Phylogenetic assessment of the earthworm *Aporrectodea caliginosa* species complex (Oligochaeta: Lumbricidae) based on mitochondrial and nuclear DNA sequences

Marcos Pérez-Losada [a,*], Maigualida Ricoy [b], Jonathon C. Marshall [c], Jorge Domínguez [b]

[a] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[b] Departamento de Ecoloxía e Bioloxía Animal, Universidade de Vigo, E-36310, Spain
[c] Department of Zoology, Weber State University, 2505 University Circle, Ogden, UT 84408, USA

ARTICLE INFO

ABSTRACT

The *Aporrectodea caliginosa* species complex includes the most abundant earthworms in grasslands and agricultural ecosystems of the Paleartic region. Historically this complex consisted of the following taxa: *A. caliginosa s.s.* Savigny, 1826, *A. trapezoides* Dugés (1828), *A. tuberculata* (Eisen, 1874), and *A. nocturna* Evans (1946). These four taxa are morphologically very similar and difficult to differentiate because of their morphological variability. Consequently, their taxonomic status and their phylogenetic relationships have been a matter of discussion for more than a century. To study these questions, we sequenced the COII (686 bp), 12S (362 bp), 16S (1200 bp), ND1 (917 bp), and tRNA-s$_{Asn-Asp-Val-Leu-Ala-Ser-Leu}$ (402 bp) mitochondrial and 28S (809 bp) nuclear gene regions for 85 European earthworms from 27 different localities belonging to the *A. caliginosa* species complex and four outgroup taxa. DNA sequences were analyzed using maximum parsimony, maximum likelihood, and Bayesian approaches of phylogenetic inference. The resulting trees were combined with morphological, ecological, and genomic evidence to test species boundaries (i.e., integrative approach). Our molecular analyses showed that *A. caliginosa s.s.* and *A. tuberculata* form a sister clade to *A. trapezoides*, *A. longa*, and *A. nocturna*, which indicates that *A. longa* is part of the *A. caliginosa* species complex. We confirm the species status of all these taxa and identify two hitherto unrecognized *Aporrectodea* species in Corsica (France). Moreover our analyses also showed the presence of highly divergent lineages within *A. caliginosa*, *A. trapezoides*, and *A. longa*, suggesting the existence of cryptic diversity within these taxa.

1. Introduction

Although morphology has traditionally been the basis of earthworm species delimitation (Savigny, 1826; Rosa, 1893; Michaelsen, 1900; Omodeo, 1956; Gates, 1972; Bouché, 1972; Perel, 1973, 1976; Zicsi, 1982, 1991; Mršić, 1991; Qiu and Bouché, 1998), earthworm taxonomy is somewhat restricted by the structural simplicity of these invertebrates, which lack complex appendices or highly specialized copulatory apparatuses. Moreover, as earthworms are soft-bodied animals, there is a scarce fossil record (Piearce et al., 1990) and it has therefore been difficult to discern ancestral and evolved characters. Lumbricidae earthworms are no exception, as their taxonomy is still far from being resolved despite being the most widely studied and one of the most broadly distributed earthworm groups (Pop, 2004). The identification of adult lumbricids is principally based on the type of *prostomium*, arrangement of the setae, position and form of the *clitellum*, *tubercula pubertatis*, and some internal organs such as the seminal vesicles and the spermathecae. However, these morphological and anatomical characters are variable, and different taxa may display overlapping variability in the same character (Pop et al., 2003). The lack of taxonomically useful characters has led to many morphologically similar species being lumped into a single species with various morphotypes or as a species complex that includes various taxa of uncertain taxonomic category (Bouché, 1972; Gates, 1972; Sims and Gerard, 1985; Briones, 1993, 1996). Another contributing factor to this poor earthworm taxonomy has been the insistence by some specialists that convenience of identification must be a priority in systematics of Lumbricidae, without regarding to details of evolutionary history.

Considering the important role that earthworms play as key organisms in terrestrial ecosystems (Domínguez et al., 2004), the failure to recognize accurate species boundaries within this group compromises many aspects of applied ecological, biodiversity, sys-

* Corresponding author.
E-mail addresses: mlosada@genoma-llc.com, mlosada323@gmail.com (M. Pérez-Losada).

tematic, and evolutionary studies (Domínguez et al., 2004, 2005; Pérez-Losada et al., 2005; King et al., 2008).

In this study we investigate phylogenetic relationships and delimit species boundaires within the species complex *Aporrectodea caliginosa* (Lumbricidae), the most abundant earthworm from Paleartic grassland regions and the most commonly found in agricultural ecosystems across the temperate zone. Historically, it was thought the *A. caliginosa* species complex included three species, *A. caliginosa s.s.* Savigny, 1826, *A. trapezoides* Dugés (1828), and *A. nocturna* (Evans, 1946), and one subspecies, *A. c. tuberculata* (Eisen, 1874), although this view has been challenged several times. As in other lumbricids, these four taxa are morphologically very similar and the characters that differentiate them are highly variable, making species identification a difficult task. *A. caliginosa* and *A. tuberculata*, for example, lack pigmentation, whereas *A. trapezoides* and *A. nocturna* are brown; however, it is possible to find specimens with intermediate pigmentation. On the other hand, the position of the *clitellum* in the *A. caliginosa* species complex occurs within the same range of segments, but the form and position of the *tubercula pubertatis* differ—they appear as two protuberances in *A. caliginosa*, *A. tuberculata*, and *A. nocturna*, and as two lateral bands in *A. trapezoides*; however, it is also possible to find specimens with extended protuberances that form a band, and bands with protuberances.

Because of their similarity, the taxonomic status of the taxa within the *A. caliginosa* species complex has been a matter of debate for more than a century. Based on morphological data, *A. caliginosa s.s.*, *A. trapezoides*, and *A. nocturna* were initially described as distinct species, whereas *A. tuberculata* was described as a subspecies of *A. caliginosa*. Michaelsen in 1900 noticed that some of these taxa were closely related and included them in a species complex, but he suggested that they belonged to a single species with two subspecies: *A. caliginosa caliginosa* and *A. c. trapezoides*, and considered the other taxa as synonymous to *A. caliginosa*. Omodeo (1952) and Casellato (1987) considered *A. trapezoides* the polyploidal variety of *A. caliginosa s.s.* Gates (1972) disagreed with Michaelsen (1900) and separated them into four distinct species [*A. caliginosa s.s.* (namely *A. turgida* Eisen 1873), *A. tuberculata*, *A. trapezoides*, and *A. nocturna*]. However, the same year, Bouché (1972) split them into two species and placed them into a different genus, *Nicodrilus caliginosus* (= *A. caliginosa*) and *N. nocturnus* (= *A. nocturna*), with the former species composed of three subspecies: *N. c. caliginosus* (= *A. c. caliginosa*), *N. c. alternisetosus* (= *A. tuberculata*), and *N. c. meridionalis* (= *A. trapezoides*). Later, Sims and Gerard (1985) suggested that these four taxa formed part of a highly variable single species (*A. caliginosa s.l.*), which displayed four forms or phenotypic varieties: *A. caliginosa s.s.*, *A. caliginosa var. trapezoides*, *A. caliginosa var. tuberculata*, *A. caliginosa var. nocturna*. Finally, almost a century after Michaelsen's study, Briones (1996) resurrected his initial proposal suggesting that the *A. caliginosa* species complex is composed of one species with two subspecies (*A. caliginosa caliginosa* and *A. c. trapezoides*).

