

**The complete mitochondrial DNA sequence of the pantropical earthworm *Pontoscolex corethrurus* (Rhinodrilidae, Clitellata): Mitogenome characterization and phylogenetic positioning**

Ana C. Conrado<sup>1</sup>, Hugo Arruda<sup>2</sup>, David W.G. Stanton<sup>3</sup>, Samuel W. James<sup>4</sup>, Peter Kille<sup>3</sup>, George Brown<sup>5</sup>, Elodie Silva<sup>5</sup>, Lise Dupont<sup>6</sup>, Shabnam Taheri<sup>6</sup>, Andrew J. Morgan<sup>3</sup>, Nelson Simões<sup>2</sup>, Armino Rodrigues<sup>7</sup>, Rafael Montiel<sup>8</sup> and Luis Cunha<sup>3,5\*</sup>

**Affiliations**

<sup>a</sup> Universidade Federal do Paraná, Departamento de Ciências do Solo, Curitiba, Paraná, Brazil

<sup>b</sup> CIRN/Departamento de Biologia, Universidade dos Açores, Ponta Delgada, Portugal

<sup>c</sup> Cardiff School of Biosciences, Cardiff University, Museum Avenue, Cardiff, UK CF10 3AX

<sup>d</sup> Maharishi University of Management, Sustainable Living Dept., Fairfield, IA, USA

<sup>e</sup> EMBRAPA-Florestas, Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Colombo, PR, Brazil

<sup>f</sup> Université Paris Est Créteil, Institut d'écologie et des sciences de l'environnement de Paris (IEES), Créteil, France

<sup>g</sup> Centro de Vulcanologia e Avaliação de Riscos Geológicos (CVARG), Universidade dos Açores, Ponta Delgada, Portugal

<sup>h</sup> Laboratorio Nacional de Genómica para la Biodiversidad, Unidad de Genómica Avanzada, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, México

\*Correspondence: Luis Cunha, Cardiff School of Biosciences, Cardiff University, Park Place, Cardiff, UK. E-mail: [luisnevescunha@gmail.com](mailto:luisnevescunha@gmail.com)

**Key Words:** *Pontoscolex corethrurus*, mitochondrial genome, Rhinodrilidae, earthworm

**Abstract:**

*Pontoscolex corethrurus* (Müller, 1857) plays an important role in soil terrestrial ecosystems and has been widely used as an animal model for a large variety of ecological studies, in particular due to its common presence and generally high abundance in human-disturbed tropical soils. In this study we describe the complete mitochondrial genome of the peregrine earthworm *Pontoscolex corethrurus*. This is the first record of a mitochondrial genome within the Rhinodrilidae family. Its mitochondrial genome is 14 835 bp in length containing 37 genes (13 protein-coding genes (PCG), 2 rRNA genes and 22 tRNA genes). It has the same gene content and structure as in other sequenced earthworms, but unusual among invertebrates it has several overlapping open reading frames. All genes are encoded on the same strand, most of the PCGs use ATG as the start codon except for NAD3, which uses GTG as the start codon. The T+A content of the mitochondrial genome is 59.9% (31.9% A, 27.9% T, 14.9% G, and 25.3% for C). The annotated genome sequence has been deposited in GenBank under the accession number KT988053.

**Keywords:** *Pontoscolex corethrurus*, mitochondria, mitochondrial genome, Azores, peregrine species, earthworm

## Introduction

Excluding a few aquatic taxa, earthworms (Annelida: Clitellata) are mostly terrestrial and include 5500 species (Blakemore et al., 2006). Believed to have originated in the Guyana Shield (Righi, 1984), the earthworm *Pontoscolex corethrurus* (Müller, 1857) is a pan-globally distributed species found in most tropical regions. It mainly occurs in human-disturbed areas and can be used as an indicator of ecosystem disturbance (Brown et al., 2006). The species formerly belonged in the Glossoscolecidae family, but was recently allocated to the Rhinodrilidae family by James (2012), following the phylogeny of James and Davidson (2012). It is also the most well-known earthworm species in the humid tropics, frequently used in ecological and agronomic studies (Bhattacharjee and Chaudhuri, 2002; Buch et al., 2013; Chapuis-Lardy et al., 2010; Dupont et al., 2012; Hamoui, 1991; Marichal et al., 2010). Being a geophagous endogeic species, *P. corethrurus* shows high plasticity regarding its tolerance to soil physicochemical characteristics, including variable moisture, high temperatures, exceptionally high carbon dioxide and low oxygen levels, and is capable of inhabiting nutrient-poor soils (Cunha et al., 2014; Hamoui, 1991; Lavelle et al., 1987), as well as rotten logs (Buch et al., 2011).

