

The complete mitochondrial DNA sequences of two sibling species of lumbricid earthworms, *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* (Bouché, 1972) (Annelida, Crassicitellata): comparison of mitogenomes and phylogenetic positioning

Csaba Csuzdi^{1*}, Jachoon Koo^{2*}, Yong Hong³

1 Department of Zoology, Eszterházy Károly Catholic University, Eger, Hungary **2** Division of Science Education and Institute of Fusion Science, College of Education, Jeonbuk National University, Jeonju 54896, Republic of Korea **3** Department of Agricultural Biology, College of Agriculture & Life Science, Jeonbuk National University, Jeonju 54896, Republic of Korea

Corresponding author: Yong Hong (yonghong@jbnu.ac.kr)

Academic editor: Sergei Subbotin | Received 7 January 2022 | Accepted 3 April 2022 | Published 29 April 2022

<http://zoobank.org/C43A88B7-0CF9-46A1-8F5E-51B35E43C488>

Citation: Csuzdi C, Koo J, Hong Y (2022) The complete mitochondrial DNA sequences of two sibling species of lumbricid earthworms, *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* (Bouché, 1972) (Annelida, Crassicitellata): comparison of mitogenomes and phylogenetic positioning. ZooKeys 1097: 167–181. <https://doi.org/10.3897/zookeys.1097.80216>

Abstract

Composting earthworms of the genus *Eisenia* play an important role in soil ecosystems. However, taxonomic classification of this genus, especially the sibling species *Eisenia fetida* and *Eisenia andrei*, is complicated because of their morphological similarity. In this study, we assessed the utility of the complete mitochondrial genome (mitogenome) for identification and differentiation of the two species. The complete mitogenomes of *E. andrei* and *E. fetida* were 15,714 and 16,560 bp, respectively. They contained 37 genes, comprising 13 protein-coding genes (PCGs), two rRNA genes, 22 tRNA genes, and a putative non-coding region, as observed in other earthworms. Sequence comparisons based on the complete nucleotide sequences excluding the non-coding region showed 85.8% similarity, whereas the predicted amino acid sequences of the 13 PCGs were 92.7% similar between the two species. In particular, distinct features were found in the non-coding regions of the mitogenomes. They include a control region associated with putative mitogenome replication and an extended sequence. The extended sequence showed sig-

* These authors contributed equally to this work.

nificant differences between the two species and other known earthworm species, suggesting its potential as a feasible molecular marker for species identification. Phylogenetic analysis of the 36 mitogenomes of earthworm species corroborated the monophyly of the genus *Eisenia* and the taxonomic distinctness of the sibling species pair, *E. fetida* and *E. andrei*.

Keywords

Compost worms, mitogenome, Oligochaeta, phylogeny, sibling species

Introduction

The earthworm species *Eisenia fetida* was described as *Enterion fetidum* by Savigny (1826). Eisen (1873) relegated this species to his newly described genus *Allolobophora* Eisen, 1873 and remarked that it is easily recognized by its peculiar color pattern consisting of reddish-brown bands separated by yellowish intersegments. Later, Malm (1877) selected *Enterion fetidum* as the type species of the genus *Eisenia* Malm, 1877. For a long time, the characteristic striated pattern was a primary identifiable characteristic of the species until Avel (1937) recognized that the classical *Eisenia fetida* existed in two morphological variants: a typical striped form and an evenly pigmented form that might represent a separate species (Avel 1937).

André (1963) carried out breeding experiments with earthworms and recognized that reproductive isolation exists between the striped and evenly colored forms of *E. fetida*, and that the crossbred offspring are sterile. Consequently, he described the uniformly pigmented form as *Eisenia fetida* var. *unicolor*. Variety names proposed after 1961 were considered intrasubspecific and invalid; therefore, Bouché (1972) proposed a new name for var. *unicolor*, *Eisenia fetida andrei* Bouché, 1972. Since then, various authors have treated the subspecies *E. f. andrei* differently. Reynolds (1977), Sims (1983), Easton (1983), Csuzdi and Zicsi (2003), and Blakemore (2008, 2013) regarded it as a color morph and synonym of *Eisenia fetida* (Savigny, 1826). Others, such as Sims and Gerard (1985), Qiu and Bouché (1998), Lehmitz et al. (2014), and Martin et al. (2016), considered *E. fetida* and *E. andrei* to be two distinct valid species.

Eisenia fetida is an important composting worm and ecotoxicological test organism (Domínguez et al. 2005; Römbke et al. 2016). Therefore, intensive studies have been carried out since the early 1980s to determine whether the two types of *E. fetida* (striped and unicolor) represent two morphological variants or two separate species. The first clear indication that *E. fetida* and *E. andrei* might represent two separate species was presented by Jaenike (1982) who used an electrophoretic survey to demonstrate complete reproductive isolation between the two species. Later, Reinecke and Viljoen (1991) and Domínguez et al. (2005), using crossbreed experiments, reported complete reproductive isolation between the two species (no viable cocoons were observed in interspecific crosses) and noted that the two species differed in their life histories. Furthermore, *E. andrei* exhibited higher reproduction rates.

Recently, Römbke et al. (2016) carried out a detailed barcoding study of the *Eisenia fetida* / *E. andrei* complex using samples from 28 laboratories in 15 countries. The two species formed two distinct clades on the neighbor-joining tree, and the *E. fetida* clade consisted of two subclades, *fetida* 1 and *fetida* 2. The mean uncorrected p-distances were 14.2% between *fetida*1 and *andrei*, 14.3% between *fetida*2 and *andrei*, and 11.2% between the two *fetida* subclades; these values exceed the species-level threshold suggested by Chang and James (2011). Therefore, Römbke et al. (2016) concluded that the complex consists of three taxa: *E. andrei* and two cryptic taxa, *E. fetida* 1 and *E. fetida* 2. Moreover, they found that *E. andrei* was always correctly identified from its morphology, whereas *E. fetida* was often misidentified as *E. andrei* (Römbke et al. 2016).