Molecular data coming from enzyme electrophoresis (Bøgh, 1992), karyotyping (Mezhzherin et al., 2008), random amplified polymorphic DNA (RAPD) (Dyer et al., 1998), and 16S and cytochrome oxidase I DNA barcode sequences (Pop et al., 2006) have not solved this taxonomic riddle either. Taxon-wise, these analyses included different *Aporrectodea* species, which makes them difficult to compare, and all of them are lacking *A. nocturna*; moreover allozymes and RAPD have limited resolution and Pop et al. (2006) only included two *Aporrectodea* species (*A. caliginosa* and *A. trapezoides*) in their study. Nonetheless, all these studies combined suggest the possibility that *A. caliginosa*, *A. trapezoides*, and *A. tuberculata* are different species and that *A. trapezoides* may be of hybrid origin.

Therefore, given the complexity of *Aporrectodea* alpha-taxonomy and the limitations of the analytical methods and marker types used in some of the previous studies, here we use multi-locus DNA sequencing to assess phylogenetic relationships and species boundaries within the *A. caliginosa* species complex. To this end, we will examine twelve mitochondrial and nuclear DNA gene regions in European samples of *A. caliginosa s. s.*, *A. trapezoides*, *A. tuberculata*, and *A. nocturna* (ingroup) and four outgroups (*A. limicola*, *A. longa*, *A. molleri*, and *A. rosea*). DNA sequences will be analyzed using maximum likelihood, maximum parsimony, and Bayesian approaches of phylogenetic inference. Resulting trees will be then combined with morphological, ecological and other genomic evidence to determine species boundaries (i.e., integrative approach) within the *A. caliginosa* species complex.

2. Material and methods

2.1. *Aporrectodea* earthworm sampling

A total of 68 specimens of *A. caliginosa s. s.*, *A. tuberculata*, *A. trapezoides*, and *A. nocturna* (*A. caliginosa* species complex) were collected in 27 different locations from western and central Europe (Fig. 1 and Table 1). Additionally, 17 specimens belonging to other *Aporrectodea* species (*A. limicola*, *A. longa*, *A. molleri*, and *A. rosea*) were collected to be used as the outgroup (Fig. 1 and Table 1). *Aporrectodea* is considered paraphyletic (Pop et al., 2006), but to our knowledge, no one has comprehensively studied their phylogenetic relationships; hence, our outgroup choice was based in species availability. All *Aporrectodea* specimens in this study were identified following the taxonomic key in Blakemore (2006).

2.2. DNA extraction, amplification, and sequencing

Total genomic DNA was extracted using the DNAeasy Tissue kit (Qiagen). Regions of the nuclear 28S rDNA and mitochondrial 16S rDNA, 12S rDNA, NADH dehydrogenase (ND1), cytochrome oxidase subunit II (COII) and tRNA Asn, Asp, Val, Leu, Ala, Ser, and Leu genes were amplified using the polymerase chain reaction (PCR). We used similar PCR conditions to those in Pérez-Losada et al. (2005) and the primers listed in Table 2. PCR products were resolved by 1.5% agarose gel electrophoresis, visualized by SYBR Green, and purified using a MultiScreen PCRµ96 (Millipore) kit. Automated sequences were generated in both directions from different runs on an Applied Biosystems (ABI) 377XL automated sequencer. We used the ABI Big-dye Ready-Reaction kit and followed the standard cycle sequencing protocol, but using a 16th of the suggested reaction size. All PCR products gave unequivocal nucleotide chromatograms. All DNA sequences were deposited in GenBank under the Accession Nos. FJ967163 – FJ967792.

2.3. Data analysis

Nucleotide sequences from each gene region (all tRNAs were combined into a single gene region) were aligned using MAFFT v5.7 (Katoh et al., 2005) under iterative refinement methods incorporating the most accurate local (L-INS-i and E-INS-i) and global (G-INS-i) pairwise alignment information. Default settings were chosen for all the parameters involved under each algorithm. Multiple sequence alignments (MSA) for each gene resulting from these three methods were concatenated and maximum likelihood (ML) trees were estimated using PhyML (Guindon and Gascuel, 2003). The G-INS-i pairwise alignment (4554 sites) generated the trees with the best likelihood scores; hence, we used this MSA for our subsequent phylogenetic analyses. Phylogenetic congruence among gene regions (COII: 686 bp, 12S: 362 bp, 16S:
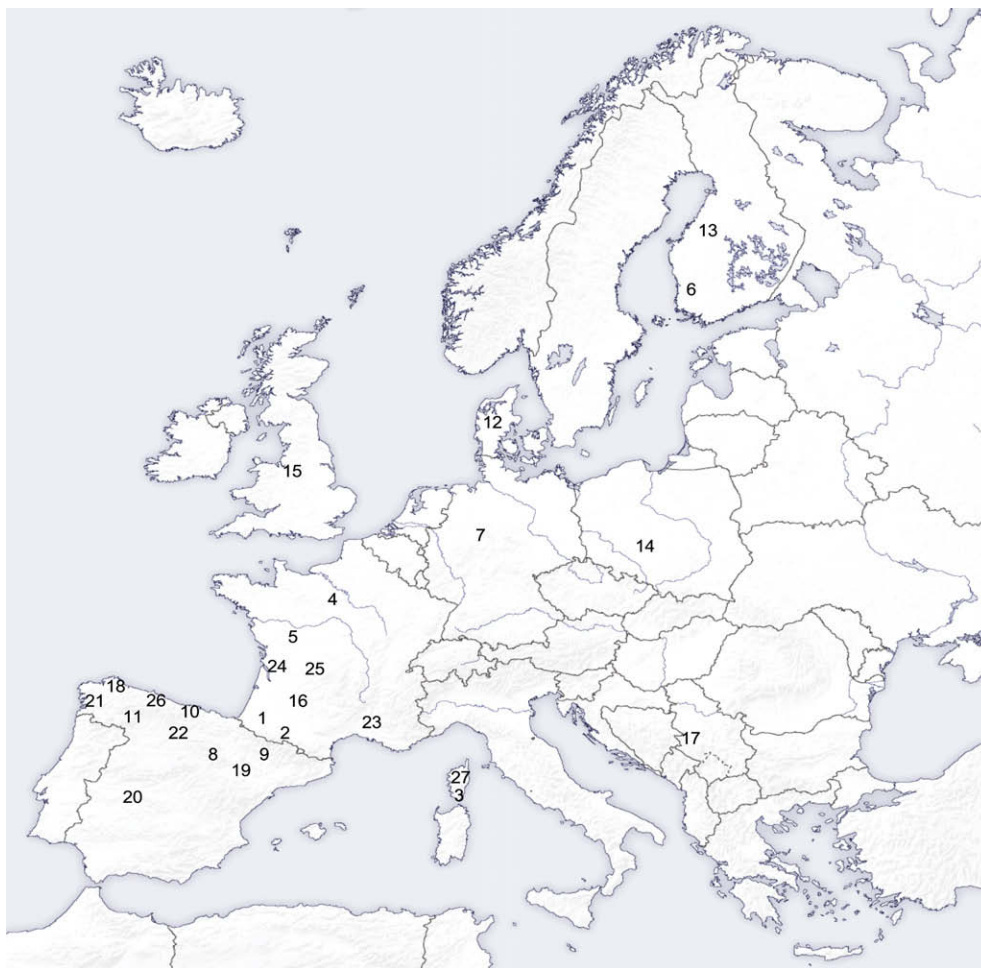
Fig. 1. Localities sampled. See Table 1 for details.