In animals, the mitochondrial DNA (mtDNA) typically contains 37 genes, encoding 13 proteins for the enzymes required for oxidative phosphorylation, the two ribosomal RNA units (rRNA), and 22 transfer RNAs (tRNAs) necessary for the translation of the proteins encoded by mtDNA (Anderson et al., 1981; Boore, 1999; Zhao et al., 2015). Molecular data have become increasingly important in recent years. Remarkable progress has been made over the past several years in the field of the molecular systematics of annelids. Compared with individual genes, the mitochondrial genome is still a promising tool for inferring phylogenetic relationships due to its high content of information, and has been applied in some phylogenetic studies involving earthworms (Zhang et al., 2015; Zhang et al., 2016b).

In this study, we sequenced the complete mtDNA sequence of *Pontoscolex corethrurus* for the first time and analyzed its mitochondrial genomic structure. Additionally, we conducted phylogenetic analyses based on the mitochondrial sequence data available elsewhere with the purpose of investigating the phylogenetic position of *P. corethrurus* within Clitellata. The information reported in this article will facilitate further investigations of phylogenetic relationships of different Annelida species.

## Material and Methods

### *Sample collection and DNA extraction*

A group of clitellate (adult) *P. corethrurus* were collected in São Miguel Island (Azores, Portugal) inside pineapple greenhouses (Locality: Fajã de Baixo, 37°45'12.2"N, 25°38'21.3"W) during January 2015. Animals were euthanized in 10% ethanol and preserved in 96% ethanol for later work. A piece of body wall tissue was used for genomic DNA extraction using standard phenol/chloroform (Sambrook and Russell, 2001) procedure followed by ethanol precipitation and kept at 4 °C for subsequent use.

### *Mitochondrial DNA amplification*

The complete *Pontoscolex corethrurus* mitogenome was amplified using seven sets of primers (Table 1) designed based on sequences retrieved from a previous study (Cunha et al. 2014).

**Table 1.** Details of the primers used to amplify the mitochondrial DNA of *Pontoscolex corethrurus*.

Primer code	Orientation	Annealing position (bp)	Nucleotide sequence (5'-3')	Melting Temperature (°C)
FP_1	Forward	2154..2175	CTCTACTATGTACCCAGGAGTG	57.46
FP_2	Forward	3121..3140	CCTAGTCCTAACTATCACCT	53.66
FP_3	Forward	4947..4965	TTCATGGGCCATCCTCACA	59.14
FP_4	Forward	5302..5320	CATTACAACATGCACGTCA	55.85
FP_5	Forward	7983..8004	ACTAGTGTCACTTACAACAACC	57.16
FP_6	Forward	8766..8787	AGTAGCCGCTATAATAGTCCTT	57.91
FP_7	Forward	10459..10478	AAAGCTTGGCGGTGCTTCAC	63.23
RP_1	Reverse	2758..2775	GCGGCCAAGATAAAGCAC	57.67
RP_2	Reverse	3740..3762	TAGAGGCGGTAAGGAGAAAGTAT	58.61
RP_3	Reverse	5691..5708	CAGAGGCGAGGTAAATTC	53.85
RP_4	Reverse	6356..6373	TGTTCAAGGGCTAGGATTG	54.99
RP_5	Reverse	8649..8670	TGATAAGGGGGAAAGTCTGATC	56.84
RP_6	Reverse	10328..10349	TGATTTGGGGTTCAGAGCCGTAG	61.59
RP_7	Reverse	11242..11263	CCTAGTGTGTGTCAGGACGCTT	64.75

Long PCR targets were amplified using different combinations of the primer sets, and and

initially sequenced with the same forward or reverse primers. Subsequent primer walking method was used to close the sequencing gaps. To ensure the accuracy of the sequence, every two contiguous segments overlapped by at least 80 bp. PCRs were performed using ~40 ng of DNA and 0.4  $\mu\text{M}$  forward and reverse primers, 0.2 mM dNTP mix and 1.25 U  $\mu\text{l}^{-1}$  Platinum HiFi DNA polymerase buffered with 1X Mg-free buffer (ThermoFisher Scientific, UK). PCR amplification buffer was supplemented with  $\text{MgCl}_2$  to achieve a final concentration of 1.75 mM in a total volume of 25  $\mu\text{l}$  reaction mixture. The reaction was denatured at 95 °C for 3 min, cycled 35 times at 95 °C for 30 s, 30 s at the required primer annealing temperature and 72 °C for 1 min per 1000 bp depending on target fragment length. Negative controls were included in all PCR amplifications to confirm the absence of contaminants. Before sequencing, PCR cleanups were performed using Exo-SAP-IT (Amersham Pharmacia, UK) reagents. Exonuclease 1 (0.25  $\mu\text{l}$ ) and Shrimp Alkaline Phosphatase (0.5  $\mu\text{l}$ ) were mixed with the PCR product (10  $\mu\text{l}$ ) and incubated at 37 °C for 45 min followed by 80 °C for 15 min. DNA was sequenced using ABI PRISM® BigDye v3.1 Terminator Sequencing technology (Applied Biosystems, USA) on an ABI PRISM® 3100 DNA automated Sequencer.