It is worth mentioning that the native range of *E. fetida* and *E. andrei* is unknown. All of the above-mentioned studies were based on laboratory stocks or specimens collected from compost or manure heaps. Perel (1998) hypothesized that the native range of *E. fetida* is somewhere in the forest-steppe zone of Central Asia; therefore, Latif et al. (2017) barcoded 62 new specimens of this complex collected from different anthropogenic and natural habitats in Iran. Surprisingly, all Iranian material appeared in the *E. andrei* clade, irrespective of striped or uniform pigmentation. Moreover, the *E. andrei* clade showed high genetic structuring in contrast to the almost uniform genetic composition found by Römbke et al. (2016). Automatic barcode gap discovery (ABGD) analysis identified two species corresponding to *Eisenia andrei* and *Eisenia fetida* with high genetic structuring inside both species, but neither of the subclades reached the unambiguous species threshold [15% K2P distance according to Chang and James (2011)].

Comparison of mitogenomes may reveal important genome-level characteristics, helping us understand genome structure, gene order, phylogenetic relationships, and evolutionary lineages. The earthworm mitogenome is a circular, double-stranded, covalently closed DNA molecule containing 13 protein-coding genes (PCGs), two ribosomal RNA genes (rRNAs), 22 transfer RNA genes (tRNAs), and one non-coding region (Zhang et al. 2016). Although Lumbricidae is the most important earthworm family in the Northern Hemisphere temperate zone and contains many widespread and invasive cosmopolitan species, only a few complete or nearly complete mitogenomes are available for this family (Boore and Brown 1995; Shekhovtsov and Peltek 2019; Zhang et al. 2019; Shekhovtsov et al. 2020). In the present study, we sequenced the complete mitochondrial genome of the sibling species *E. fetida* and *E. andrei* to clarify its taxonomic position and to gain a better understanding of the mitogenomes of Lumbricidae.

Material and methods

Sample preparation and DNA extraction

Adult *E. andrei* were collected from a farm in Sangseo-myeon, Buan-gun, Jeollabuk-do, Korea (33°41'23.80"N, 126°38'33.67"E; 40 m a.s.l.) on March 26, 2021. *Eisenia fetida* adults were collected near a house at Seolcheon-myeon, Muju-gun, Jeollabuk-do, Korea

(33°58'00.61"N, 127°47'47.88"E; 408 m a.s.l.) on April 2, 2021, and preserved in 99% ethanol until DNA extraction. A voucher specimen of each species was deposited at Jeonbuk National University, Jeonju City, Korea, under accession numbers JBNU0011 and JBNU0012. Total genomic DNA was prepared from a small portion of body segments of a single adult earthworm using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). The remaining tissue was stored at -20 °C in 90% ethanol to preserve the specimens.

TruSeq DNA Library construction

The sequencing library was prepared by random fragmentation of genomic DNA, followed by 5' and 3' adapter ligations. Briefly, 100 ng genomic DNA was fragmented using adaptive focused acoustic (AFA) technology (Covaris Inc., Woburn, MA, USA). The fragmented DNA was end-repaired and ligated to TruSeq indexing adapters using the Illumina TruSeq DNA Nano Library Prep Kit according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). The resulting libraries were quantified through a qPCR-based assay using the KAPA Library Quantification Kit for Illumina Sequencing platforms according to the manufacturer's instructions (Kapa Biosystems, Woburn, MA, USA). The libraries were qualified using an Agilent Technologies 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA).

DNA sequencing and assembly

Paired-end (2 × 150 bp) sequencing was performed using an Illumina HiSeq-X platform (Illumina Inc., USA) at Macrogen Inc. (Seoul, Korea). For each species, > 39 million reads (5.1–5.9 Gb) were generated. To reduce bias in the analysis, adapter trimming and quality filtering were performed using Trimmomatic version 0.36 (Bolger et al. 2014). After filtering, the number of total reads of *E. andrei* and *E. fetida* was > 29 million (4.4 Gb) and > 34 million (5.2 Gb), respectively. De novo assembly of raw sequencing reads was performed using various *k*-mer lengths in SPAdes version 3.13.0 (Bankevich et al. 2012). Mitochondrial contigs were assembled into a single contig using BlastN alignment (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) against the *Lumbricus terrestris* Linnaeus, 1758 mitogenome (GenBank accession number, NC_001673) as the reference sequence. The assembled mitochondrial sequences for *E. andrei* were connected to a single circular molecule, whereas the conformation of the contig for *E. fetida* was unclear because the 12 bp TA-repeat sequence overlapped at both ends. This region corresponds to heteroplasmic tandem repeats in the mitochondrial control region (Liu et al. 2020). To close the circular genome, pairs of PCR primers (5'-ACCACCAGAGTTCTCGTTCG-3' and 5'-GCCAATATCGGCCCAAACC-3') were designed to amplify the control region. The reaction was performed in an nTaq-tenuo (Enzynomics Inc., Seoul, Korea) with the following program: 95 °C for 3 min; 35 cycles of 95 °C for 20 s, 55 °C for 30 s, and 72 °C for 1 min; and a final extension of 5 min at 72 °C. The amplicons were directly sequenced using Sanger sequencing (Macrogen Inc., Seoul, Korea) to determine the complete mitogenome of *E. fetida*.

