

Title: Phylogenetic analysis of oligochaete Tubificinae (Clitellata, Annelida) based on mitochondrial sequence data.

Authors: Ainara Achurra¹, M. Arantzazu Elejalde² and Pilar Rodriguez¹

Addresses:

1 Department of Zoology and Animal Cell Biology, Faculty of Science and Technology, University of the Basque Country UPV/EHU, Apdo. 644, Bilbao 48080, Spain.

2 Department of Invertebrates, Royal Belgian Institute of Natural Sciences. Vautierstraat 29, B-1000 Brussels, Belgium.

E-mails: ainara.achurra@ehu.es, pilar.rodriguez@ehu.es,
mirenarantzazu.elejalde@ehu.es

Abstract

Partial sequences of the mitochondrial genes 16S rRNA and cytochrome c oxidase subunit I (COI) were used (1) to resolve the internal relationships of the subfamily Tubificinae (Clitellata, Annelida); and (2) to test the existence of cryptic species within the stygobiont oligochaete *Troglodrilus galarzai*. Phylogenies were estimated using maximum likelihood, Bayesian inference and parsimony. Although trees were incompletely resolved at intergeneric level, a close relationship between *Heterochaeta* and *Troglodrilus* was revealed, and the taxonomic status of *Lophochaeta ignota* and *Heterochaeta costata* separated from *Tubifex* was corroborated by mitochondrial molecular data. Maximum genetic divergence between allopatric populations of *T. galarzai* was 18% for COI (uncorrected pairwise distance), suggesting cryptic speciation within this nominal species.

Introduction

Tubificine worms are clitellate annelids inhabiting superficial and subterranean freshwaters as well as brackish and marine waters. The diagnostic morphological characters of the subfamily Tubificinae Eisen, 1879 are: solid prostates (when present) with stalk-like attachments to atria; no obvious coelomocytes in the body cavity; and sperm in spermathecae usually (but not in all species) arranged in spermatzeugmata (Giani *et al.* 1984). Modified genital chaetae can occur, usually associated to the spermathecal pores but sometimes also to the penis. Currently about 32 genera are known within the subfamily Tubificinae, seven of which are monospecific and endemic to different regions of the world. Brinkhurst (1991) studied the phylogeny of the subfamily using 17 morphological characters and 29 tubificine genera. Many of the taxa remained unresolved, but several lineages were defined by a series of changes in the atrium, the vas deferens, the penis structure and the genital chaetae. Specific molecular analyses of the subfamily Tubificinae have not been conducted yet despite new molecular tools have been used effectively in assessing the phylogeny of several clitellate taxa (e.g. Martin *et al.* 2010 and Zhou *et al.* 2010; previously reviewed by Halanych and Janosik 2006). This article provides the first molecular phylogenetic study of the subfamily Tubificinae, using partial sequences of the mitochondrial genes cytochrome c oxidase subunit I (COI) and 16S rRNA. A special focus on understanding the taxonomic status of the genera *Lophochaeta* Štolc, 1886 and *Heterochaeta* Claparède, 1863 is given, since both have been regarded as synonyms of *Tubifex* Lamarck, 1816 by some authors (Brinkhurst and Jamieson 1971) but as distinct taxa by others (Holmquist 1985).

Troglodrilus galarzai is a stygobiont (i.e. groundwater limited) species in a monospecific genus. Only four populations (all in southern Europe) are known: Santa Eufemia-Ereñozar and Gorbeia karstic units in the northern Iberian Peninsula, and the gallery of Montgelas and Crotot cave in France. Juget *et al.* (2006) examined among-population variability of several morphological characters in the species. Differences, although suggested, were not statistically significant and measurements overlapped. Here, potential cryptic speciation within *T. galarzai* is tested through assessing the genetic divergence for COI and 16S rRNA sequences between Spanish and French populations.

Materials and methods

Taxa, sampling and collections

In present study, the ingroup tubificine taxa (Tab. 1) represents 3 of the 4 main lineages in the morphology-based phylogeny estimated by Brinkhurst (1991). As outgroups, *Rhyacodrilus okamikae* and 2 enchytraeid species (*Fridericia tuberosa* and *Bucholzia fallax*) were selected. DNA sequences of a total 21 tubificine species were studied. Specimens newly sequenced for this study were 6 *Troglodrilus galarzai*, 2 *Isochaetides gianii*, 2 *Lophochaeta ignota*, 2 *Varichaetadrilus bizkaensis*, 1 *Embolocephalus velutinus*, 1 *Tubifex tubifex* and 1 *Rhyacodrilus okamikae*. Three specimens of *T. galarzai* were collected in south eastern France and loaned to us by M. des Châteliers (University of Lyon). Remaining material was collected by P. Rodriguez and A. Achurra in caves and springs in the northern Iberian Peninsula.

Specimens were sorted from the sediment samples and mature worms were identified alive, killed in 30% ethanol and cut into two parts. The anterior part was fixed in 4% formaldehyde and the posterior part was preserved in 96% ethanol for molecular analyses. Anterior body fragments were further stained in Ehrlich's hematoxylin and dissected or whole-mounted in Canada balsam for morphological study. Anterior fragments of the sequenced individuals are deposited, as microscope slides, in the National Museum of Natural Sciences, Madrid, Spain (MNCN). One specimen of *T. tubifex* from a culture kept in the laboratory at the University of the Basque Country (UPV/EHU) was used as quality control for the gene sequencing procedure and the phylogenetic analyses. Several specimens from the same culture were used in a previous phylogenetic research by Crottini *et al.* (2008) (clade 2e; Genbank accession numbers: EU117525-EU117539).

DNA extraction, fragment amplification and sequencing

Individual DNA was extracted from the posterior section of the worm using the DNAeasy Tissue Kit (QIAGEN). Polymerase chain reaction (PCR) was used to amplify two mitochondrial gene fragments, cytochrome oxidase subunit I (COI) and 16S rRNA, using universal primers (Folmer *et al.* 1994 and Palumbi *et al.* 1991, respectively). The PCR conditions were: denaturation at 94°C for 5 min, annealing at 45-57°C for 30 s; elongation at 94°C for 30 s; total number of cycles 32 and 35 (for COI and for 16S rRNA, respectively). The cycling ended with an extension phase at 72°C for 7-8 min. Reaction products were run on 1.5% agarose gels and stained with ethidium bromide to

verify positive amplifications. Amplicons were sequenced on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems). New sequences were submitted to GenBank; accession numbers of these, as well as of the additional ones obtained for the phylogenetic analysis, are given in Table 1.