#### *Sequence editing and analysis*

Sequence trace files were corrected and aligned with the MEGA v.7 (Kumar et al., 2016). The sequence overlap and mitogenome assembly were performed using CLC main workbench v.6 (Qiagen). The annotation of the 13 protein-coding genes and two rRNA genes were determined using the MITOS v.2 web server (Bernt et al., 2013) and manually curated using other published annelid mitogenomes as shown in Table 2, whereas the tRNA genes were identified using the program tRNAscan-SE 1.21 (Lowe and Eddy, 1997). The annotated genome sequence was deposited in GenBank under accession number KT988053.

#### *Phylogenetic analyses*

To clarify the phylogenetic position of *P. corethrurus* within the Clitellata, the complete mitogenome sequences of eight representative Clitellata species (Table 2) were incorporated together with the presently obtained *P. corethrurus* mitogenome sequence for phylogenetic analysis. Sequences were aligned using MEGA v.7 (Kumar et al., 2016), and adjustments were made manually. Phylogenetic analyses were based on the concatenated sequences of 13 protein-

coding genes and the two rRNA units. The possible bias of substitution saturation at each codon position of protein-coding genes and two rRNA genes was investigated using DAMBE v.4.5.57 (Xia and Xie, 2001) and MEGA v. 7 (Kumar et al., 2016).

Two different methods, Bayesian inference (BI) and maximum likelihood (ML) were used to construct the phylogenetic tree. Bayesian analyses were undertaken with MrBayes v.3.1.2 (Ronquist and Huelsenbeck, 2003) under the best-fit model of nucleotide evolution selected in MrModeltest v.2.3 (Nylander, 2004). Analyses were run for 1,000,000 generations, and sampled every 100 generations to assess convergence. Initial trees with non-stationary log-likelihood values were discarded as part of a burn-in procedure and combined the remaining trees that resulted in convergent log-likelihood scores from both independent searches. These trees were used to construct a consensus tree.

**Table 2.** Representative Clitellata species included in this study for comparison.

Species	Family	Length (bp)	GenBank accession number
<i>Pontoscolex corethrurus</i>	Rhinodrilidae	14,835	<b>Present study</b>
<i>Tonoscolex birmanicus</i>	Megascolecidae	15,170	KF425518
<i>Amyntas gracilis</i>	Megascolecidae	15,161	NC_027258
<i>Duplodicrodrilus schmardae</i>	Megascolecidae	15,156	NC_029867
<i>Metaphire guillemi</i>	Megascolecidae	15,174	NC_029869
<i>Perionyx excavatus</i>	Megascolecidae	15,083	NC_009631
<i>Lumbricus terrestris</i>	Lumbricidae	14,998	NC_001673
<i>Drawida japonica</i>	Moniligastridae	14,648	NC_028050
<i>Hirudo nipponia</i>	Hirudinidae	14,414	NC_023776

Maximum likelihood analysis (ML) was performed with MEGA v.7 (Kumar et al., 2016). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, alpha parameter = 0.9143)). The rate variation model allowed for some sites to be

evolutionarily invariable ([+I], 17.7596% sites). The analysis involved nine mitogenome sequences (Table 2). All positions containing gaps and missing data were eliminated. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analysed (Felsenstein, 1985).

## Results and discussion

### *Mitochondrial genomic structure*

The mitochondrial genome of *P. corethrurus* was determined to be 14 835 bp in length, comprising 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs), and one putative control region with a length of 318 bp (Figure 1).

The structure of the mitochondrial genome of *Pontoscolex corethrurus* is detailed in Table 3. The gene order and orientation are similar to the previous earthworm mitochondrial genomes (Boore and Brown, 1995; Zhang et al., 2015) but slightly smaller and more condensed (with several intergenic overlaps, see Table 3). The gene organization is similar to other earthworm species (e.g. *Lumbricus terrestris*: Boore and Brown, 1995).