1200 bp, ND1: 917 bp, tRNAs: 402 bp, and 28S: 809 bp) was assessed using the Wiens' (1998) protocol. No areas of strongly supported incongruence were observed in our alignment. Gene regions were analyzed both in combination as a single dataset and as multiple concatenated partitions. Maximum parsimony (MP) trees were inferred using the combined dataset (one partition). MP heuristic searches were performed in PAUP* v4b10 (Swofford, 2002) using 100 random addition (RA) replicates, a maxtree of 10,000 trees per replicate, and tree-bisection-reconnection (TBR). ML analysis of the concatenated dataset (6 partitions) was performed in RAxML (Stamatakis, 2006) using 1000 RA. Modeltest 3.06 (Posada and Crandall, 1998) was used to select the appropriate models of evolution for each gene partition under the Akaike Information Criterion AIC (Posada and Buckley, 2004). The general time reversible model of evolution, with proportion of invariable sites and gamma distribution was selected for each data partition. Clade support under the MP and ML approaches was assessed using the non-parametric bootstrap procedure (Felsenstein, 1985) with 1000 bootstrap replicates and one RA per replicate. The concatenated dataset (6 partitions) was also analysed using Bayesian methods coupled with Markov chain Monte Carlo (BMCMC) inference as implemented in MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003). Four independent BMCMC analyses were run with each consisting of four chains. Each Markov chain was started from a random tree and run for 10^7 cycles, sampling every 1000th generation. Model parameters were unlinked and treated as unknown variables with uniform default priors and they were esti-

mated as part of the analysis. Convergence and mixing were monitored using Tracer v1.4 (Rambaut and Drummond, 2003). All sample points prior to reaching stationary were discarded as burn-in. The posterior probabilities (pP) for individual clades obtained from separate analyses were compared for congruence and then combined and summarized on a 50% majority-rule consensus tree (Huelsenbeck and Imennov, 2002; Huelsenbeck et al., 2002).

Confidence in our best hypotheses of phylogenetic relationships were tested by first creating alternative hypotheses in MacClade as indicated in Pérez-Losada et al. (2004) and then comparing them under both likelihood and Bayesian frameworks. Likelihood topological tests were conducted using the Shimodaira and Hasegawa (S–H) (1999) test as implemented in PAUP*. Ten thousand replicates were performed for every topology test resampling the partial likelihoods for each site (RELL model). Bayesian topological tests were performed as described in Huelsenbeck et al. (2002).

Several methods for empirically testing species boundaries have been proposed and compared (Sites and Marshall, 2003, 2004; Marshall et al., 2006; Pons et al., 2006; Sei and Porter, 2007). Here we used an integrative approach of species delimitation that takes into account multiple lines of evidence by combining phylogenetic relatedness with other factors like shared morphological, chromosomal, and ecological characters, and genomic evidence. This general integrative approach has been reviewed and argued for by several researchers (Will et al., 2005; Rissler and Apodaca, 2007; Bond and Stockman, 2008) and explicitly applied in various forms

Table 1
Taxa sampled, codes, localities, and GPS coordinates. The code indicates the taxa name, the locality and the haplotypes. Locality # are showed in Fig. 1.

Taxon	Code	Locality #	Locality	Coordinates
A. caliginosa	Aca.FrAd.1,2,3	1	France (Adé)	43°07′55.8″ N 00°02′15.4″ W
	Aca.FrAgn.1,2	2	France (Antignac)	42°49′21.4″ N 00°36′16.6″ E
	Aca.FrCcg.1	3	France-Corsica (Bains-de-Taccana)	41°50′02.4″ N 08°57′45.9″ E
	Aca.FrPar.1,2	4	France (Paris)	48°43′14.91″ N 02°57′15.42″ E
	Aca.FrSd.1,2	5	France (Soudan)	46°25′11.5″ N 00°04′09.2″ W
	Aca.FlJk.1	6	Finland (Jokioinen)	60°48′02.82″ N 23°27′39.77″ E
	Aca.GGoe.1,2	7	Germany (Goettingen)	51°11′40.26″ N 10°16′23.02″ E
	Aca.SpEch.1	8	Spain-Navarra (Echarri)	42°46′02.8″ N 01°49′56.9″ W
	Aca.SpQr.1	9	Spain-Navarra (Quinto Real)	43°05′46.4″ N 01°31′46.1″ W
	Aca.SpBb.1,2	10	Spain-Bilbao	43°17′54.1″ N 03°02′32.3″ W
	Aca.SpOu.1,2	11	Spain-Ourense	42°07′55.26″ N 08°03′04.59″ W
A. tuberculata	Atu.DkSk.1,2	12	Denmark (Silkeborg)	56°12.25′N 09°30′W
	Atu.FlJk.1,2	6	Finland (Jokioinen)	60°48′02.82″ N 23°27′39.77″ E
	Atu.FlJy.1,2	13	Finland (Jyväskylä)	62°14′44.75″ N 25°41′27.49″ E
	Atu.PlZm.1	14	Poland (Lomianki)	52° 20′ N 20° 53′ E
	Atu.UkLc.1	15	United Kingdom (Lancaster)	54°02′N 02°45′W
A. trapezoides	Atr.FrAd.1,2	1	France (Adé)	N 43°09′59.4″ W 00°0′20.6″
	Atr.FrMsg.1,2	16	France (Monsegur)	44° 39′19.27″ N 0°4′50.54″ E
	Atr.FrSd.1	5	France (Soudan)	46°25′11.5″ N 00°04′09.2″ W
	Atr.PlZm.1	14	Poland (Lomianki)	52° 20′ N 20° 53′ E
	Atr.SbKg.1	17	Serbia (Kragujevac)	44°00′N 20°59′E
	Atr.SpLg.1,2	18	Spain (Lugo)	43°11′34.2″ N 07°13′46.2″ W
	Atr.SpMc.1,2	19	Spain-Navarra (Murchante)	42°01′33.1″N 001°39′22.9″W
	Atr.SpBb.1	10	Spain-Bilbao	43°17′54.1″ N 03°02′32.3″ W
	Atr.SpOu.1	11	Spain (Ourense)	42°08′13.5″ N 08°02′52.5″ W
	Atr.SpTld.1	20	Spain (Toledo)	39°51′23.36″N 04°06′21.42″W
	Atr.SpVg.1	21	Spain (Vigo)	42°10′01.92″ N 08°41′03.51″ W
	Atr.SpVt.1,2,3	22	Spain (Vitoria)	42°55′35.3″ N 02°43′46.0″W
A. nocturna	Ano.FrAvg.1,2	23	France (Avignon)	43°54′43.6″ N 004°53′07.7′ E′
	Ano.SpVg.1,2,3,4	21	Spain (Vigo)	42°10′01.92″ N 08°41′03.51″ W
A. longa	Alo.FrMny.1	24	France (Marnay)	46°23′51.0″ N 0°21′47.2″ E
	Alo.FrVrr.1	25	France (Verrieres)	46° 23′51.0″ N 0°04′09.2″ W
	Alo.FrSd.1,2,3,4	5	France (Soudan)	46°25′11.5″ N 00°04′09.2″ W
	Alo.FrPar.1	4	France (Paris)	48°43′14.91″ N 02°57′15.42″ E
	Alo.SpCbr.1,2	26	Spain (Cantabria)	43°23′37.0″ N 04°01′0.1″ W
	Alo.UkLc.1	15	United Kingdom (Lancaster)	54°02′N 02°45′W

Table 1 (continued)

Taxon	Code	Locality #	Locality	Coordinates
A. limicola	Ali.UkLc.1,2	15	United Kingdom (Lancaster)	54°02′N 02°45′W
A. rosea	Aro.SpVg.1	21	Spain (Vigo)	42°09′53.48″ N 08°40′56.25″ W
A. molleri	Amo.SpOu.1	11	Spain (Ourense)	42°08′13.5″ N 08°02′52.5″ W
Aporrectodea sp1	Apsp1.FrCcg.1	27	France-Corsica (Zonza)	41°44′20.9″N 09°09′36.3″W
Aporrectodea sp2	Apsp2.FrCcg.2,3	27	France-Corsica (Zonza)	41°44′20.9″N 09°09′36.3″W

Table 2
Primer sequences, length (bp) of the amplified gene regions, and position of the mDNA genes relative to the *Lumbricus terrestris* mtDNA genome and *Eisenia fetida* 28S gene.