Sequence analyses

Two data sets were evaluated. For resolving the intergeneric relationships within the subfamily, we followed recommendation by Milinkovitch *et al.* (1996) and considered a few specimens of each genus (Tab. 1) in order to avoid redundant phylogenetic information. For assessing low-level phylogenetic relationships, including cryptic speciation in *T. galarzai*, as many specimens as possible were evaluated (all the specimens in Table 1).

Before phylogenetic analyses were run, DNA sequences were aligned using the web version of MUSCLE version 3.7 (Edgar 2004) on the European Bioinformatics Institute (EBI) server applying default settings. The resulted alignments were contrasted with those obtained using Clustal X version 1.8 (Thompson *et al.* 1997) and were manually refined. Saturation levels of the COI fragments were determined using the program DAMBE version 4.2.13 (Xia and Xie 2001). First, second and third codon positions were analysed separately. No mutational saturation was evident at first and second nucleotide positions but the third position showed saturation when genetic distances were over 6%. The need to include COI third position in the analyses, as they contain much of the phylogenetic structure in the data, has been corroborated by Källersjö *et al.* (1999) for the subfamily Phallodrilinae (Annelida, Clitellata). Therefore, analyses were conducted both excluding and including the third nucleotide codon position. Genetic distances among sequences [uncorrected pairwise distances (p) and Kimura 2-parameter (K2P) distances] were calculated for each data set (16S and COI) in PAUP version 4.0b10 (Swofford 2002), and considering all the sequences in Table 1. Bayesian analyses (BA) were performed using the MrBayes version 3.1.2 (Ronquist and Huelsenbeck 2003). Sequences were combined in a partitioned model with individual models for the different genes and COI codon positions. The best models of sequence evolution were selected using the Akaike Information Criterion (AIC) implemented in MrModeltest version 2.3 (Nylander 2004) in conjunction with PAUP version 4.0b10. The model selected for the 16S sequences was the general time reversible model (GTR) with a proportion of invariable sites (I) and a gamma-shaped rate variation across sites

(G). For the COI first position GTR+I was used; for the COI second position GTR+G; and for the COI third position the SYM model with a proportion of invariable sites (I) and a gamma-shaped rate variation across sites (G). The Markov Chain Monte Carlo (MCMC) search was run with four chains (one cold and three heated) for 10 million generations, with trees sampled every 100 generations (the first 25000 burn-in trees were discarded). We checked for stationarity and convergence of the chains with the software TRACER 1.3 (Rambaut and Drummond 2004). Posterior probabilities were estimated for branches based on the saved trees and plotted on the majority-rule consensus tree.

For the maximum parsimony (MP) analyses, gaps were treated as missing data. A heuristic search was performed, with 10 random addition replicates, using the tree bisection reconnection (TBR) option generating multiples trees to determine the most parsimonious one. Parsimony bootstrap support values were calculated through 1000 bootstrap replicates (10 addition sequence replicates per bootstrap replicate). The weight of transversions (Tv) and transitions (Ts) was varied depending on the fragment and was estimated by maximum likelihood. Weighting was 2:1 for the COI gene and 1:1 for the 16S rRNA gene.

We also constructed trees applying maximum likelihood (ML) method using the online version of RaxML BlackBox (Stamakis *et al.* 2008), with 100 bootstrap replicates by using the GTRGAMMA model. All model parameters were estimated by the program from its own maximum parsimony starting trees. The alignment was divided into the same four partitions as used for the Bayesian analysis.

Results

Phylogenetic analyses

The aligned 16S rRNA fragment consists of 495 base pairs (bp). The 26 sequences of the ingroup are A:T rich (68%), with nucleotide composition of T (28.8%), C (16.7%), A (39.2%) and G (15.2%). Parsimony-informative characters are 134 (26.9%). A clade formed by *Heterochaeta* + *Troglodrilus* is supported by the three analyses (BA posterior probabilities: PP = 1.00, bootstrap values under ML: BV = 88, bootstrap values under MP: BV = 86). *Isochaetides* is grouped as the sister taxon to the clade *Heterochaeta* + *Troglodrilus* in the BA tree (PP = 0.99) and the 3 genera are grouped together with *Lophochaeta* and *Limnodrilus* in a clade (PP = 0.88); neither ML nor MP analyses support these relationships. Another clade is formed by *Varichaetadrilus* +

Potamothrix + *Ilyodrilus* + *Tubificoides* + *Tubifex* (PP = 0.99, BV under ML = 57, BV under MP = 50). *Varichaetadrilus* and *Potamothrix* are grouped together (PP = 1.00, BV under ML = 64, BV under MP = 77) within this clade.

The aligned COI fragment comprises 658 bp. The 21 sequences of the ingroup data matrix are A:T rich (66%), with a nucleotide percent composition of T (29%), C (2%), A (37%) and G (13%). The COI gene shows 306 (47%) parsimony-informative characters (only 22% when excluding the third nucleotide position). The COI trees contain a basal ingroup polytomy, both when using all nucleotide sequence data and when excluding the third codon position, revealing that the COI analyses do not resolve the relationships among tubificine genera.

The combined data set of the COI and 16S rRNA fragments consists of 1153 bp. A total of 557 (47.9%) characters are constant, 391 (33.7%) variable characters are parsimony uninformative and 345 (29.7%) characters are parsimony informative. The combined analysis yield better-resolved consensus trees (Fig. 2). The split between *Clitellio* and the rest of the tubificine taxa is supported by the three analyses (PP = 1.00, both excluding and including the third codon position; BV under MP excluding the third codon position = 90; BV under ML including the third codon position = 93; BV under MP excluding the third codon position = 96). *Troglodrilus* and *Heterochaeta* form a strongly-supported monophyletic clade (Clade 1: PP = 1.00, both excluding and including the third codon position; BV under MP excluding the third codon position = 98; BV under ML excluding/including the third position = 98/100). *Lophochaeta* + *Isochaetides* + *Limnodrilus* form a second clade (Clade 2: PP = 1.00, both including and excluding the third codon position; BV under ML including the third position = 74). The remaining genera (*Varichaetadrilus* + *Tubifex* + *Ilyodrilus* + *Tubificoides*) are grouped together (Clade 3: PP excluding the third codon position = 0.84, PP including the third codon position = 0.89; BV under ML including the third position = 64).