Mitogenome annotation

The annotation and visualization of mitochondrial genomes were performed using the online MITOS software (Donath et al. 2019), and manual curation was performed using BLAST searches in the NCBI database for various earthworm mitochondrial genomes deposited in NCBI (Table 1). A comparative map of mitochondrial genomes was created using Geneious Prime 2021 software (<https://www.geneious.com>). The *cox1* sequence was used as an anchor for linearized maps of the mitochondrial genomes. The annotated complete genome sequences were registered in GenBank under accession numbers OK513069 for *E. andrei* and OK513070 for *E. fetida*. The associated biosample numbers were SAMN26185682 for *E. andrei* and SAMN26185683 for *E. fetida*. All sequencing datasets, including SRA, are available in the NCBI BioProject database under the accession number PRJNA769829.

Phylogenetic analyses

To clarify the phylogenetic position of the two species, the available complete or near-complete mitogenomes were obtained from GenBank, comprising 24 species of Megascolecidae, 14 species of Lumbricidae, and one species of Rhinodrilidae. *Drawida japonica* (Michaelsen, 1892) from the exquisiclitellate family Moniligastridae was used as the outgroup.

Two sets of sequence matrices were composed: one containing the PCGs, 12S, and 16S RNA genes, and the other consisting only of PCGs. Sequences were aligned with MAFFT ver. 7 (Katoh and Standley 2013) using the G-INS-i option and concatenated in MegaX (Kumar et al. 2018); the resulting matrices were 13,505 and 11,241 bp, respectively. The protein-coding alignment was translated into amino acid sequences and aligned in MAFFT ver. 7 using the G-INS-i option; the resulting matrix was with 3714 amino acid positions.

The best-fitting evolutionary model for each partition (PCG, 16S, 12S) was selected using ModelFinder (Kalyaanamoorthy et al. 2017) implemented in the IQTree web server (<http://iqtree.cibiv.univie.ac.at/>) by applying the Akaike information criterion (AIC; Akaike 1973) and Bayesian information criterion (BIC; Schwarz 1978). GTR + I + Γ was selected as the best-fitting evolutionary model for PCGs and 12S RNA, TIM2 I + Γ was selected for 16S RNA, and MtMAM I + Γ for the amino acid sequences.

Bayesian inference of the phylogeny was estimated with MrBayes v.3.2.6 (Ronquist et al. 2012) as implemented in CIPRES Science Gateway V. 3.3. (Miller et al. 2010). The analysis was performed with default parameters, and each of the two independent runs was set to 10 million generations and sampling every 1000th generation (10,000 trees). Twenty percent of the trees were discarded as burn-in, and the remaining trees were combined and summarized in a 50% majority-rule consensus tree. As the TIM2 model was not implemented in MrBayes, the closest complex model GTR + I + Γ was used instead. Maximum likelihood phylogenetic inference was performed using the IQTree web server with default options (Nguyen et al. 2015 <http://iqtree.cibiv.univie.ac.at/>).

Table 1. List of Megadrili mitogenomes used in this study.

Species	Genbank No.	Total length (bp)	*Non-coding region (bp)	Topology
<i>Amyntas aspergillus</i>	KJ830749	15,115	565	Circular
<i>Amyntas carnosus</i>	KT429008	15,160	601	Circular
<i>Amyntas corticis</i>	KM199290	15,126	573	Circular
<i>Amyntas cucullatus</i>	KT429012	15,122	569	Circular
<i>Amyntas gracilis</i>	KP688582	15,161	582	Circular
<i>Amyntas hupeiensis</i>	KT429009	15,069	477	Circular
<i>Amyntas instabilis</i>	KT429007	15,159	577	Circular
<i>Amyntas jiriensis</i>	KT783537	15,151	618	Circular
<i>Amyntas longisiphonus</i>	KM199289	15,176	491	Circular
<i>Amyntas moniliatus</i>	KT429020	15,133	562	Circular
<i>Amyntas pectiniferus</i>	KT429018	15,188	618	Circular
<i>Amyntas redactus</i>	KT429010	15,131	572	Circular
<i>Amyntas robustus</i>	KT429019	15,013	432	Circular
<i>Amyntas rongshuiensis</i>	KT429014	15,086	546	Circular
<i>Amyntas spatiosus</i>	KT429013	15,152	595	Circular
<i>Amyntas triastriatus</i>	KT429016	15,160	582	Circular
<i>Amyntas yunoshimensis</i>	LC573969	15,109	581	Circular
<i>Metaphire californica</i>	KP688581	15,147	567	Circular
<i>Metaphire guillelmi</i>	KT429017	15,174	594	Circular
<i>Metaphire hilgendorfi</i>	LC573968	15,186	649	Circular
<i>Metaphire vulgaris</i>	KJ137279	15,061	484	Circular
<i>Duplodicodrilus schmardae</i>	KT429015	15,156	595	Circular
<i>Perionyx excavatus</i>	EF494507	15,083	504	Circular
<i>Tonoscolex birmanicus</i>	KF425518	15,170	595	Circular
<i>Aporrectodea rosea</i>	MK573632	15,086	512	Circular
<i>Lumbricus rubellus</i>	MN102127	15,464	433	Circular
<i>Lumbricus terrestris</i>	U24570	14,998	384	Circular
** <i>Eisenia balatonica</i>	MK642872	14,589	-	Linear
** <i>Eisenia nana</i>	MK618511	14,599	-	Linear
** <i>Eisenia nordenskioldi</i>	MK618509	14,572	-	Linear
** <i>Eisenia nordenskioldi</i>	MK618510	14,592	-	Linear
** <i>Eisenia nordenskioldi</i>	MK618513	14,567	-	Linear
** <i>Eisenia nordenskioldi</i>	MK642867	14,576	-	Linear
** <i>Eisenia nordenskioldi</i>	MK642868	14,556	-	Linear
** <i>Eisenia nordenskioldi pallida</i>	MK618512	14,567	-	Linear
** <i>Eisenia nordenskioldi pallida</i>	MK642869	14,553	-	Linear
** <i>Eisenia spelaea</i>	MK642870	14,738	-	Linear
** <i>Eisenia tracta</i>	MK642871	14,589	-	Linear
<i>Eisenia andrei</i>	OK513069	15,714	1151	Circular
<i>Eisenia fetida</i>	OK513070	16,560	1988	Circular
<i>Drawida japonica</i>	KM199288	14,648	3	Circular
<i>Pontoscolex corethrurus</i>	KT988053	14,835	318	Circular

*Putative non-coding region between *trnR* and *trnH*.