Primer sequences	Length (bp)	Genetic position
tRNA-Asn-COII-tRNA-Asp: LumbF1: 5′-GGC ACC TAT TTG TTA ATT AGG-3′ tRNA-Asn-COII-tRNA-Asp: LumbR2: 5′-GTG AGG CAT AGA AAT ACA CC-3′	tRNA-Asn: 27 COII: 686 tRNA-Asp: 58	1556–1576 2339–2358
12S-tRNA-Val-16S-LumbF1: 5′-CTT AAA GAT TTT GGC GGT GTC-3′ 12S-tRNA-Val-16S-LumbR1: 5′-CCT TTG CAC GGT TAG GAT AC-3′	12S: 362 tRNA-Val: 67 16S: 713	10586–10603 11699–11718
12S-tRNA-Val-16S-LumbF4: 5′-CAG CTT GTG TAC TGC CGT CGT AAG-3′ 12S-tRNA-Val-16S-LumbR2: 5′-GCA ATG TTT TTG TTA AAC AGT CG-3′	12S: 271 tRNA-Val: 67 16S: 626	10672–10695 11620–11642
16S-tRNA-Leu-Ala-Ser-Leu-LumbF2: 5′-CGA CTG TTT AAC AAA AAC ATT GC-3′ 16S-tRNA-Leu-Ala-Ser-Leu-LumbR2: 5′-GTT TAA ACC TGT GGC ACT ATT C-3′	16S: 649 tRNA Leu-Ala-Ser-Leu: 220	11620–11642 12469–12490
tRNA-Leu-ND1-LumbF2: 5′-GAA TAG TGC CAC AGG TTT AAA C-3′ tRNA-Leu-ND1-LumbR1b: 5′-TTA ACG TCA TCA GAG TTA TC-3′	tRNA-Leu: 30 ND1: 917	12469–12490 13468–13487
28s-RD3.3f: 5′-GAA GAG AGA GTT CAA GAG TAC G-3′ 28s-rD5b: 5′-CCA CAG CGC CAG TTC TGC TTA C-3′	952	280–301 1240–1261
28S-F1: 5′-GAG TAC GTG AAA CCG TCT AG-3′ 28S-R1: 5′-CGT TTC GTC CCC AAG GCC TC-3′	809	295–314 1125–1144

by several others to date (Wiens and Penkrot, 2002; Dettman et al., 2003; Marshall et al., 2006; Sei and Porter, 2007; Stockman and Bond, 2007; Bond and Stockman, 2008).

3. Results

Our phylogenetic analyses showed no major disagreements among the MP (Fig. 2) and ML (Fig. 3) and Bayesian (Fig. 4) topologies. The few topological differences observed were mainly the result of a larger number of polytomies in the MP analysis. Three specimens from the Island of Corsica (France), which morphologically could not be identified and were regarded as *Aporrectodea* sp1 and *Aporrectodea* sp2, fell within the outgroup, showed high genetic divergence among them (as indicated by their branch lengths) and formed a sister clade to *A. limicola*. This suggests the presence of two hitherto unrecognized earthworm species in this island. To the contrary, the *A. longa* samples, which were initially selected as part of the outgroup, formed two paraphyletic clades with variable support that fell within the *A. caliginosa* species complex, sister related to *A. nocturna*. Monophyly of *A. longa* was rejected by the S–H test ($P = 0.0043$) and presented a pP < 0.001. All the samples from the putative *A. caliginosa* species complex (including *A. longa*) clustered together [bootstrap proportion (bp) = 96–100 and pP = 1.0] into two deep sister clades, one composed of *A. caliginosa s. s.* and *A. tuberculata* and another of *A. trapezoides*, *A. longa*, and *A. nocturna*. These two assemblages were also supported by high bp (89–100) and pP (0.99–1.0) values. Within the *A. caliginosa* species complex, the *A. tuberculata*, *A. caliginosa s.s.*, and *A. nocturna* samples formed monophyletic clades, but the *A. trapezoides* samples formed two paraphyletic ones. Monophyly of *A. trapezoides* was

not rejected by the S–H test ($P = 0.171$), although it presented a pP of 0.036 (i.e., monophyly is rejected). All ingroup subclades corresponding to different morphological species showed large genetic differences among them. No interdigitation of haplotypes was observed among putative *Aporrectodea* species, despite the fact that all of them shared sampling localities (i.e., sympatry). Therefore, in groups that fail to form monophyletic clusters we may have a lack of molecular evidence supporting these clusters but this is not the same as evidence supporting alternative clusters and may just be an issue of marker resolution.

Our ML phylogenetic tree (Fig. 3) showed deep intraspecific structuring within two taxa belonging to the *A. caliginosa* species complex *s.s.* (*A. caliginosa s.s.* and *A. trapezoides*) and *A. longa*. The eleven samples of *A. caliginosa s.s.* were grouped into two main sister clades. The twelve *A. trapezoides* samples were grouped into two main paraphyletic clades. A subclade of this taxon was composed of genetically very similar specimens. Finally, the six *A. longa* samples were grouped into two main paraphyletic clades. This raises the possibility of the existence of unrecognized species within these groups. within species clades, specimens from the same location (Aca.FrAd, Atr.FrAd, Atr.SpVt, and Alo.FrSd) fell in separated subclades; although this might indicate the presence of old lineages within those localities, given that these species are all peregrines (Blakemore, 2006) to some extent, human transport seems a more reasonable explanation to this diversity pattern.