Genetic variation and intraspecific relationships

Uncorrected pairwise (p) distances and Kimura 2-parameter (K2P) distances between taxa for COI and 16S DNA sequences are summarized in Table 2. When considering all specimens in Table 1, 16S-based trees show a similar topology to those estimated with fewer specimens (Appendix 1), while COI-based trees show an unresolved basal polytomy (trees not shown). Populations of *Troglodrilus galarzai* from France (3 specimens) and Iberian Peninsula (3 specimens) have different exclusive haplotypes.

Uncorrected “p” distance between both populations is 17.8-18.1% for COI (20-21% K2P distance) and 9.2-10% for 16S (10-11% K2P distance). Both populations are clustered monophyletically as sister groups with maximum or very high support (Appendix 1). Uncorrected maximum “p” distance between populations of *Lophochaeta ignota* from Iberian Peninsula and northern Europe is 3% for COI (3% K2P) and 1% for 16S (1% K2P). Uncorrected “p” distance between *L. ignota* and *Tubifex tubifex* is 18-24% for COI (23-27% K2P) and 17.9-21.5% for 16S (21-26% K2P). Uncorrected “p” distance between *T. tubifex* and *Heterochaeta costata* is 20-22% for COI (23-26% K2P) and 18-21% for 16S (21-25% K2P). Uncorrected “p” distance between the various populations of *T. tubifex* is 10.7-14.3% for 16S (7-16% K2P); maximum uncorrected “p” distance for COI is 23% (26% K2P). The 16S haplotype of the new sequenced specimen of *T. tubifex* is identical to the haplotype 13a in Crottini *et al.* (2008), both from the same population of *T. tubifex* cultured at the laboratory of the University of the Basque Country. The clades for *T. tubifex* in the 16S-based trees in present study are consistent with those obtained by Crottini *et al.* (2008), except for clade 4 (Fig. 2 in Crottini *et al.* 2008), which resulted in close relationship with *Varichaetadrilus bizkaiensis*. We re-examined the histological sections of 5 specimens corresponding to clade 4 (Genbank accession numbers EU117505-EU117509) and identified them as *Potamothrix bavaricus*. The representatives of the genus *Limnodrilus* group together in a clade according to 16S data, but the species *Limnodrilus hoffmeisteri* appears to be paraphyletic (Appendix 1; uncorrected “p” distance 5-12%, 5-13% K2P).

Discussion

High-level phylogenetic relationships within Tubificinae

Trees based on mitochondrial molecular data do not resolve intergeneric or intrasubfamilial relationships with the exception of Clade 1 (*Troglo-drilus* + *Heterochaeta*). Both genera into this clade share a thick cuticular penial sheath; penis within a penial sac with a circular fold; and a moderately long vas deferens entering the atrium apically. Our results suggest either a freshwater origin for *Heterochaeta* or a littoral marine or brackish water ancestor for *Troglo-drilus*. The known hypothesis of stygofauna colonizing subterranean freshwaters from littoral marine habitats was first proposed for oligochaetes by Giani and Rodriguez (1988) referred to stygobiont phallo-drilines. Since then, other authors have also hypothesized a marine ancestor for several subterranean taxa (Sambugar *et al.* 1999; Pinder *et al.* 2006; Des Chatelliers *et*

al. 2009). The available data on the origin of tubificine genera are scarce and phylogenetic relationships within the subfamily still unresolved. However, we provide the first molecular evidence for a close relationship between stygobiont and estuarine oligochaetes.

Clade 2 (Fig. 2) consists of three taxa which share long vasa deferentia and well-developed penial sacs: *Lophochaeta ignota*, a freshwater species with a Palearctic distribution; *Limnodrilus*, a cosmopolitan freshwater genus; and *Isochaetides gianii*, the only stygobiont species known in the genus, which is mainly restricted to the Baikalian region. On the contrary, the 16S analysis (Fig. 1) shows *Isochaetides* as the sister taxon to *Heterochaeta* + *Troglodrilus*.

A third clade (Clade 3: *Varichaetadrilus* + *Tubifex* + *Ilyodrilus* + *Tubificoides*) is suggested in the combined analysis (Fig. 2). The genera *Varichaetadrilus*, *Ilyodrilus* and *Tubificoides* possess tubular atria with histologically differentiated sections along the longitudinal (proximal-distal) axis (Holmquist 1985; Rodriguez and Giani 1984: Fig. 2; Erséus 1989). In contrast, the atrium has no histological differentiation along this axis in *Tubifex*. The 16S analysis (Fig. 1) appears to be concordant with the histology of the atrium since it supports a clade formed by *Varichaetadrilus* + *Potamothrix* + *Ilyodrilus* + *Tubificoides*, excluding *Tubifex*. The genus *Potamothrix* also shows a longitudinal differentiation of the atrial histology, forming several sections along this axis [bipartite or tripartite atrium, as defined by Finogenova and Poddubnaja (1990)]. Other tubificine genera, such as *Troglodrilus* or *Lamadrilus*, have a lateral histological differentiation between the concave and convex sides of the atrium, apparently associated to the prostatic gland junction. At this point, the molecular phylogeny is still too unstable to identify the true intergeneric relationships in clades 2 and 3 and a broader gene and taxon sampling should resolve this issue.

Our study supports *Clitellio arenarius* as the sister species to all other ingroup tubificine taxa (Fig. 2), which suggests a singular taxonomic status for the species. Its phylogenetic position has been extensively discussed (Brinkhurst and Jamieson 1971; Erséus 1990; Gustavsson 1995) and at present, *Clitellio* is considered to be a tubificine genus based on the presence of penes and spermatozuogmata, as well as on the sperm ultrastructure (Ferraguti and Ruprecht 1992; Marotta *et al.* 2008).

Low-level phylogenetic relationships within Tubificinae

Mitochondrial data highly supports the recognition of *Lophochaeta ignota* apart from *Tubifex* (Figs 1 and 2), which is in accordance with some earlier investigations based on morphological characters (Holmquist 1985; Rodriguez and Achurra 2011). The species has a largely Palearctic distribution (although also reported in the Great Lakes of North America and Lake Titicaca), inhabiting superficial and subterranean freshwater habitats (Rodriguez and Achurra 2011). Interestingly, the genetic distances between the populations of Cork (Ireland) (Beauchamp *et al.* 2001) and Biscay (Spain) of *L. ignota* were low (Tab. 2) although about 1,500 km separate these regions. Our analyses (Figs 1 and 2) also show that *Tubifex* and *Heterochaeta* are well-separated genera and support the recognition of the genus *Heterochaeta*, which was based on differences in the male duct structure with other tubificines, particularly *Tubifex* (Holmquist 1985).