**Table 3.** Organisation and structure details of the mitochondrial genome of *P. corethrurus*.

Gene	Direction	From	To	Size (bp)	Start	Stop	Anti-codon	Intergenic bases (bp)
COI	+	1	1540	1540	ATG	T--		0
tRNA- <sup>Asn</sup>	+	1541	1602	62			GTT	0
COII	+	1603	2289	687	ATG	TAG		-1
tRNA- <sup>Asp</sup>	+	2289	2351	63			GTC	2
ATP8	+	2354	2513	160	ATG	T--		0
tRNA- <sup>Tyr</sup>	+	2514	2576	63			GTA	-1
tRNA- <sup>Gly</sup>	+	2576	2638	63			TCC	3
COIII	+	2642	3419	778	ATG	T--		0
tRNA- <sup>Gln</sup>	+	3420	3488	69			TTG	0
ND6	+	3489	3954	466	ATG	T--		0
Cytb	+	3955	5094	1140	ATG	TAA		-2
tRNA- <sup>Trp</sup>	+	5092	5154	63			TCA	1
ATP6	+	5156	5851	696	ATG	TAA		-2
tRNA- <sup>Arg</sup>	+	5850	5910	61			TCG	0
<i>Putative Control Region</i>	+	5911	6228	318				
tRNA- <sup>His</sup>	+	6229	6288	60			GTG	0

ND5	+	6289	8010	1722	ATG	TAA		3
tRNA <sup>-Phe</sup>	+	8014	8073	60			GAA	4
tRNA <sup>-Glu</sup>	+	8078	8141	64			TTC	0
tRNA <sup>-Pro</sup>	+	8142	8204	63			TGG	4
tRNA <sup>-Thr</sup>	+	8209	8272	64			TGT	0
ND4L	+	8273	8569	297	ATG	TAA		-7
ND4	+	8563	9921	1359	ATG	TAA		-2
tRNA <sup>-Cys</sup>	+	9920	9986	67			GCA	1
tRNA <sup>-Met</sup>	+	9988	10050	63			CAT	-1
s-rRNA	+	10050	10838	789				-7
tRNA <sup>-Val</sup>	+	10832	10894	63			TAC	-2
l-rRNA	+	10893	12104	1212				0
tRNA <sup>-Leu 1</sup>	+	12105	12166	62			TAG	2
tRNA <sup>-Ala</sup>	+	12169	12230	62			TGC	1
tRNA <sup>-Ser 2</sup>	+	12232	12293	62			TGA	1
tRNA <sup>-Leu 2</sup>	+	12295	12360	66			TAA	0
ND1	+	12361	13290	930	ATG	TAG		-1
tRNA <sup>-Ile</sup>	+	13290	13354	65			GAT	0
tRNA <sup>-Lys</sup>	+	13355	13417	63			TTT	0
ND3	+	13418	13770	353	GTG	TA-		-1
tRNA <sup>-Ser 1</sup>	+	13770	13832	63			TCT	0
ND2	+	13833	14835	1003	ATG	T--		0

The nucleotide composition is asymmetric (31.92% A, 25.26% T, 14.88% G, and 31.45% for C) with an overall A+T content of 59.86%. One remarkable trait of metazoan mitogenomes is the strand-specific bias in nucleotide composition (Hassanin et al., 2005; Reyes et al., 1998). Such bias is measured as G/C-skew  $(G\%-C\%)/(G\%+C\%)$  and A/T-skew  $(A\%-T\%)/(A\%+T\%)$ , respectively (Perna and Kocher, 1995). The overall GC- and AT-skews of the H-strand of *P. corethurus* mitogenome were -0.258 and 0.066, respectively, indicating a compositional bias associated with an excess of C over G nucleotides and a slight excess of A over T nucleotides on the H-strand.

### *Protein-coding genes*

The *P. corethurus* genome contained the expected 13 protein-coding genes with a total of 11,131 bp in size, accounting for 75.03% of the whole mitogenome. Most of the PCGs are initiated with ATG codons, except for ND3 gene, which uses GTG as the initiation codon. Six PCGs (COX1, ATP8, COX3, ND6, ND3, and ND2) are terminated with an incomplete codon T or TA,



which could be completed to TAA by polyadenylation post-transcriptionally (Ojala et al., 1981). COX2 and ND1 use TAG as a termination codon.

Nucleotide composition and codon usage frequencies were calculated from a concatenated sequence of all protein-coding genes on the H-strand. The base composition of protein-coding genes revealed a negative bias for A (14.67%), especially at second codon positions (12.88%, Table 4). For all protein genes, T was the most frequent nucleotide at the first and third positions whereas G was most frequent at the second position.

#### *Ribosomal and transfer RNA genes*

Like other mitochondrial genomes (Inoue et al., 2000; Zardoya et al., 1995), twenty-two tRNA genes were identified (Supplementary figure 1). The tRNA genes were scattered throughout the mitochondrial genome and ranged in size from 60 to 67 bp (Table 3). The *P. corethurus* mitogenome also contained a small subunit of rRNA and a large subunit of rRNA, which were 789 bp and 1212 bp in length, respectively. As in other Clitellata genomes, these genes were located between the tRNA<sup>Met</sup> and tRNA<sup>Val</sup> genes and between tRNA<sup>Val</sup> and tRNA<sup>Leu</sup> genes, respectively (Zhang et al., 2016b).