4. Discussion

All of our MP, ML and BMCMC phylogenetic analyses based on 12 different mitochondrial and nuclear genes revealed two deep
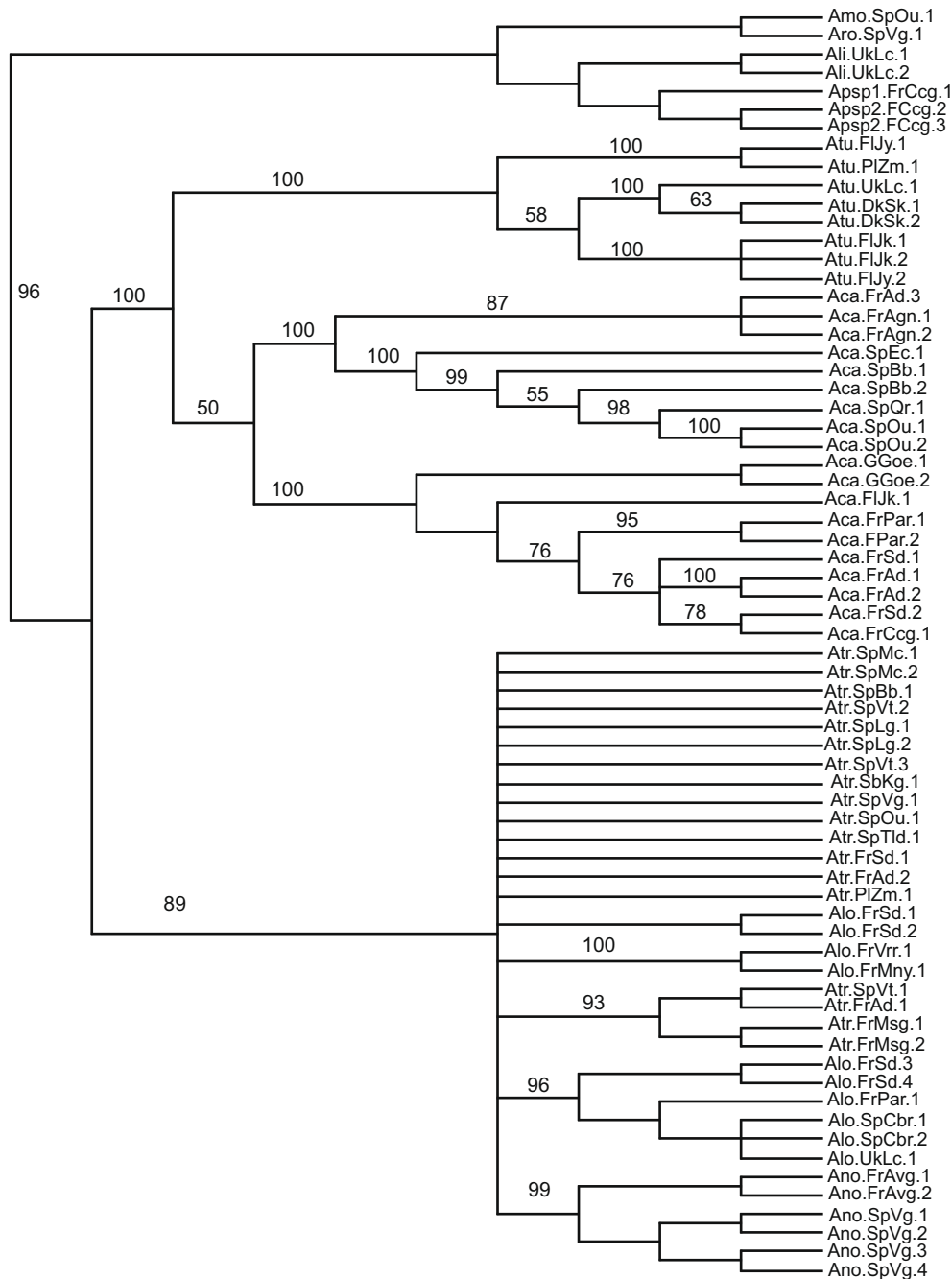
Fig. 2. Strict consensus tree of 10^6 most parsimonious trees ($L = 5365$). Bootstrap proportions (if $\geqslant 50\%$) are shown for each node.

sister clades, one composed of *A. caliginosa* (2 subclades) and *A. tuberculata*, and another composed of *A. nocturna*, *A. trapezoides* (2 subclades), and *A. longa* (2 subclades). *A. caliginosa*, *A. tuberculata*, and *A. nocturna* formed monophyletic assemblages, but *A. trapezoides* and *A. longa* resulted paraphyletic. As expected, the 28S gene was less variable than the mitochondrial genes due to its lower substitution rate. Hence the 28S ML tree showed less resolution at shallow level than any of the mitochondrial genes alone or combined. Nonetheless, both nuclear and mitochondrial trees showed the same basic assemblages described above. *A. caliginosa s.s.* and *A. tuberculata* sister relationship was weakly supported (bp $\leqslant 50\%$ and p$P < 0.6$) and *A. nocturna* was clustered (bp $> 70\%$ and p$P = 0.71$) with one of the *A. longa* clades. Our ML phylogenetic tree, however, showed deep phylogenetic structuring among those subclades, which is indicative of high (ancient) genetic divergence.

Generally accepted valid species such as *A. longa* presented levels of genetic divergence similar to those observed among *A. caliginosa* taxa. Moreover, no evidence of gene flow was observed between subclades despite the fact that many of these putative species occur in sympatry.

The integrative approach of species delimitation can greatly aid species identification (e.g., Yoder et al., 2005; Marshall et al., 2006; Sanders et al., 2006; Schlick-Steiner et al., 2006; Rissler and Apodaca, 2007; Roe and Sperling, 2007; Bond and Stockman, 2008). Several morphological, ecological, and genomic features support our phylogenetic assemblages. *A. caliginosa s.s.* and *A. tuberculata* (clade 1) have gray or light pigmentation, medium size and live in horizontal galleries on the soil (i.e., endogeic species) (Bouché, 1972). *A. trapezoides*, *A. longa*, and *A. nocturna* (clade 2) are characterized by a brown or dark pigmentation, larger size and live in vertical
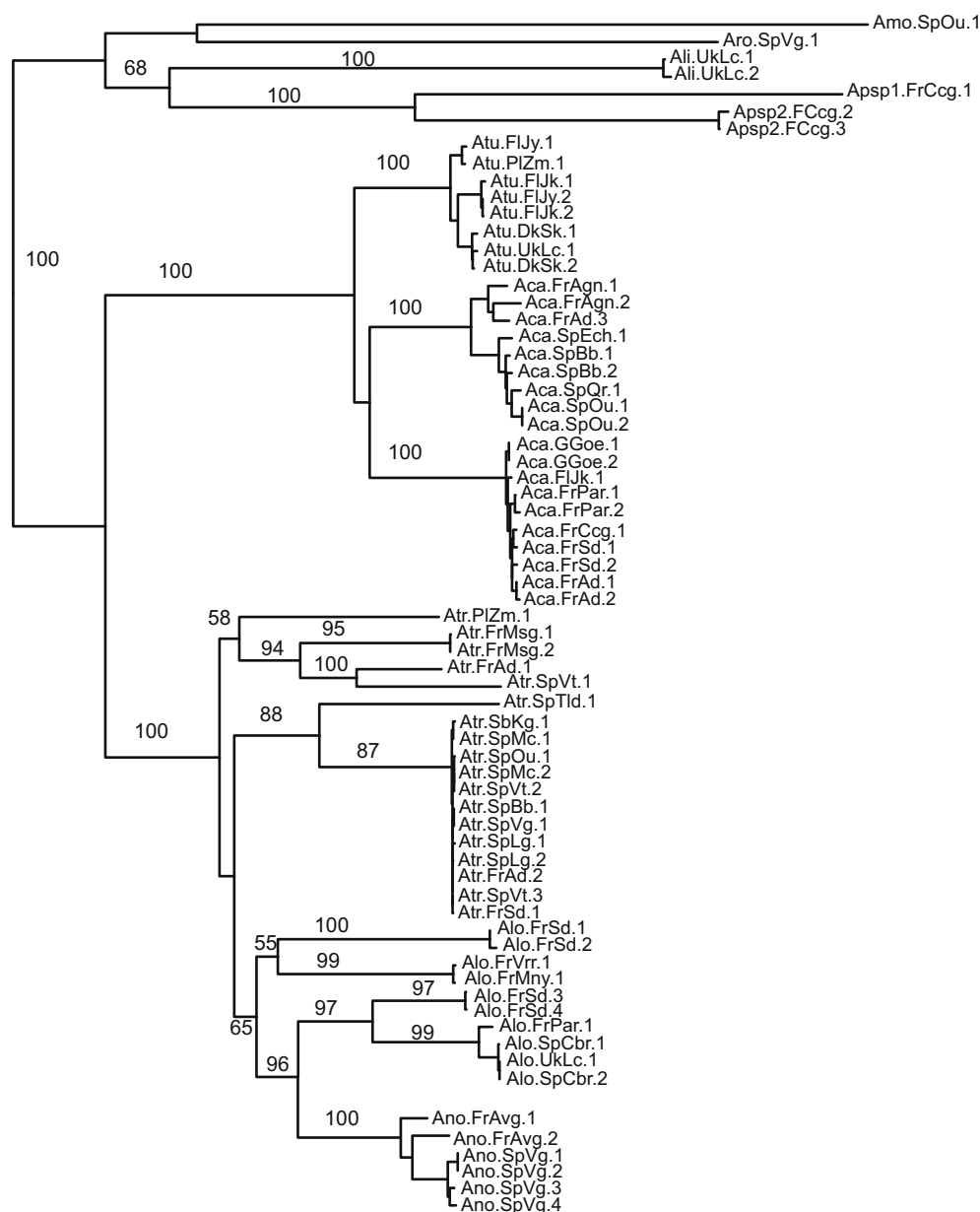
Fig. 3. Maximum likelihood mix-model tree. Branch lengths are shown proportional to the amount of change along the branches. Bootstrap proportions (if ⩾50%) are shown for each node.

galleries in the soil (i.e., anecic species) (Bouché, 1972). These differences in body size and ecology would explain the deep phylogenetic divergence observed in our trees between these two clades. Within clade 1, *A. tuberculata* can be separated from *A. caliginosa s.s.* based on the absence of genital tumescences in the segment number 33 in the former and its presence in the latter. This character is considered to be highly plastic, since the degree of development of the tumescences seems to reflect an increased sexual activity of the specimen (Sims and Gerard, 1985). However, the lack of genital tumescences in the segment number 33 in all *A. tuberculata* specimens remained constant in all the analyzed specimens.