The monophyly of *T. tubifex* + *T. blanchardi* within the subfamily, as reported by Crottini *et al.* (2008), is confirmed here with the addition of other genera (Appendix 1). The identification of clade 4 of Crottini *et al.* (2008) as *P. bavaricus* reduces the number of *T. tubifex* clades delimited in their study from five to four. However, a high genetic distance between clades is still retained (Tab. 2), suggesting the existence of cryptic speciation, as pointed by those authors.

The clade formed by *Varichaetadrilus bizkaiensis* and *Potamothrix bavaricus* in the 16S-based tree (Fig. 1) has much support in morphology. Both species share long tubular atria with differentiated histological sections, short thin-walled vasa deferentia entering the atrium apically, penis in small penial sacs, and spermathecal ampullae with short ducts opening in lateral line. Other common characters are smooth hair and pectinate chaetae in dorsal chaetae bundles.

Limnodrilus hoffmeisteri is resolved as a paraphyletic taxon with respect to *L. cervix* (Appendix 1), while they were sister species in the study by Beauchamp *et al.* (2001). This contradictory result may be due to the fact that these authors only studied North American populations and our analysis includes also European specimens. A different argument may be related to the identification of *Limnodrilus* species. The taxonomic characters currently used for this purpose are primarily the shape of the chaetae and the length, width and shape of the cuticular penial sheath. Several species are known to have a high intraspecific variability of these characters (Kennedy 1969; Dzwillo 1984). Thus, species limits may be blurred. Further molecular analyses

including also other common *Limnodrilus* species (e.g. *L. profundicola*, *L. claparedeianus*) should give insight into the species boundaries within this genus.

Cryptic speciation in *Troglo-drilus galarzai*

French and Spanish populations of *T. galarzai* are at the moment morphologically impossible to differentiate following Juget *et al.* (2006). Mutual exclusivity of haplotypes, as well as mutual monophyly and high genetic distance between populations (maximum uncorrected “p” distance 18%) reveal the existence of cryptic speciation in *T. galarzai*. Lefébure *et al.* (2006) suggested an interspecific distance threshold of 16% (patristic distances) for COI in crustaceans, but the limit seems to vary according to the studied group (Shih *et al.* 2009). Recently, Erséus and Gustafsson (2009) have suggested a distance about 10% or more between congeneric species in Clitellata and distances in the range of 5–23% have been reported by other authors (Bely and Wray 2004: 5-17% between naidine species within a genus; Erséus and Kvist 2007: 19.3-22.9% between *Tubificoides* species collected from the same area; Gustafsson *et al.* 2009: 17.71% between potential cryptic species in *Lumbriculus variegatus*; Martin *et al.* 2010: 10.6% between two *Rhyacodriloides* species). In addition, the geographical allopatry of the studied populations also supports the cryptic speciation. The authors have done a new detailed morphological study including the four known populations of *T. galarzai* that will be published elsewhere.

The subfamily Tubificinae: the state of the art

Our results are limited by the use of mitochondrial genes, and nuclear genes may reveal new relationships. The suitability of the COI gene for phylogenetic reconstructions has been questioned for Phallo-drilinae (Nylander *et al.* 1999). However, the relationships revealed by the COI within the subfamily Naidinae (Bely and Wray 2004) were totally congruent with those ones based on nuclear sequence data (Envall *et al.* 2006). A second limitation may be that the evaluated genera still represent about 40% of the total tubificine genera known to date. To summarize, the molecular data set analyzed in this study has enabled us to corroborate several previous taxonomic hypotheses, and to recognize a few new relationships. Large taxon sampling, other molecular markers and integration of morphological characters seem crucial for further resolution of phylogenetic relationships within the Tubificinae.

Acknowledgments

Special thanks to Benjamin Gomez for making possible to the first author the access to the laboratory in the Faculty of Pharmacy (UPV/EHU) to carry out the molecular work. We are grateful to Christer Erséus and two anonymous referees who, with their suggestions, helped to improve the original manuscript. Thanks for the critical reading and comments to José Galián (University of Murcia, Spain) and Steve Fend (USGS). We are also indebted to Michel des Chatelliers (University of Lyon) for his sending material of *Troglodrilus galarzai* for the molecular analysis, and to Marco Ferragutti and Angelica Crottini (University of Milan) for loaning us histological preparations of *Tubifex tubifex* from their collection for re-examination. The first author had a PhD grant by the UPV/EHU (Euskarazko tesia egiteko ikerketa pertsonala prestatzeko laguntzen deialdia 2007-2011). The molecular analyses have been financed by the Catedra UNESCO (project UNESCO 04/04) and the GIU09/31 project (UPV/EHU).

References

- Beauchamp, K. A., Kathman R. D., McDowell T. S., Hedrick R. P. (2001) Molecular phylogeny of tubificid oligochaete with special emphasis on *Tubifex tubifex* (Tubificidae). *Molecular Phylogenetics and Evolution* **19**, 216–224.
- Bely, A.E., Wray G.A. (2004) Molecular phylogeny of naidid worms (Annelida: Clitellata) based on cytochrome oxidase I. *Molecular Phylogenetics and Evolution* **30**, 50–63.
- Brinkhurst, R.O. (1991) A phylogenetic analysis of the Tubificinae (Annelida, Oligochaeta). *Canadian Journal of Zoology* **69**, 392–397.
- Brinkhurst, R.O., Jamieson B.G.M. (1971) ‘Aquatic Oligochaeta of the world.’ (University of Toronto Press, Toronto, ON, and Buffalo, NY).
- Crottini, A., Marotta R., Barbuto M., Casiraghi M., Ferraguti M. (2008) The world in a river? A preliminary analysis of the 16S rDNA variability of *Tubifex* species (Clitellata: Tubificidae) from the Lambro River. *Molecular Phylogenetics and Evolution* **48**, 1189–1203.
- Des Châtelliers, M.C., Juget J., Lafont M., Martin P. (2009) Subterranean aquatic Oligochaeta. *Freshwater Biology* **54**, 678–690.
- Dzwillo, M. (1984) Phylogenetic and taxonomic problems in freshwater Oligochaeta with special emphasis on chitinous structures in Tubificinae. *Hydrobiologia* **115**, 19–23.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797.
- Envall, I., Källersjö M., Erséus C. (2006) Molecular evidence for the non-monophyletic status of Naidinae (Annelida, Clitellata, Tubificidae). *Molecular Phylogenetics and Evolution* **40**, 570–584.