#### *Non-coding regions*

As shown in Table 3, there are 22 intergenic spacer regions, ranging in size from -7 to 4 bp observed in *P. corethurus*.

As in most Clitellata, the major non-coding region in *P. corethurus* mitochondrial genome was located between tRNA<sup>Arg</sup> and tRNA<sup>His</sup>. It was determined to be 318 bp in length, less than other reported Clitellata species (Zhang et al., 2016a), and it had a base composition that was rich in A and T (A+T=67,6%).

**Table 4.** Base composition for protein-coding, tRNA, and rRNA genes of *P. corethurus* mitogenome.

Gene/Region	Base composition (%)				A+T (%)	Size (bp)
	T	C	A	G		
COI	27.4	26.7	27.9	18.1	55.3	1,540

COII	25.76	24.31	34.93	14.99	60.7	687
ATP8	24.38	26.88	36.88	11.88	61.25	160
COIII	27.25	27.51	25.71	19.54	52.96	778
ND6	28.33	25.97	30.26	15.45	58.58	466
Cytb	27.98	27.11	29.65	15.26	57.63	1,140
ATP6	29.6	29.17	30.6	10.63	60.2	696
ND5	27.7	26.66	31.71	13.94	59.41	1,722
ND4L	25.93	28.28	33	12.79	58.92	297
ND4	27.45	27.52	32.3	12.73	59.75	1,359
ND1	28.39	25.81	29.78	16.02	58.17	930
ND3	32.58	26.06	27.2	14.16	59.77	353
ND2	30.01	28.22	30.71	11.07	60.72	1,003
<hr/>						
Protein Coding	<hr/>					
1st	30.19	25.39	17.39	27.04	47.58	3,710
2st	24.93	27.57	12.88	34.61	37.81	3,710
3st	36.08	27.89	13.75	22.29	49.83	3,710
Total	30.4	26.95	14.67	27.98	45.07	11,131
tRNA	30.5	17.6	34.9	17	65.45	1,391
rRNA	24.1	22	37.6	16.3	61.67	2,001
Putative Control Region	31.45	18.87	36.16	13.52	67.61	318
Overall	27.94	25.26	31.92	14.88	59.86	14,835

### *Phylogenetic analyses within the Clitellata*

The phylogenetic trees (the 50% majority-rule consensus tree is shown in Figure 2) were highly consistent regardless of the analytic method used, and were statistically supported by high posterior probability and intermediate bootstrap values. This phylogenetic analysis represented the first investigation of *P. corethurus* relationships within the Clitellata based on the complete mitogenome. As indicated by the tree, different species from the same family clustered together (Megascolecidae: *M. guillemi*, *D. schmardae*, *A. gracilis*, *P. excavatus* and *T. birmanicus*), and the species from Lumbricidae and the *P. corethurus* formed a monophyletic group. The species *D. japonica* belongs to the Moniligastridae, the sister group to Crassiclitellata (earthworms), which explains its phylogenetic position. The Moniligastridae are not Crassiclitellata because they have a single cell layer in the clitellum.

### **Conclusion**

For the first time, the sequencing, annotation and analysis of the mitochondrial genome of a member of Rhinodrilidae was completed in this study. The mitogenome of *P. corethurus* was

found to be 14,835 bp in length and showed a similar composition in size, low GC content and gene order to earthworm mitogenomes already available. The complete mitogenome reported here is expected to allow for further studies of the *P. corethrurus* phylogeny and for analyses on the taxonomic status of the family Rhinodrilidae.

### Acknowledgments

We thank Jose Talavera for the taxonomic identification of the studied specimen. This study was financially supported by the Portuguese Government (FCT project PTDC/AAC-AMB/115713/2009). Luis Cunha was supported by an EU Marie Curie fellowship - MSCA-IF-2014-GF-660378. George Brown, Elodie da Silva acknowledge fellowship support by CNPq. Ana Caroline Conrado was supported by a scholarship funded by CAPES. Sam James was supported by USA NSF award DEB-1136604. Dave Stanton was supported by a NERC grant (NE/M017656/1).

### Declaration of Interest Section

The authors report no conflicts of interest and are responsible for the content and writing of the paper.