The most obvious characteristic separating *A. trapezoides* from the other *Aporrectodea* species is its polyploid condition (Omodeo, 1952, 1955; Casellato, 1987; Sbordoni et al., 1987), which makes this earthworm the only polyploid taxon within the complex. Two *A. trapezoides* varieties have been described based on this genomic characteristic, a triploid variety and a tetraploid one

(Omodeo, 1952; Casellato and Rodighiero, 1972; Casellato, 1987), which are assumed to have arisen by parthenogenetic reproduction. In our ML tree (Fig. 3) we found evidence of this type of reproduction since there is a group of samples that are genetically very similar to each other despite the geographical distance among them. Besides, *A. trapezoides* has been regarded as male sterile (Gates, 1972) because of the presence in adult individuals of male organs retained in juvenile state, suggesting its parthenogenetic reproduction. However, this reduction of male structures has been reported to be very heterogeneous (Briones, 1996), as we have also found in the specimens analyzed. *A. trapezoides* would be then considered a paraphyletic species; however, species-level paraphyly is more common than often thought and Funk and Omland (2003) reviewed many such cases. Templeton (1998) argued that recognition of paraphyletic species is preferred over the alternative of elevating all monophyletic assemblages within to species and thus producing new species by "remote control".
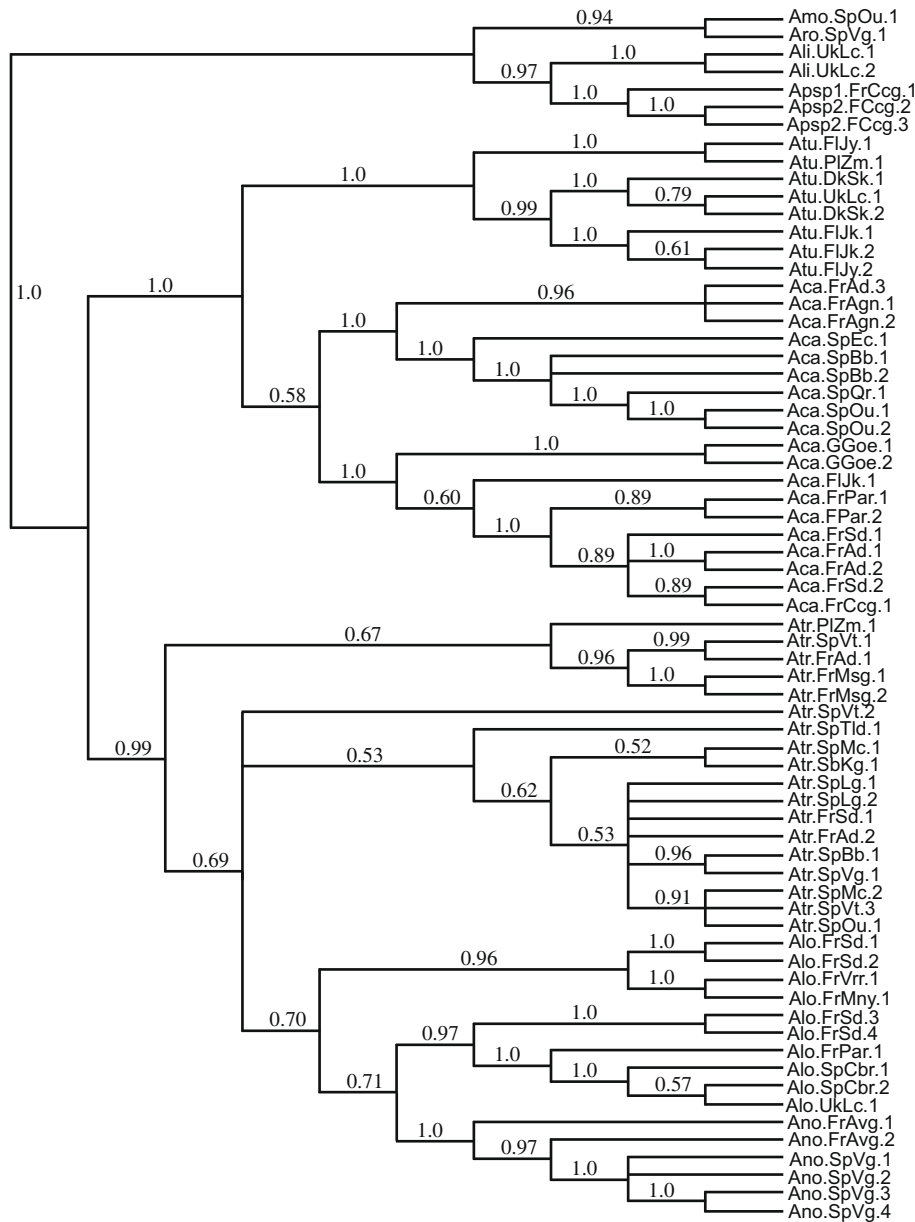
Fig. 4. Fifty percent majority-rule consensus BMCMC tree under mix-models. Clade posterior probabilities (if ⩾50%) are shown for each node.

The species status of *A. longa* has been widely accepted because of its morphological differences. The larger size of this earthworm as well as a somewhat flattened body, the position of the protuberant *clitellum* covering eight or nine segments and the *tubercula pubertatis* seen as band-like over segments 32–34, among other features, make *A. longa* different from other *Aporrectodea* species. *A. longa* is also taxonomically considered the closest species to *A. trapezoides* and *A. nocturna* (Gates, 1972; Blakemore, 2006). Our molecular trees seem to support this relationship.

Aporrectodea nocturna can be differentiated from *A. trapezoides* and *A. longa* based on the shape of the *clitellum*, which is cylindrical in the former and saddle-shaped in latter. Moreover, while *A. nocturna*, *A. caliginosa*, and *A. tuberculata* present *tubercula pubertatis* as two protuberances, *A. longa* and *A. trapezoides* have band-like ones (Gates, 1972; Sims and Gerard, 1985; Blakemore, 2006). Additionally, *A. nocturna* has reddish brown pigmentation and larger size in comparison to *A. caliginosa* and *A. tuberculata*, which lack pigmentation.