Erséus, C. (1989) Four new West Atlantic species of *Tubificoides* (Oligochaeta, Tubificidae). *Proceedings of the Biological Society of Washington* **102**, 878–886.

Erséus, C. (1990) Cladistic analysis of the subfamilies within the Tubificidae (Oligochaeta). *Zoologica Scripta* **19**, 57–63.

Erséus, C., Gustafsson D. (2009) Cryptic speciation in clitellate model organisms. In ‘Annelids in Modern Biology’. (Ed. Dan Shain) pp. 31–46 (Wiley–Blackwell: Hoboken, NJ).

Erséus, C., Kvist S. (2007) COI variation in Scandinavian marine species of *Tubificoides* (Annelida: Clitellata: Tubificidae). *Journal of the Marine Biological Association of the United Kingdom* **87**, 1121–1126.

Ferraguti, M., Ruprecht D. (1992) The double sperm line in *Clitellio arenarius* (Annelida, Oligochaeta, Tubificinae). *Italian Journal of Zoology* **59**, 349–362.

Finogenova, N.P., Poddubnaja T.L. (1990) One more revision of the genus *Potamothrix* Vejdovský et Mrázek, 1902 (Oligochaeta, Tubificidae). *Zoologische Jahrbücher. Abteilung für Systematik, Ökologie und Geographie der Tiere* **117**, 55–83.

Folmer, O., Black M., Hoeh W., Lutz R., Vrijenhoek R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.

Giani, N., Martinez-Ansemil E., Brinkhurst R.O. (1984) Revision du statut taxonomique des Aulodrilinae (Tubificidae, Oligochaeta). *Bulletin de la Société d'Histoire Naturelle de Toulouse* **120**, 17–22.

Giani, N., Rodriguez P. (1988) Description de quelques especes nouvelles de Tubificidae (Oligochaeta) de grottes et de sources karstiques de la peninsule iberique. *Stygologia* **4**, 121–137.

Gustavsson, L.M. (1995) Morphological investigation of the tubificid *Clitellio arenarius* (Annelida, Oligochaeta). *Sarsia* **80**, 159–171.

Gustafsson, D.R., Price D.A., Erséus C. (2009) Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Molecular Phylogenetics and Evolution* **51**, 182–189.

Halanych, K.M., Janosik A.M. (2006) A review of molecular markers used for Annelid phylogenetics. *Integrative and Comparative Biology* **46**, 533–543.

Holmquist, C. (1985) A revision of the genera *Tubifex* Lamarck, *Ilyodrilus* Eisen, and *Potamothrix* Vejdovsky & Mrazek (Oligochaeta, Tubificidae) with extensions to some connected genera. *Zoologischer Jahrbücher, Abteilung für Systematik*, **112**, 311–366.

Juget, J., des Chatelliers M., Rodriguez P. (2006) *Troglodrilus* (Annelida, Oligochaeta, Tubificidae), a new genus from subterranean habitats in southwestern Europe. *Hydrobiologia* **564**, 7–17.

Källersjö, M., Albert V.A., Farris J.S. (1999) Homoplasy Increases Phylogenetic Structure. *Cladistics* **15**, 91–93.

Kennedy, C.R. (1969) The variability of some characters used for species recognition in the genus *Limnodrilus* Claparède (Oligochaeta: Tubificidae). *Journal of Natural History* **3**, 53–60.

Lefébure, T., Douady C.J., Gouy M., Gibert J. (2006) Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* **40**, 435–447.

Marotta, R., Ferraguti M., Erséus C., Gustavsson L.M. (2008) Combined-data phylogenetics and character evolution of Clitellata (Annelida) using 18S rDNA and morphology. *Zoological Journal of the Linnean Society* **154**, 1–26.

- Martin, P., Martinez-Ansemil E., Sambugar B. (2010) The Baikalian genus *Rhyacodriloides* in Europe: phylogenetic assessment of Rhyacodriloidinae subfam. N. within the Naididae (Annelida). *Zoologica Scripta* **39**, 462–482.
- Milinkovitch, M.C., LeDuc R.G., Adachi J., Farnir F., Georges M., Hasegawa M. (1996) Effects of Character Weighting and Species Sampling on Phylogeny Reconstruction: A Case Study Based on DNA Sequence Data in Cetaceans. *Genetics* **144**, 1817–1833.
- Nylander, J.A.A., Erséus C., Källersjö M. (1999) A test of monophyly of the gutless Phallodrilinae (Oligochaeta, Tubificidae) and the use of a 573 bp region of the mitochondrial cytochrome oxidase I gene in analysis of annelid phylogeny. *Zoologica Scripta* **28**, 305–313.
- Nylander, J.A.A. (2004) MrModeltest v2. Program distributed by the author. (Evolutionary Biology Centre, Uppsala University.)
- Palumbi, S.R., Martin A., Romano S., McMillan W.O., Stice L., Grabowski G. (1991) ‘A Simple Fool's Guide to PCR.’ (Honolulu: University of Hawaii Press.)
- Pinder, A.M., Eberhard S.M., Humphreys W.F. (2006) New phallodrilines (Annelida: Clitellata: Tubificidae) from Western Australian groundwater. *Zootaxa* **1304**, 31–48.
- Rambaut, A., Drummond A.J. (2004) Tracer. Version 1.3. University of Oxford.
- Rodriguez, P., Achurra A. (2011) New species of aquatic oligochaetes (Annelida: Clitellata) from groundwaters in karstic areas of northern Spain, with taxonomic remarks on *Lophochaeta ignota*. *Zootaxa* **2332**, 21–39.
- Rodriguez, P., Giani N. (1984) Description de *Varichaetadrilus bizkaiensis* n. sp.; nouvelle espèce de Tubificidae (Oligochaeta) d'Espagne. *Bulletin de la Société d'Histoire Naturelle de Toulouse* **120**, 127–130.

Ronquist, F., Huelsenbeck J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **20**, 289–290.

Sambugar, B., Giani N., Martínez-Ansemil E. (1999) Groundwater oligochaetes from southern-Europe. Tubificidae with marine phyletic affinities: new data with description of a new species, review and consideration on their origin. *Mémoires de Biospéologie* **XXVI**, 107–116.