** Incomplete mitochondrial genome sequence lacking the entire non-coding region and *trnR*.

Results

The complete mitochondrial genomes of *Eisenia fetida* and *Eisenia andrei* consisted of 16,560 and 15,714 base pairs, respectively. The setup of the mitogenomes of both species followed the typical Bauplan of the earthworm mitogenome assembly, consisting of 13 PCGs, 22 transfer RNAs, two ribosomal RNA genes, and a control region (Fig. 1; Table 2).

All genes were encoded on the heavy DNA strand, and both genomes showed biased base composition, with 63.5% AT and 36.4% GC content in *E. fetida* and 62.8% and 37.2% in *E. andrei*.

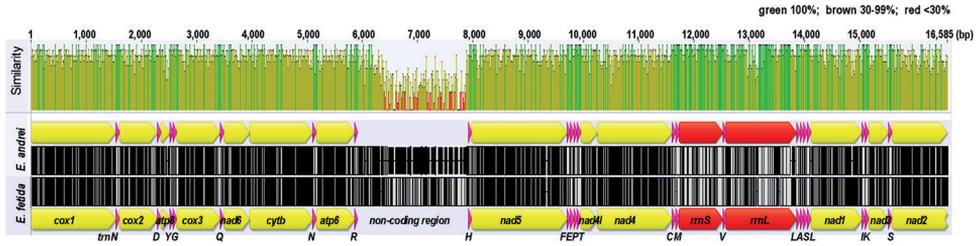


Figure 1. Comparison of mitogenomes of *Eisenia andrei* and *E. fetida*. The map is based on sequence similarity and was constructed using Geneious Prime 2021 software. Sequence similarity is represented by green (100%), brown (30–99%), and red (<30%). *cox1* was used as an anchor to linearized genomes. Organization of mitochondrial genes is shown in Table 2. Non-coding region is defined as the region between *trnR* and *trnH*.

Table 2. Comparative analysis of gene organization of *Eisenia andrei* and *E. fetida* mitogenomes (bp = base pairs).

Gene	Strand	<i>E. fetida</i>		<i>E. andrei</i>		Similarity
		Size (bp)	start/stop codon	Size (bp)	start/stop codon	
<i>cox1</i>	+	1540	ATG/T	1540	ATG/T	86%
<i>trnN</i>	+	61		61		98%
<i>cox2</i>	+	687	ATG/TAG	687	ATG/TAA	86%
<i>trnD</i>	+	61		61		89%
<i>atp8</i>	+	163	ATG/T	160	ATG/T	78%
<i>trnY</i>	+	63		63		95%
<i>trnG</i>	+	63		64		92%
<i>cox3</i>	+	778	ATG/T	778	ATG/T	86%
<i>trnQ</i>	+	69		69		91%
<i>nad6</i>	+	469	ATG/T	469	ATG/T	85%
<i>cytb</i>	+	1140	ATG/TAA	1140	ATG/TAA	85%
<i>trnW</i>	+	62		63		90%
<i>atp6</i>	+	696	ATG/TAA	696	ATG/TAA	82%
<i>trnR</i>	+	61		63		93%
*NC	+	1988		1151		60%
<i>trnH</i>	+	62		62		90%
<i>nad5</i>	+	1722	ATG/TAA	1722	ATG/TAA	83%
<i>trnF</i>	+	62		63		92%
<i>trnE</i>	+	63		63		95%
<i>trnP</i>	+	64		64		94%
<i>trnT</i>	+	65		63	-	97%
<i>nad4L</i>	+	297	ATG/TAA	297	ATG/TAA	88%
<i>nad4</i>	+	1359	ATG/TAG	1359	ATG/TAG	83%
<i>trnC</i>	+	65		65		97%
<i>trnM</i>	+	63		63		100%
<i>rrnS</i>	+	794		794		94%
<i>trnV</i>	+	64		63		94%
<i>rrnL</i>	+	1282		1278		89%
<i>trnL</i>	+	62		63		94%
<i>trnA</i>	+	62		62		94%
<i>trnS</i>	+	67		67		94%
<i>trnL</i>	+	64		62		95%
<i>nad1</i>	+	919	ATG/T	919	ATG/T	85%
<i>trnI</i>	+	64		64		97%
<i>trnK</i>	+	65		65		97%
<i>nad3</i>	+	354	ATG/TAG	354	ATG/TAG	82%
<i>trnS</i>	+	64		64		97%
<i>nad2</i>	+	1003	ATG/T	1003	ATG/T	81%

* Non-coding regions between *trnR* and *trnH*.

Table 3. Comparison of deduced amino acid sequences of 13 protein-coding genes between *Eisenia andrei* and *E. fetida*.