### References

- Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M., Coulson, A.R., Drouin, J., Eperon, I., Nierlich, D., Roe, B.A., Sanger, F., 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290, 457-465.
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69, 313-319.
- Bhattacharjee, G., Chaudhuri, P., 2002. Cocoon production, morphology, hatching pattern and fecundity in seven tropical earthworm species — a laboratory-based investigation. *J. Biosci.* 27, 283-294.
- Blakemore, R., Ito, M., Kaneko, N., 2006. Alien earthworms in the Asia/Pacific region with a checklist of species and the first records of *Eukerria saltensis* (Oligochaeta: Ocnerodrilidae) and *Eiseniella tetraedra* (Lumbricidae) from Japan, and *Pontoscolex corethrurus* (Glossoscolecidae) from Okinawa, in: Koike, F., Clout, M.N., Kawamichi, M., De Poorter, M., Iwatsuki, K. (Eds.), *Assessment and control of biological invasion risks*. Shoukadoh Book Sellers/IUCN, Kyoto, Japan, pp. 173-181.

- Boore, J.L., 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767-1780.
- Boore, J.L., Brown, W.M., 1995. Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. *Genetics* 141, 305-319.
- Brown, G.G., James, S.W., Pasini, A., Nunes, D.H., Benito, N.P., Martins, P.T., Sautter, K.D., 2006. Exotic, peregrine, and invasive earthworms in Brazil: diversity, distribution, and effects on soils and plants. *Caribb. J. Sci.* 42, 339-358.
- Buch, A.C., Brown, G.G., Niva, C.C., Sautter, K.D., Lourençato, L.F., 2011. Life cycle of *Pontoscolex corethrurus* (Müller, 1857) in tropical artificial soil. *Pedobiologia* 54, S19-S25.
- Buch, A.C., Brown, G.G., Niva, C.C., Sautter, K.D., Sousa, J.P., 2013. Toxicity of three pesticides commonly used in Brazil to *Pontoscolex corethrurus* (Müller, 1857) and *Eisenia andrei* (Bouché, 1972). *Appl. Soil Ecol.* 69, 32-38.
- Chapuis-Lardy, L., Brauman, A., Bernard, L., Pablo, A.-L., Toucet, J., Mano, M., Weber, L., Brunet, D., Razafimbelo, T., Chotte, J.-L., 2010. Effect of the endogeic earthworm *Pontoscolex corethrurus* on the microbial structure and activity related to CO<sub>2</sub> and N<sub>2</sub>O fluxes from a tropical soil (Madagascar). *Appl. Soil Ecol.* 45, 201-208.
- Cunha, L., Montiel, R., Novo, M., Orozco-terWengel, P., Rodrigues, A., Morgan, A.J., Kille, P., 2014. Living on a volcano's edge: genetic isolation of an extremophile terrestrial metazoan. *Heredity* 112, 132-142.
- Dupont, L., Decaëns, T., Lapied, E., Chassany, V., Marichal, R., Dubs, F., Maillot, M., Roy, V., 2012. Genetic signature of accidental transfer of the peregrine earthworm *Pontoscolex corethrurus* (Clitellata, Glossoscolecidae) in French Guiana. *Eur. J. Soil Biol.* 53, 70-75.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 783-791.
- Hamoui, V., 1991. Life-cycle and growth of *Pontoscolex corethrurus* (Müller, 1857)(Oligochaeta, Glossoscolecidae) in the laboratory. *Rev. Ecol. Biol. Sol* 28, 469-478.
- Hassanin, A., Léger, N., Deutsch, J., 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. *Syst. Biol.* 54, 277-298.
- Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2000. Complete mitochondrial DNA sequence of the Japanese sardine *Sardinops melanostictus*. *Fish. Sci.* 66, 924-932.
- James, S.W., 2012. Re-erection of Rhinodrillidae Benham, 1890, a senior synonym of Pontoscolecidae James, 2012 (Annelida: Clitellata). *Zootaxa* 3540, 67-68.
- James, S.W., Davidson, S.K., 2012. Molecular phylogeny of earthworms (Annelida: Crassicitellata) based on 28S, 18S and 16S gene sequences. *Invertebr. Syst.* 26, 213-229.

- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, msw054.
- Lavelle, P., Barois, I., Cruz, I., Fragoso, C., Hernandez, A., Pineda, A., Rangel, P., 1987. Adaptive strategies of *Pontoscolex corethrurus* (Glossoscolecidae, Oligochaeta), a peregrine geophagous earthworm of the humid tropics. *Biol. Fertil. Soils* 5, 188-194.
- Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955-964.
- Marichal, R., Martinez, A.F., Praxedes, C., Ruiz, D., Carvajal, A.F., Oszwald, J., del Pilar Hurtado, M., Brown, G.G., Grimaldi, M., Desjardins, T., 2010. Invasion of *Pontoscolex corethrurus* (Glossoscolecidae, Oligochaeta) in landscapes of the Amazonian deforestation arc. *Appl. Soil Ecol.* 46, 443-449.
- Müller, F., 1857. II.—Description of a new species of Earth-worm (*Lumbricus corethrurus*). *J. Nat. Hist.* 20, 13-15.
- Nylander, J., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University 2.
- Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290, 470-474.
- Perna, N.T., Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* 41, 353-358.
- Reyes, A., Gissi, C., Pesole, G., Saccone, C., 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol. Biol. Evol.* 15, 957-966.
- Righi, G., 1984. *Pontoscolex* (Oligochaeta, Glossoscolecidae), a new evaluation. *Stud. Neotrop. Fauna Environ.* 19, 159-177.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574.
- Sambrook, J., Russell, D., 2001. *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York.
- Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.* 92, 371-373.
- Zardoya, R., Garrido-Pertierra, A., Bautista, J.M., 1995. The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. *J. Mol. Evol.* 41, 942-951.
- Zhang, L., Jiang, J., Dong, Y., Qiu, J., 2015. Complete mitochondrial genome of four pheretimoid earthworms (Clitellata: Oligochaeta) and their phylogenetic reconstruction. *Gene* 574, 308-316.
- Zhang, L., Jiang, J., Dong, Y., Qiu, J., 2016a. Complete mitochondrial genome of an *Amyntas* earthworm,

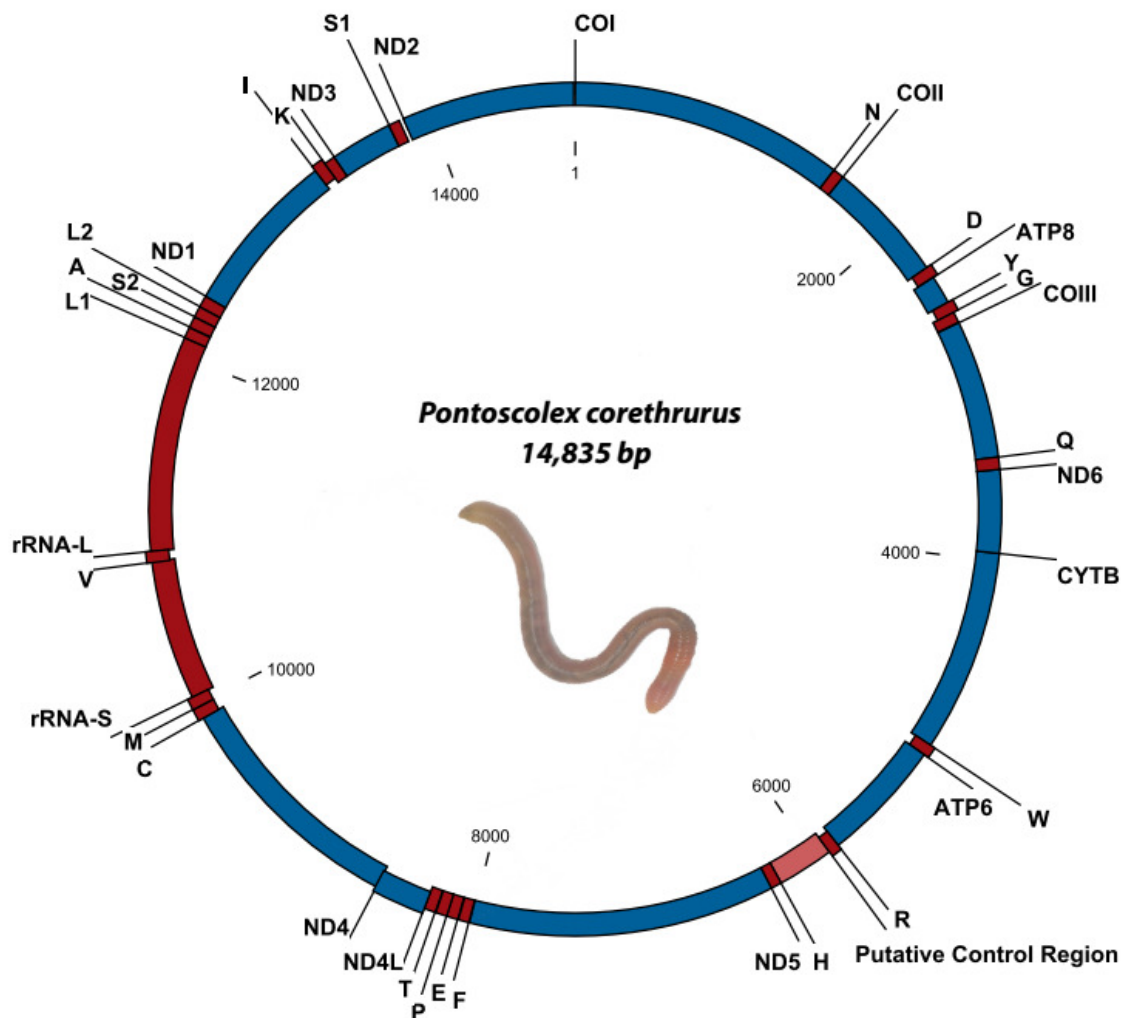
*Amyntas aspergillus* (Oligochaeta: Megascolecidae). Mitochondrial DNA 27, 1876-1877.