Shih, H.T., Kamrani E., Davie P.J.F., Liu M.Y. (2009) Genetic evidence for the recognition of two fiddler crabs, *Uca iranica* and *U. albimana* (Crustacea: Brachyura: Ocypodidae), from the northwestern Indian Ocean, with notes on the *U. lactea* species-complex. *Hydrobiologia* **635**, 373–382.

Stamatakis, A., Hoover P., Rougemont J. (2008) A Rapid Bootstrap Algorithm for the RAxML Web-Servers. *Systematic Biology* **75**, 758–771.

Swofford, D.L. (2002) ‘PAUP. Phylogenetic Analysis Using Parsimony Version 4.’ (Sinauer Associates: Sunderland, Massachusetts.)

Thompson, J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876–4882.

Xia, X., Xie Z. (2001) DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity* **92**, 371–373.

Zhou, H., Fend S.V., Gustafson D.L., De Wit P., Erséus C. (2010) Molecular phylogeny of Nearctic species of *Rhynchelmis* (Annelida). *Zoologica Scripta* **39**, 378–393.

Table 1. List of taxa included in the study, collection sites for newly sequenced specimens, GenBank accession number for the respective sequences and vouchers deposited in the National Museum of Natural Sciences, Madrid, Spain. GenBank numbers in bold denote new sequences. GenBank numbers marked with an asterisk refer to sequences used in the analysis of intergeneric relationships.

Taxon	Collection site	16S	COI	Voucher
Annelida, Clitellata, Tubificinae				
<i>Aulodrilus plurisetus</i> (Piguet, 1906)	Apraiz spring , Santa Eufemia-Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 524225, Y 4801350, Z 22	AJ225900*, AJ225899	AF054190*, EF675228	MNCN 16.03/3055
<i>Clitellio arenarius</i> (Müller, 1776)		AY885615*		
<i>Embolecephalus velutinus</i> (Grube, 1879)		HQ603823*		
<i>Heterochaeta costata</i> Claparède, 1863	Argatxa spring, Santa Eufemia-Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 527790, Y 4800925, Z 2	AY340460*	AF054189*	MNCN 16.03/3056, MNCN 16.03/3057
<i>Ilyodrilus templetoni</i> (Southern, 1909)		EF089341*	HQ616696* , HQ616697*	
<i>Isochaetides gianii</i> Rodriguez & Achurra, 2010		HQ603816* , HQ603817*		
<i>Limnodrilus cervix</i> Brinkhurst, 1963	Apraiz spring , Santa Eufemia-Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 524225, Y 4801350, Z 22	AF325984*, AF325983	EF089357*, EF089358*	MNCN 16.03/3058, MNCN 16.03/3059
<i>Limnodrilus hoffmeisteri</i> Claparède, 1862		AF325978, AF325980, AF325985, AY885613*, EU117546		
<i>Limnodrilus udekemianus</i> Claparede, 1862		AY885612*, AF325986		
<i>Lophochaeta ignota</i> Štolc, 1886	Apraiz spring , Santa Eufemia-Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 524225, Y 4801350, Z 22	HQ603818* , HQ603819	HQ616698* , HQ616699	MNCN 16.03/3058, MNCN 16.03/3059
northern Iberian Peninsula population				
northern Europe population		AF325988, AY885610*, AF325987	EF089360*	
<i>Potamothrix bavaricus</i> (Oeschmann, 1913)	Ubegi spring, Gorbeia karstic unit, Basque Country, Spain, UTM coordinates: X 0516270, Y 4766001, Z 973	EU117505*	HQ616690*	MNCN 16.03/3060, MNCN 16.03/3061,
<i>Troglodrilus galarzai</i> (Giani & Rodriguez, 1988)		HQ603810*		
northern Iberian Peninsula population				

	Artzegi spring, Gorbeia karstic unit, Basque Country, Spain, UTM coordinates: X 0520052, Y 4762908, Z 807	HQ603811* , HQ603812	HQ616691* , HQ616692	MNCN 16.03/3062
French population	Montgelas gallery, France (see Juget et al., 2006 for detailed description of the locality)	HQ603813* , HQ603814 , HQ603815*	HQ616693* , HQ616694 , HQ616695*	MNCN 16.03/3063, MNCN 16.03/3064, MNCN 16.03/3065
<i>Tubifex blanchardi</i> Vejdovsky, 1891 <i>Tubifex tubifex</i> (Müller, 1774) clade 2 in Crottini <i>et al.</i> 2008		EU117477* EU117502, AF426857, AF326038, EU117533, EU117530, EU117497 EU117492*, EU117488 EU117502		
clade 3 in Crottini <i>et al.</i> 2008 clade 5 in Crottini <i>et al.</i> 2008 northern Iberian Peninsula population	Culture from the University of the Basque Country, worms collected in a mountain river, Basque Country, Spain	HQ603822*	HQ616702*	MNCN 16.03/3066
other populations			EF089365*, EF089366, EF089375, EF089382, EF089376, EF179543, EF179544, EF089368, EF089377, EF089372, EF089373, EF089371, EF089374, EF089369, EF089380, EF089381, EF089370, EF089386, EF089385, EF089384	
<i>Tubifex smirnowi</i> Lastockin, 1927 <i>Tubificoides amplivasatus</i> (Erséus, 1975) <i>Tubificoides benedii</i> (Udekem, 1855)		AY885620* AY340483* AY885611*	EF675217, EF675219 EF675195, EF675199, EF675207, EF675194, EF675203, EF675193, EF675204*	
<i>Tubificoides bermudae</i> Râsmark and Erséus, 1986 <i>Tubificoides swirencowi</i> (Michaelsen, 1926) <i>Tubificoides kozloffii</i> Baker, 1983 <i>Varichaetadrilus bizkaiensis</i> Rodriguez &	Apraiz spring, Santa Eufemia-	AY885614* HQ603820* , HQ603821*	EF675224, EF675226* EF675227* HQ616700* ,	MNCN

Giani, 1984	Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 524225, Y 4801350, Z 22		HQ616701*	16.03/3067, MNCN
Annelida, Clitellata, Rhyacodrilinae <i>Rhyacodrilus okamikae</i> Giani & Rodriguez, 1988	Okamika cave, Santa Eufemia- Ereñozar karstic unit, Basque Country, Spain, UTM coordinates: X 0536852, Y 4797504, Z -5	HQ603824*	HQ616703*	16.03/3068 MNCN 16.03/3069
Annelida, Clitellata, Enchytraeidae <i>Buchholzia fallax</i> Michaelsen, 1887 <i>Fridericia tuberosa</i> Rota, 1995		AY885581*		
		AY885580*	AF064047*	

Table 2. Maximum uncorrected “p” distances (values on the top) and maximum Kimura 2-parameter distances (values on the bottom) of 16S rRNA (values in the top-right triangle) and COI (values in the low-left triangle) gene fragments between taxa. All values are based on the pairwise analysis of the 48 taxa data set for the 16S rRNA and 53 taxa data set for the COI.