Protein	<i>Eisenia fetida</i>	<i>Eisenia andrei</i>	Similarity (%)
COX1	513 aa	513 aa	99.4
COX2	228 aa	228 aa	95.2
ATP8	54 aa	53 aa	79.6
COX3	259 aa	259 aa	97.7
NAD6	156 aa	156 aa	93.6
CYTB	379 aa	379 aa	96.0
ATP6	231 aa	231 aa	93.1
NAD5	567 aa	573 aa	90.8
NAD4L	98 aa	98 aa	92.9
NAD4	452 aa	452 aa	92.0
NAD1	306 aa	306 aa	92.6
NAD3	117 aa	117 aa	92.3
NAD2	334 aa	334 aa	89.1

The overall mitogenome sequence similarity between the two species was 80.8%, and it increased to 85.8% when the control region was excluded. The 13 PCGs were 78%–86% similar (Table 2). Among the PCGs, *nad4l* showed the highest similarity (88%) and *atp8* the lowest (78%). The average similarity of the 13 PCGs between the two species was 84%.

However, the deduced amino acid sequences of the 13 PCGs showed, on average, 92.7% similarity between the species; COX1 was the most similar (99.4%) and ATP8 the most dissimilar (79.6%) (Table 3). Sequence variation between the two species was lower at the amino acid level than at the DNA level. In particular, *cox1* showed 86% similarity at the DNA level but more than 99% similarity at the amino acid level.

Phylogenetic reconstruction of the available Lumbricidae complete or nearly complete mitogenomes using the 13 PCGs and the 12S and 16S RNA genes highly supported the Lumbricidae family (1 posterior probability and 100% bootstrap support). In addition, the genus *Eisenia* was resolved monophyletic, and the close relationship of the *E. fetida/andrei* species pairs was confirmed (Fig. 2). Interestingly, the included *Eisenia* sequences formed two well-supported subclades: one consisting of the European *E. spelaea* (Rosa, 1901) and the *E. fetida/andrei* species pair, and the other comprising the Asian taxa of the *E. nordenskioldi* (Eisen, 1879) species complex (including *E. tracta* Perel, 1985 and *E. nana* Perel, 1985), and the Asian specimens of the European *E. balatonica* (Pop, 1943). A nearly identical tree topology was obtained using the translated amino acid sequences. The only notable difference was in the swapped position of *Perionyx excavatus* Perrier, 1872 and *Tonoscolex birmanicus* within the Megascolecidae clade (Fig. 3).

Discussion

The mitogenomes of *E. fetida* and *E. andrei* show the same setup as other lumbricid mitogenomes (Boore and Brown 1995; Shekhovtsov and Peltek 2019; Zhang et al. 2019; Shekhovtsov et al. 2020). The nucleotide composition of the mitogenomes was also

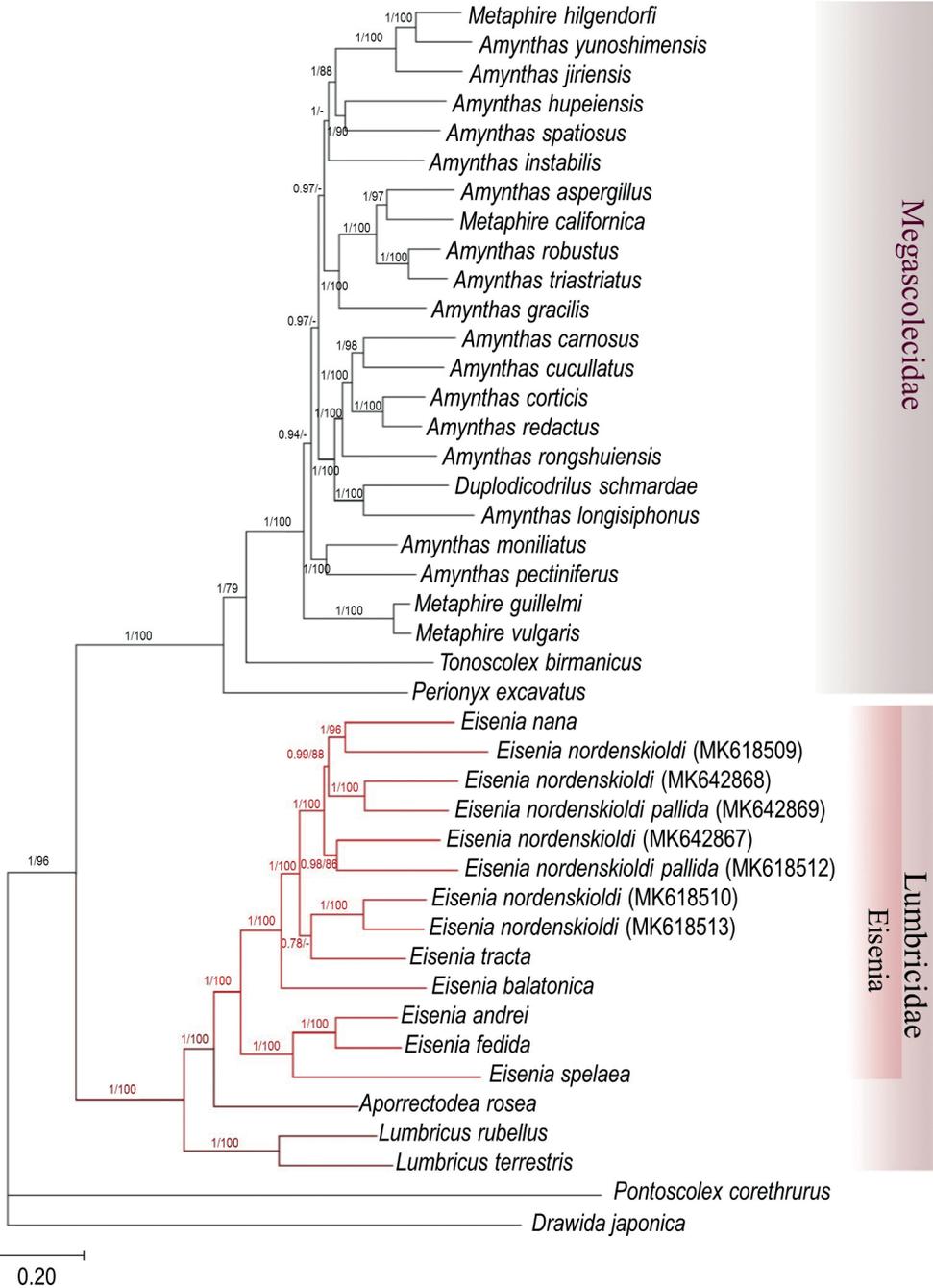


Figure 2. Phylogenetic analysis of 42 Megadrili species, including *E. andrei* and *E. fetida*, based on nucleotide sequences of 13 protein-coding genes and the 12S and 16S RNA genes. The numbers above branches present Bayesian posterior probabilities/maximum likelihood bootstrap values (values under 0.75 and 75% are not shown).

