Hoffmeister, 1845 und *Proctodrilus* gen. n. (Oligochaeta:Lumbricidae). *Acta Zoologica Academiae Scientiarum Hungaricae*, 31, 275–289.
- Zicsi, A., & Šapkarev, J. (1982). Eine neue Cernovitovia-Art aus Jugoslawien (Oligochaeta: Lumbricidae). *Acta Zoologica Academiae Scientiarum Hungaricae*, 28, 181–182.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.