Lineage	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.
1. <i>T. galarzai</i> IP	p K2P	10 11	15 17	20 23	23 28	23 28	22 26	20 23	22 26	24 29	24 28	21 24	22 25	25 30	24 29	-	-	22 26	23 27	24 29	27 33	24 28	25 30
2. <i>T. galarzai</i> F	18 21	p K2P	12 13	17 19	20 24	20 23	20 24	18 21	20 23	22 25	22 26	20 23	21 25	22 26	23 28	-	-	21 25	22 26	23 27	27 33	24 29	23 28
3. <i>H. costata</i>	19 22	18 21	p K2P	19 21	19 21	19 22	20 23	19 22	20 23	21 24	19 22	19 22	20 23	21 25	22 26	-	-	21 25	19 22	24 28	22 26	22 26	22 26
4. <i>I. gianii</i>	21 25	23 28	21 25	p K2P	18 20	18 21	20 24	20 23	17 19	21 24	23 27	20 23	23 27	20 23	24 30	-	-	22 26	20 24	23 27	24 29	21 24	22 26
5. <i>L. ignota</i> IP	18 21	20 23	21 25	19 22	p K2P	1 1	18 21	18 20	17 19	18 21	21 24	18 20	22 26	22 26	23 27	-	-	22 26	17 19	17 20	22 27	21 24	23 27
6. <i>L. ignota</i> IP	20 23	21 24	21 25	20 24	3 3	p K2P	18 21	18 20	16 18	18 21	20 24	18 21	22 26	22 26	22 26	-	-	21 25	17 20	17 19	22 26	21 25	23 27
7. <i>L. hoffmeisteri</i>	22 26	21 25	22 26	21 25	21 24	24 20	p K2P	11 12	14 15	21 24	20 23	18 21	23 28	22 26	23 27	-	-	22 24	20 23	22 26	24 28	19 22	25 31
8. <i>L. cervix</i>	-	-	-	-	-	-	-	p K2P	15 16	22 26	19 22	17 20	22 26	20 24	22 26	-	-	20 24	19 21	21 24	23 28	19 21	22 27
9. <i>L. udekemianus</i>	-	-	-	-	-	-	-	-	p K2P	19 22	20 23	17 19	22 26	20 23	21 25	-	-	21 25	18 21	20 23	22 26	19 22	22 26
10. <i>I. templetoni</i>	22 26	22 27	21 25	23 28	21 24	21 25	21 25	-	-	p K2P	17 19	13 15	19 22	16 19	19 21	-	-	18 21	17 20	23 29	19 22	21 25	22 26
11. <i>V. bizkaiensis</i>	22 26	23 27	22 25	23 28	21 24	22 25	20 23	-	-	25 30	p K2P	8 9	19 22	15 17	19 21	-	-	18 21	17 19	23 27	20 23	20 23	23 28
12. <i>P. bavaricus</i>	-	-	-	-	-	-	-	-	-	-	-	p K2P	17 19	13 14	17 19	-	-	15 17	15 17	21 25	20 23	18 21	21 24
13. <i>T. amplivasatus</i>	25 31	25 31	23 28	25 31	23 28	24 29	24 29	-	-	25 30	25 30	-	p K2P	18 21	10 12	-	-	21 25	21 25	21 25	24 30	21 25	25 30
14. <i>T. benedii</i>	23 28	23 28	24 28	23 28	20 23	21 25	23 27	-	-	23 28	23 27	-	21 26	p K2P	19 21	-	-	19 22	18 20	23 28	22 25	20 24	22 27
15. <i>T. bermudae</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	p K2P	-	-	21 24	20 23	25 31	24 29	22 26	26 32
16. <i>T. kozlovi</i>	23 29	24 30	23 28	24 29	24 29	24 29	24 28	-	-	24 30	23 27	-	19 23	21 24	-	p K2P	-	-	-	-	-	-	-
17. <i>T. swirencowi</i>	24	25	23	23	22	23	24	-	-	23	25	-	23	21	-	22	p	-	-	-	-	-	-

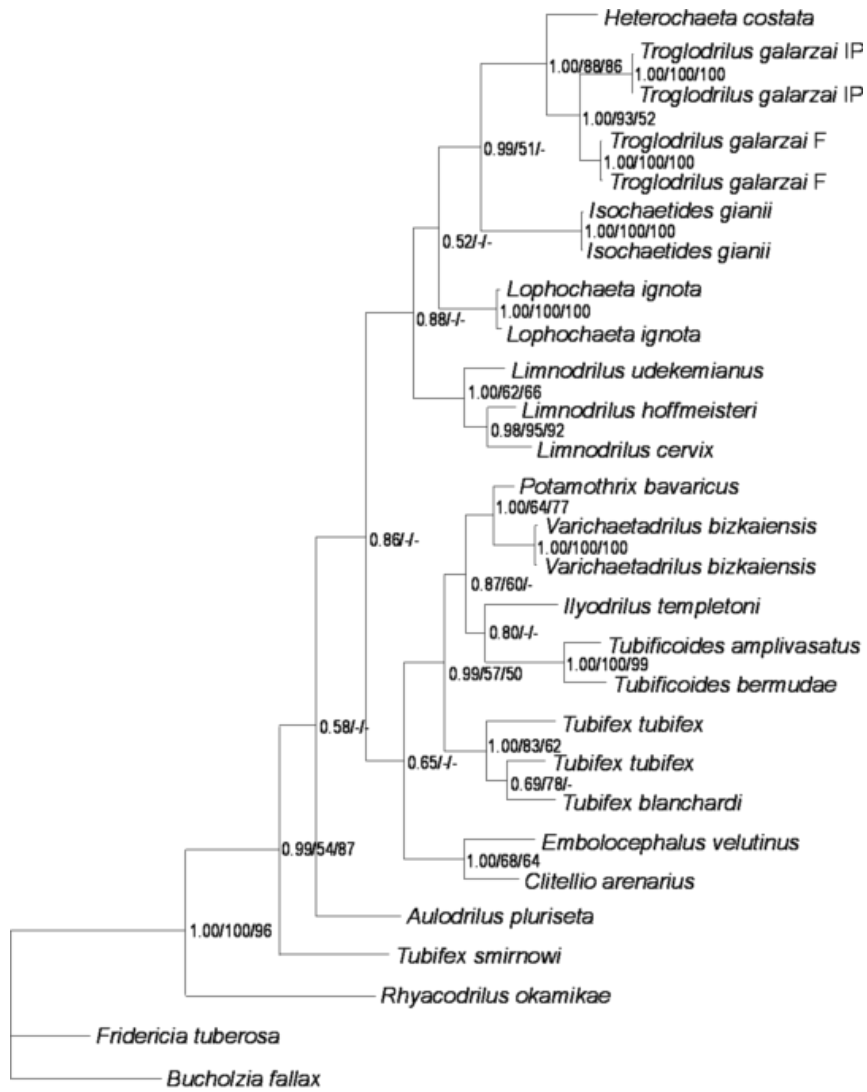
	29	30	27	27	26	27	30			28	30		28	28		25	K2P						
18. <i>T. tubifex</i>	25 25	22 23	22 26	26 28	23 27	24 29	23 27	-	-	24 29	27 28	-	24 30	24 27	-	25 30	25 31	p K2P	13 14	21 25	21 27	19 23	20 25
19. <i>T. blanchardi</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	p K2P	21 24	20 24	18 21	21 25
20. <i>A. pluriseta</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	p K2P	21 25	20 23	20 23
21. <i>E. velutinus</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	p K2P	16 18	22 26
22. <i>C. arenarius</i>	23 28	23 27	20 24	25 30	22 27	22 27	24 28	-	-	22 26	24 30	-	25 31	22 26	-	25 30	21 29	25 30	-	-	-	p K2P	20 23
23. <i>T. smirnowi</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	p K2P