Finally, enzyme electrophoresis (Bøgh, 1992), karyotyping (Mezhzherin et al., 2008), and RAPDs (Dyer et al., 1998) also suggest that *A. caliginosa s.s.*, *A. tuberculata*, and *A. longa* (Bøgh, 1992) and *A. caliginosa s.s.*, *A. trapezoides*, and *A. longa* (Dyer et al., 1998), respectively, are genetically different species. Therefore, all these genetic, morphological, genomic, and ecological evidence suggests that *A. caliginosa s.s.*, *A. tuberculata*, *A. trapezoides*, *A. longa*, and *A. nocturna* constitute valid species. This interpretation agrees with some of the initial species descriptions and Gates (1972) proposal based on morphological evidence. Alternative proposals that suggest a lesser number of species or subspecies (Michaelsen, 1900; Omodeo, 1952; Gerard, 1964; Vedovini, 1969; Bouché, 1972; Casellato, 1987; Sims and Gerard, 1985; Sbordoni et al., 1987; Briones, 1996) are not supported by our analyses.

Our phylogenetic trees also revealed two hitherto unrecognized *Aporrectodea* species in Corsica, *Aporrectodea* sp1 and sp2. These taxa were morphologically very similar to *A. trapezoides* and *A. caliginosa s.s.*, but differed in two morphological features: (1) one of

the specimens (*Aporrectodea* sp1) lacks the spermathecae and the other two (*Aporrectodea* sp2) presented them between segments 12/13 and 13/14, whereas in the *A. caliginosa* complex the spermathecae constantly appear between segments 10/11 and 11/12; (2) *Aporrectodea* sp1 and sp2 present highly muscular septa between segments 6–11, whereas in the *A. caliginosa* complex the septa are moderately thickened. Nonetheless, the description of this new species is beyond the scope of this paper.

Our phylogenetic analyses also showed deep phylogenetic structuring within *A. caliginosa* s.s., *A. trapezoides*, and *A. longa*, where samples were clustered into two subclades each. Given the range of morphological variation in these species, we did not find discriminatory morphological features in any case between the specimens from different subclades, but the genetic divergence they present may indicate otherwise. A recent phylogenetic analysis (King et al., 2008) of mitochondrial COI and 16S genes from British earthworms has also reported very highly divergent lineages within *Aporrectodea longa*, *A. rosea*, *Allolobophora chlorotica* and *Lumbricus rubellus* and suggested the existence of multiple cryptic species within these taxa. Our results support this pattern, hence suggesting an *unprecedented* diversity within Lumbricidae earthworms (King et al., 2008).

5. Conclusions

Hence, how many species do constitute the *Aporrectodea caliginosa* species complex? Using an integrative approach to species delimitation (Templeton, 1989; Sites and Marshall, 2003, 2004; Will et al., 2005; Rissler and Apodaca, 2007; Bond and Stockman, 2008) this study suggests at least five valid species: *A. caliginosa* s.s., *A. tuberculata*, *A. nocturna*, *A. trapezoides* and *A. longa*. However, the possibility of new unrecognized subspecies or even species within these taxa is also raised. The taxonomic implications of this study are very important. *A. caliginosa* is the most abundant earthworm in grasslands from Paleartic regions and the most commonly found in agricultural ecosystems across the world. All future research on evolution, biogeography, ecology, conservation, and biodiversity and studies of more applied aspects (e.g., soil pollution and ecotoxicology) on these *Aporrectodea* taxa should be aware of their specific status and their biological differences. Finally our study also highlights the importance of using multiloci sequence data and phylogenetic analysis for delimiting earthworm species boundaries and assessing their evolutionary relationships (Pop, 2004).

References

Blakemore, R.J., 2006. Cosmopolitan Earthworms—An Eco-Taxonomic Guide to the Peregrine Species of the World, second ed. VermEcology, Japan. 600 pp.

Bøgh, P.S., 1992. Identification of earthworms (Lumbricidae): choice of method and distinction criteria. Megadrilogica 4, 163–174.

Bond, J.E., Stockman, A.K., 2008. An integrative method for delimiting cohesion species: finding the population-species interface in a group of California trapdoor spiders with extreme genetic divergence and geographical structuring. Syst. Biol. 57, 628–646.

Bouché, M.B., 1972. Lombriciens de France. Écologie et Systématique. INRA, Paris. 671 pp.

Briones, M.J.I., 1993. Two ecotypes in *Allolobophora caliginosa* (Oligochaeta, Lumbricidae). Acta Oecologica 14, 317–325.

Briones, M.J.I., 1996. A taxonomic study of the *Allolobophora caliginosa* complex (Oligochaeta, Lumbricidae): a preliminary study. Can. J. Zool. 74, 240–244.

Casellato, S., Rodighiero, R., 1972. Karyology of Lumbricidae. III Contribution. Caryologia 25, 513–524.

Casellato, S., 1987. On polyploidy in oligochaetes with particular reference to Lumbricidae. In: Pagliai, A.M.B., Omodeo, P. (Eds.), Proceedings of the International Symposium on Earthworms. Mucchi Editore, Modena, pp. 75–84.

Dettman, J.R., Jacobson, D.J., Turner, E., Pringle, A., Taylor, J.W., 2003. Reproductive isolation and phylogenetic divergence in *Nuerospora*. Evolution 57, 2721–2741.

Domínguez, J., Bohlen, P.J., Parmelee, R.W., 2004. Earthworms increase nitrogen leaching to greater soil depths in row crop agroecosystems. Ecosystems 7, 672–685.

Domínguez, J., Ferreiro, A., Velando, A., 2005. Are *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* Bouché, 1972 (Oligochaeta, Lumbricidae) different biological species? Pedobiologia 49, 81–87.

Dugés, A., 1828. Recherche sur la circulation, la respiration et al reproduction des Annélides sétigéres abranches. Ann. Sci. Nat. 15, 284–336.

Dyer, R., Fowler, J.C.S., Baker, G.H., 1998. Detecting genetic variability in exotic earthworms, *Aporrectodea* spp. (Lumbricidae), in Australian soils using RAPD markers. Soil Biol. Biochem. 30, 159–165.

Eisen, G., 1874. New Englands och Canadas Lumbricider. Ofvers. K. Vetensk. Akad. Forh. 30, 41–49.

Evans, A.C., 1946. A new species of earthworm of the genus *Allolobophora*. Annu. Mag. Nat. Hist. (Ser.11) 14, 98–101.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Funk, D.J., Omland, K.E., 2003. Specie-level paraphyly and polyphyly: frequency causes, and consequences, with insights from animal mitochondrial DNA. Ann. Rev. Ecol. Evol. Syst. 34, 397–423.

Gates, G.E., 1972. Contributions to North American earthworms (Annelida: Oligochaeta). No. 3: towards a revision of the earthworm family Lumbricidae IV. The *trapezoides* species group. Bull. Tall Timbers Res. Station 12, 1–146.

Gerard, B.M., 1964. Synopses of the British fauna: (6) Lumbricidae. Linnean Society of London, pp. 58.

Guindon, S., Gascuel, O., 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Huelsenbeck, J.P., Imennov, N.S., 2002. Geographic origin of human mitochondrial DNA: accommodating phylogenetic uncertainty and model comparison. Syst. Biol. 51, 155–165.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51, 673–688.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

King, R.A., Tibble, A.L., Symondson, W.O., 2008. Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricide earthworms. Mol. Ecol. 17, 4684–4698.

Marshall, J.C., Arévalo, E., Benavides, E., Sites, J., Sites Jr., J.W., 2006. Delimiting species: Comparing methods for Mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. Evolution 60, 1050–1065.

Mezhzherin, S.V., Vlasenko, R.P., Garbar, A.V., 2008. Genetic structure and peculiarities of earthworms *Aporrectodea* (superspecies) *caliginosa* (Oligochaeta: Lumbricidae) complex on the territory of Ukraine. Cytol. Genet. 42, 50–57.

Michaelsen, W., 1900. Das Tierreich 10: Vermes, Oligochaeta. Friedman & Sohn, Germany, pp. 715.