Zhang, L., Sechi, P., Yuan, M., Jiang, J., Dong, Y., Qiu, J., 2016b. Fifteen new earthworm mitogenomes shed new light on phylogeny within the Pheretima complex. Sci. Rep. 6.

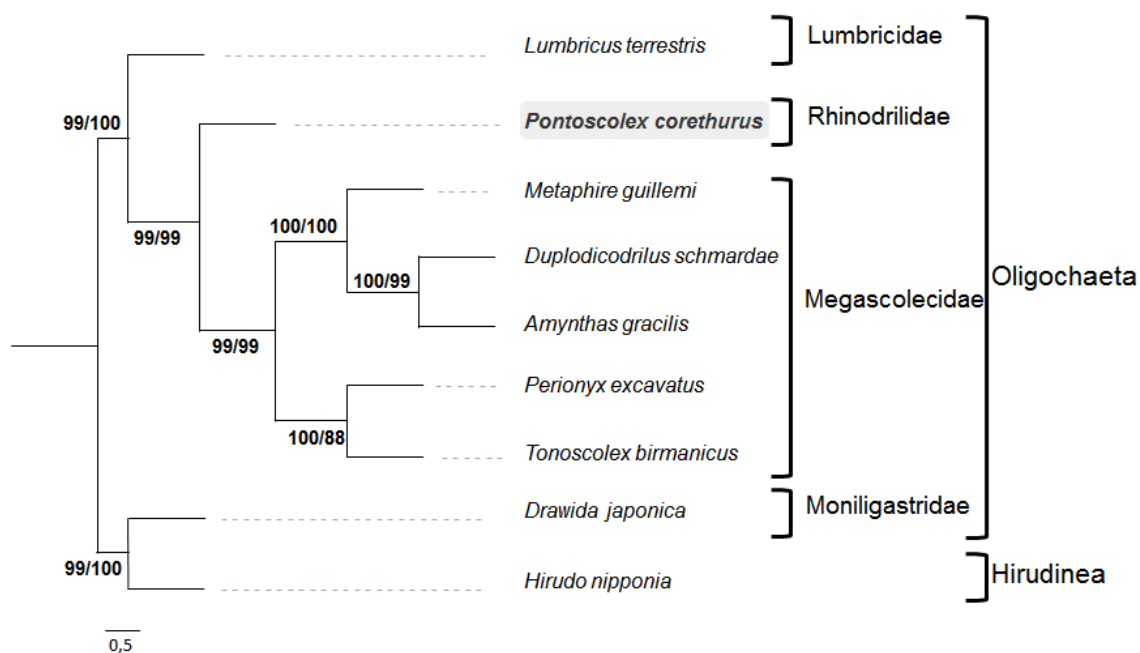
Zhao, L., Gao, T., Lu, W., 2015. Complete mitochondrial DNA sequence of the endangered fish (*Bahaba taipingensis*): Mitogenome characterization and phylogenetic implications. ZooKeys, 181.

## Figures

**Figure 1.** The mitochondrial genome of *Pontoscolex corethrurus* (Müller 1856). Gene order and positions are shown, including the putative control region. IUPAC single letter codes are used to identify transfer RNA. The L1, L2, S1, and S2 transfer RNAs are differentiated on the basis of their anti-codons TAG, TAA, TCT, and TGA, respectively.



**Figure 2.** Phylogenetic relationships among phylum Annelida based on the combined 13416 bp nucleotide positions. The posterior probability value of BI analyses and bootstrap support values of ML analyses (in the order: BI, ML) are indicated near the branches.



#### Highlights:

- The mitochondrial genome of *Pontoscolex corethurus* (Müller, 1857) was determined.
- For the first time, the sequencing, annotation and analysis of the mitochondrial genome of a member of Rhinodrilidae was completed in this study.
- The mitogenome of *P. corethurus* was found to be 14,835 bp in length and showed a similar composition in size, low GC content, and gene order to earthworm mitogenomes already available.
- The complete mitogenome reported here is expected to allow for further studies of the *P. corethurus* phylogeny and for analyses on the taxonomic status of the family Rhinodrilidae.