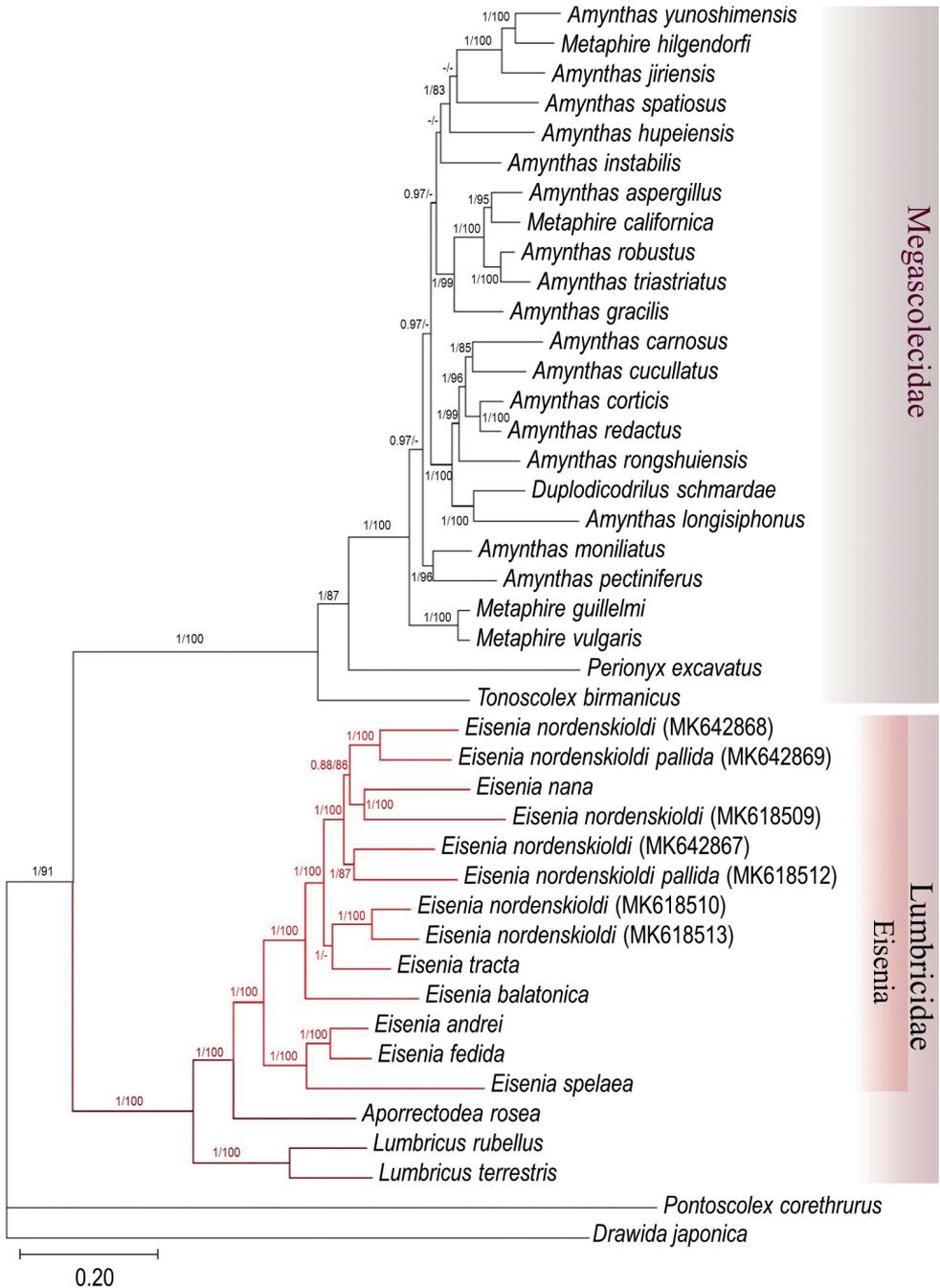


Figure 3. Phylogenetic analysis of 42 Megadrili species, including *E. andrei* and *E. fedida*, based on translated amino acid of 13 protein-coding genes. The numbers above branches are Bayesian posterior probabilities/maximum likelihood bootstrap values (values under 0.75 and 75% are not shown).