Figure captions

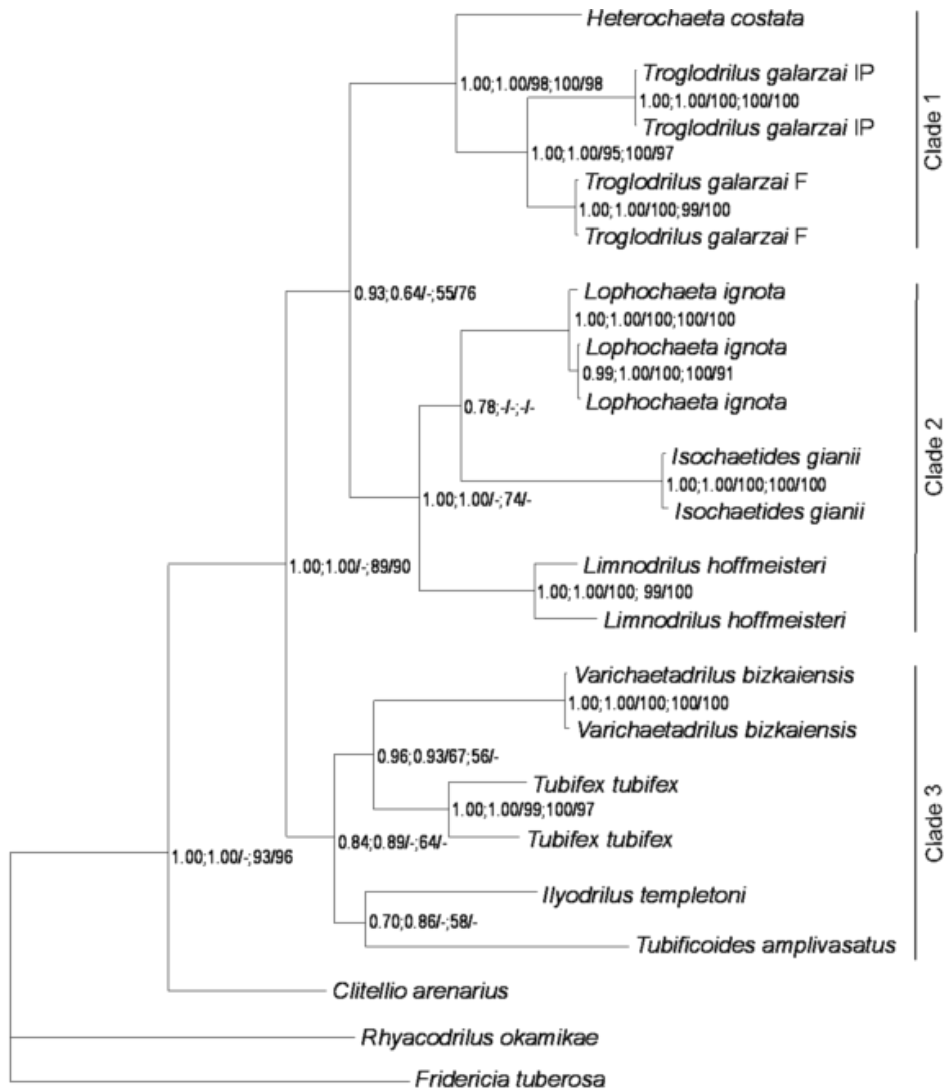
Fig. 1 Bayesian consensus tree showing the phylogenetic relationships of 18 tubificine species obtained from 16S rRNA sequences. BA posterior probabilities (>0.5), ML bootstrap values ($>50\%$) and MP bootstrap values ($>50\%$) are given at the nodes. IP = Iberian Peninsula population; F = French population.

Fig. 2 Bayesian consensus tree of combined sequences of COI and 16S rRNA genes. Five node support values are given using the formula: BA posterior probabilities excluding the third codon position; BA posterior probabilities including the third codon position / ML bootstrap values excluding the third codon position; ML bootstrap values including the third codon position / MP bootstrap values excluding the third codon position. Only support values $>0.5;0.5/50;50/50$ are given. The vertical bars denote ingroup clades discussed in the text. IP = Iberian Peninsula population; F = French population.

Appendix 1 Bayesian consensus tree showing the phylogenetic relationships using forty-six 16S rRNA sequences. BA posterior probabilities (>0.5), ML bootstrap values ($>50\%$) and MP bootstrap values ($>50\%$) are given at the nodes. IP = Iberian Peninsula population; F = French population; * = new sequenced specimen from the culture at the UPV/EHU.



0.1 substitution/site



0.1 substitution/site