Mršić, N., 1991. Monograph on earthworms (Lumbricidae) of the Balkans. Academia scientiarum et atrium Slovenica, Historia naturalis, Ljubljana 31, 1–355; II: 356–757.

Omodeo, P., 1952. Cariologia dei Lumbricidae. Caryologia 4, 173–178.

Omodeo, P., 1955. Carioilogia dei Lumbricidae. II. Contributo. Caryologia 8, 135–178.

Omodeo, P., 1956. Contributo alla revisione dei Lumbricidae. Arch. Zool. Ital. 41, 129–212.

Perel, T.S., 1973. The shape of nephridial bladders as a taxonomic character in the systematics of Lumbricidae. Zoologischer Anzeiger 191, 310–317.

Perel, T.S., 1976. A critical analysis of the Lumbricidae genera system (with key to the USSR fauna genera). Rev. Ecol. Biol. Sol. 13, 635–643.

Pérez-Losada, M., Høeg, J.T., Crandall, K.A., 2004. Unraveling the evolutionary radiation of the Thoracican barnacles using molecular and morphological evidence. A comparison of several divergence time estimation approaches. Syst. Biol. 53, 244–264.

Pérez-Losada, M., Eiroa, J., Mato, S., Domínguez, J., 2005. Phylogenetic species delimitation of the earthworm *Eisenia fetida* (Savigny, 1826) and *Eisenia andrei* (Bouché, 1972) (Oligochaeta, Lumbricidae) based on mitochondrial and nuclear DNA genes. Pedobiologia 49, 317–324.

Piearce, T.G., Oates, K., Carruthers, W.J., 1990. A fossil earthworm embryo (Oligochaeta) from beneath a Late Bronze Age midden at Potterne, Wiltshire. UK J. Zool. 220, 537–542.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence based species

delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609.

Pop, A.A., Wink, M., Pop, V.V., 2003. Using of 18S, 16S rDNA and cytochrome *c* oxidase sequences in earthworm taxonomy (Oligochaeta, Lumbricidae). Pedobiologia 47, 428–433.

Pop, V.V., 2004. Towards a consistent and convenient genera system of the Lumbricidae family (Oligochaeta). A compromise between splitter and lumper viewpoints, adapted for the Romanian earthworm fauna. In: Moreno, A., Borges, S. (Eds.), Avances en taxonomía de lombrices de tierra/Advances in Earthworm Taxonomy (Annelida: Oligochaeta). Editorial Complutense, Madrid, pp. 317–333.

Pop, A.A., Csuzdi, Cs., Wink, M., Pop, V.V., 2006. An attempt to reconstruct the molecular phylogeny of the Genus *Allolobophora* (sensu lato, Pop, 1941) using 16S rDNA and COI sequences (Oligochaeta, Lumbricidae). In: Pop, V.V., Pop, A.A. (Eds.), Advances in Earthworm Taxonomy II. University Press, Cluj, pp. 155–166.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over Likelihood Ratio Tests. Syst. Biol. 53, 793–808.

Qiu, J.P., Bouché, M.B., 1998. Revision des taxons supraspécifiques de Lumbricoidea. Doc. Pedozoologiques et Integrologiques. Montpellier 3, 179–216.

Rambaut, A., Drummond, A.J., 2003. Tracer: MCMC trace analysis tool. Available from: <http://evolve.zoo.ox.ac.uk/>.

Rissler, L.J., Apodaca, J.J., 2007. Adding more ecology into species delimitation : ecological niche models and phylogeography help define cryptic species in the Black salamander (*Aneides flavipunctatus*). Syst. Biol. 56, 924–942.

Roe, A.D., Sperling, F.A.H., 2007. Population structure and species boundaries delimitation of cryptic *Dioryctria* moths : an integrative approach. Mol. Ecol. 16, 3617–3633.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rosa, D., 1893. Revisione dei Lumbricidi. Mem. R. Acc. Sci. Torino. 43, 399–476.

Sanders, K.L., Malhotra, A., Thorpe, R.S., 2006. Combining molecular, morphological, and ecological data to infer species boundaries in a cryptic tropical pitviper. Biol. J. Linn. Soc. 87, 343–364.

Savigny, J.C., 1826. Anylyse d'un Mémoir sur les Lombrics par Cuvier. Mem. Acad. Sci. Inst. Fr. 5, 176–184.

Sbordoni, M.C., De Matthaeis, E., Omodeo, P., Bidoli, R., Rodino, R., 1987. Allozyme variation and divergence between diploid and triploid population of *Allolobophora caliginosa* (Lumbricidae; Oligochaeta): particular reference to Lumbricids. In: Pagliai, A.M.B., Omodeo, P. (Eds.), Proceedings of the International Symposium on Earthworms. Mucchi Editore, Modena, pp. 53–74.

Schlick-Steiner, B.C., Steiner, F.M., Moder, K., Seifert, B., Sanetra, M., Dryerson, E., Stauffer, C., Christian, E., 2006. A multidisciplinary approach reveals cryptic

diversity in Western Paleartic *Tetramorium* ants (Hymenoptera: Formicidae). Mol. Phylogenet. Evol. 40, 259–273.

Sei, M., Porter, A.H., 2007. Delimiting species boundaries and the conservation genetics of the endangered maritime ringlet butterfly (*Coenonympha nipisiquit* McDunnough). Mol. Ecol. 16, 3313–3325.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of loglikelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Sims, R.W, Gerard, B.M., 1985. Earthworms, Keys and Notes for the Identification and Study of the Species. Synopses of the British Fauna (New Series) No. 31. Linnean Society of London and the Estuarine and Brackish-Water Sciences Association, London.

Sites, J.W., Marshall, J.C., 2003. Delimiting species: a renaissance issue in systematic biology. Trends Ecol. Evol. 18, 462–470.

Sites, J.W., Marshall, J.C., 2004. Empirical criteria for delimiting species. Ann. Rev. Ecol. Evol. Syst. 35, 199–227.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Stockman, A.K., Bond, J.E., 2007. Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. Mol. Ecol. 2007, 3374–3392.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods), Sinauer, Sunderland, MA.

Templeton, A.R., 1989. The meaning of species and speciation: a genetic perspective. In: Otte, D., Endler, J.A. (Eds.), Speciation and its Consequences. Sinauer Associates, Sunderland, MA.

Templeton, A.R., 1998. Species and speciation: geography, population structure, ecology and gene trees. In: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, New York, NY, pp. 32–41.

Vedovini, A., 1969. Contribution à l'étude des variations de l'espèce *Allolobophora caliginosa* (Lumbricidae). Bull. Soc. Zool. France 94, 657–662.

Wiens, J., 1998. Combining data sets with different phylogenetic histories. Syst. Biol. 47, 568–581.

Wiens, J.J., Penkrot, T.A., 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). Syst. Biol. 51, 69–91.

Will, K.W., Mishler, B.D., Wheeler, Q.D., 2005. The perils of DNA barcoding and the need for integrative taxonomy. Syst. Biol. 54, 844–851.

Yoder, A.D., Olson, L.E., Hanley, C., Heckman, K.L., Rasoloarison, R., Russel, A.L., Ranivo, J., Soarimalala, V., Karanth, K.P., Raselimanana, A.P., Goodman, S.M., 2005. A multidimensional approach for detecting species patterns in Malagasy vertebrates. Proc. Natl. Acad. Sci. USA 102, 6587–6594.

Zicsi, A., 1982. Verzeichniss der bis 1971 beschriben und revidierten Taxa der Familie Lumbricidae (Oligochaeta). Acta Zool. Hung. 28, 421–454.

Zicsi, A., 1991. Über die Regenwürmer Ungarns (Oligochaeta, Lumbricidae) mit Bestimmungstabellen der Arten. Opuscula Zool. 24, 167–191.