similar to that of other Lumbricidae species: the AT content of *E. fetida* and *E. andrei* (63.5% and 62.8%, respectively) was comparable to that in Lumbricidae species (59.88–65.69%), including *Eisenia nordenskioldi*, *E. balatonica*, *E. tracta*, *E. spelaea*, *Lumbricus terrestris* Linnaeus, 1758, and *Aporrectodea rosea* (Savigy, 1826) (Shekhovtsov et al. 2020). Zhang et al. (2016) reported higher AT contents in other earthworm families; for example, Megascolecidae has an AT content of 62.6–67.6%, and the Moniligastridae (*Drawida japonica*) genome has an AT content as high as 69.7%. However, the mitogenomes of *E. fetida* (16,560 bp) and *E. andrei* (15,714 bp) were larger than those of other lumbricid species, such as *L. terrestris* (14,998 bp), *L. rubellus* Hoffmeister, 1845 (15,464 bp), and *Ap. rosea* (15,089 bp). These size differences are primarily due to the extreme length variation of the non-coding region (Shekhovtsov et al. 2020). The length of the non-coding region was 1988 bp in *E. fetida* and 1152 bp in *E. andrei* and significantly longer than those of known mitogenomes of other earthworm species [from 318 bp in *Pontoscolex corethrurus* (Müller, 1857) to 649 bp in *Metaphire hilgendorfi* (Michaelsen, 1892); Table 1]. In addition, slight differences were observed in the coding regions. The *atp8* gene consists of 54 amino acids in *fetida* and 53 amino acids in *andrei*, whereas *nad5* comprises 567 amino acids in *fetida* and 573 amino acids in *andrei*.

The family Lumbricidae is well-known for its notoriously polyphyletic genera (Domínguez et al. 2015). Unfortunately, only 16 complete or nearly complete Lumbricidae mitogenomes are available in GenBank (including our two new sequences), which prevents us from reaching a comprehensive conclusion on Lumbricidae phylogeny. However, our phylogenetic reconstructions using the available complete or nearly complete mitogenomes corroborated the monophyly of the family Lumbricidae and *Eisenia* (Domínguez et al. 2015; Shekhovtsov et al. 2020), the genus with the most mitogenome sequences (13 sequences) reported, including the type species *Eisenia fetida*. It is interesting to note that the *E. fetida* and *E. andrei* clade along with the Central European *E. spelaea* is distant from the Asian *E. nordenskioldi* species complex, *E. tracta*, and *E. nana*.

Perel (1998) hypothesized that the native range of *E. fetida* is somewhere in the forest-steppe zone of Central Asia, and that the species originally occurred under the bark of fallen logs. In addition, Latif et al. (2017) found surprisingly high morphological and genetic variability of *E. andrei* in northwestern Iran, which demonstrates that the native range of both species is somewhere in western Central Asia. This could explain their closer affinity to the Central European *E. spelaea* than to the Siberian–Far Eastern *E. nordenskioldi* species group.

Eisenia fetida and *E. andrei* are sister taxa in both tree topologies (Figs 2, 3), and the branch length between *E. fetida* and *E. andrei* is similar to those of other species on the trees. This supports their distinct species status. However, considering the genetic p-distances of the studied mitogenomes (Table 4), the *E. fetida*–*E. andrei* species pair showed the second smallest genetic distance (14.1%), whereas the p-distance between *L. rubellus* and *L. terrestris* was 18.9% or even larger between the two closely related species *E. nana* and *E. tracta* (19.2%).

References

- Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrow BN, Csáki F (Eds) Proceedings of the 2nd International Symposium on Information. Akademiai Kiado, Budapest, 267–281.
- André F (1963) Contribution à l'analyse expérimentale de la reproduction des lombriciens. Bulletin Biologique de la France et de la Belgique 97: 3–101.
- Avel M (1937) Titres et travaux scientifiques. Delmas, Bordeaux, 76 pp.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) Spades: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Blakemore RJ (2008) Cosmopolitan earthworms – an eco-taxonomic guide to the species. 3rd edn. VermEcology, Yokohama, 757 pp.
- Blakemore RJ (2013) Earthworms newly from Mongolia (Oligochaeta, Lumbricidae, *Eisenia*). ZooKeys 285: 1–21. <https://doi.org/10.3897/zookeys.285.4502>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) 30(15): 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boore JL, Brown WM (1995) Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. Genetics 138: 423–433. <https://doi.org/10.1093/genetics/138.2.423>
- Bouché MB (1972) Lombriciens de France Écologie et Systématique. Institut National de la Recherche Agronomique. Annales de Zoologie Ecologie Animale. Numéro hors-série, Paris, 671 pp.
- Chang CH, James SW (2011) A critique of earthworm molecular phylogenetics. Pedobiologia 54S: 3–9. <https://doi.org/10.1016/j.pedobi.2011.07.015>
- Csuzdi Cs, Zicsi A (2003) Earthworms of Hungary (Annelida: Oligochaeta, Lumbricidae). Hungarian Natural History Museum, Budapest, 273 pp.
- Domínguez J, Velando A, Ferreira A (2005) Are *Eisenia fetida* and *Eisenia andrei* (Oligochaeta, Lumbricidae) different biological species? Pedobiologia 49(1): 81–87. <https://doi.org/10.1016/j.pedobi.2004.08.005>
- Domínguez J, Aira M, Breinholt JW, Stojanović M, James SW, Pérez-Losada M (2015) Underground evolution: New roots for the old tree of lumbricid earthworms. Molecular Phylogenetics and Evolution 83: 7–19. <https://doi.org/10.1016/j.ympev.2014.10.024>
- Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M (2019) Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Research 47(20): 10543–10552. <https://doi.org/10.1093/nar/gkz833>
- Easton EG (1983) A guide to the valid names of Lumbricidae (Oligochaeta). In: Satchell JE (Ed.) Earthworm Ecology - From Darwin to Vermiculture. Chapman & Hall, London, 475–487. https://doi.org/10.1007/978-94-009-5965-1_41

- Eisen G (1873) On Skandinavien Lumbricider. Ofversigt af Kongliga Vetenskaps-Akademiens Förhandlingar 30: 43–56. <https://archive.org/details/biostor-135527>
- Eisen G (1879) On the Oligochaeta collected during the Swedish Expeditions to the Arctic regions in the years 1870, 1875 and 1876. Öfversigt af Kongliga Svenska Vetenskaps-Akademiens Förhandlingar 15(7): 1–49.
- Jaenike J (1982) “*Eisenia foetida*” is two biological species. Megadrilogica 4: 6–8.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods 14(6): 587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution 35(6): 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Latif R, Malek M, Csuzdi Cs (2017) When morphology and DNA are discordant: Integrated taxonomic studies on the *Eisenia fetidalandrei* complex from different parts of Iran (Annelida, Clitellata: Megadrili). European Journal of Soil Biology 81: 55–63. <https://doi.org/10.1016/j.ejsobi.2017.06.007>
- Lehmitz R, Römbke J, Jänsch S, Krück S, Beylich A, Graefe U (2014) Checklist of earthworms (Oligochaeta: Lumbricidae) from Germany. Zootaxa 3866(2): 221–245. <https://doi.org/10.11646/zootaxa.3866.2.3>
- Liu H, Xu N, Zhang Q, Wang G, Xu H, Ruan H (2020) Characterization of the complete mitochondrial genome of *Drawida gisti* (Metagnophora, Moniligastridae) and comparison with other Metagnophora species. Genomics 112(5): 3056–3064. <https://doi.org/10.1016/j.ygeno.2020.05.020>
- Malm AW (1877) Om daggmaskar, Lumbricina. Öfversigt af Sällskapet Hortikulturens Vänners i Göteborg Förhandlingar 1: 34–47.
- Martin P, Martinez-Ansemil E, Pinder A, Timm T, Wetzel MJ (2016) World checklist of freshwater Oligochaeta species. [Available online at] <http://fada.biodiversity.be/group/show/12> [accessed 09 October 2021]
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE), 1–8. <https://doi:10.1109/GCE.2010.5676129>
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32(1): 268–274. <https://doi.org/10.1093/molbev/msu300>
- Perel TS (1985) Specific Features of the Earthworm Fauna (Oligochaeta, Lumbricidae) in Altai Refugia of Nemoral Vegetation. Doklady Akademii Nauk SSSR 283(3): 752–756. [In Russian]
- Perel TS (1998) The earthworms of the fauna of Russia: Cadaster and key. Nauka, Moscow, 97 pp. [in Russian]

- Pop V (1943) Hazai és külföldi lumbricidák a Magyar Nemzeti Múzeumban. *Annales Historico-Naturales Musei Nationalis Hungarici* 34: 12–24.
- Qiu JP, Bouché MB (1998) Revision des taxons supraspécifiques de Lumbricoidea. *Documents pedozoologiques et intergrologiques* 3: 179–216.
- Reinecke AJ, Viljoen SA (1991) A comparison of the biology of *Eisenia fetida* and *Eisenia andrei* (Oligochaeta). *Biology and Fertility of Soils* 11(4): 295–300. <https://doi.org/10.1007/BF00335851>
- Reynolds JW (1977) The earthworms (Lumbricidae and Sparganophilidae) of Ontario. Royal Ontario Museum, Toronto, 141 pp.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3): 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Römbke J, Aira M, Backeljau T, Breugelmans K, Dominguez J, Funke E, Graf N, Hajibabaei M, Perez-Losada M, Porto PG, Schmelz RM, Vierna J, Vizcaino A, Pfenninger M (2016) DNA barcoding of earthworms (*Eisenia fetida/andrei* complex) from 28 ecotoxicological test laboratories. *Applied Soil Ecology* 104: 3–11. <https://doi.org/10.1016/j.apsoil.2015.02.010>
- Rosa D (1901) Un lombrico cavernicolo. *Atti della Società dei Naturalisti e Matematici di Modena* 4: 36–39. <http://www.morebooks.unimore.it/site/home/la-produzione-scientifica/documento610066458.html>
- Savigny JC (1826) Analyse d'un memoire sur les Lumbricus par Cuvier. *Mémoires de l'Académie des Sciences de l'Institut de France* 5: 176–184.
- Schwarz G (1978) Estimating the dimension of a model. *Annals of Statistics* 6(2): 461–464. <https://doi.org/10.1214/aos/1176344136>
- Shekhovtsov S, Peltek SE (2019) The complete mitochondrial genome of *Aporrectodea rosea* (Annelida: Lumbricidae). *Mitochondrial DNA. Part B, Resources* 4(1): 1752–1753. <https://doi.org/10.1080/23802359.2019.1610091>
- Shekhovtsov SV, Golovanova EV, Ershov NI, Poluboyarova TV, Berman DI, Bulakhova NA, Szederjesi T, Peltek SE (2020) Phylogeny of the *Eisenia nordenskiöldi* complex based on mitochondrial genomes. *European Journal of Soil Biology* 96: e103137. <https://doi.org/10.1016/j.ejsobi.2019.103137>
- Sims RW (1983) The scientific names of earthworms. In: Satchell JE (Ed.) *Earthworm Ecology: from Darwin to Vermiculture*. Chapman and Hall, London, 365–373. https://doi.org/10.1007/978-94-009-5965-1_40
- Sims RW, Gerard BM (1985) Earthworms. Keys and notes to the identification and study of the Species. *Synopsis of the British Fauna (New series)*, EJ. Brill, Leiden, 171 pp.
- Zhang L, Sechi P, Yuan M, Jiang J, Dong Y, Qiu J (2016) Fifteen new earthworm mitogenomes shed new light on phylogeny within the *Pheretima* complex. *Scientific Reports* 6(1): e20096. <https://doi.org/10.1038/srep20096>
- Zhang Q, Liu H, Zhang Y, Ruan H (2019) The complete mitochondrial genome of *Lumbricus rubellus* (Oligochaeta, Lumbricidae) and its phylogenetic analysis. *Mitochondrial DNA. Part B, Resources* 4(2): 2677–2678. <https://doi.org/10.1080/23802359.2019.1644242>