

Integrative taxonomy of the freshwater worm *Rhyacodrilus falciformis* s.l. (Clitellata: Naididae), with the description of a new species

SVANTE MARTINSSON, AINARA ACHURRA, MARCUS SVENSSON,
CHRISTER ERSÉUS

The genetic and morphological variation within *Rhyacodrilus falciformis* Bretscher, 1901 (Clitellata: Naididae) in Europe was explored using an integrative approach, with three unlinked genetic markers (the mitochondrial cytochrome *c* oxidase subunit I (COI) and the nuclear Histone 3 (H3) and Internal Transcribed Spacer region (ITS)) combined with morphology, to investigate whether this taxon constitutes a single or several species. Using Automatic Barcode Gap Discovery on the COI dataset, the specimens were divided into seven clusters, used as hypothetical species that were further tested with the other data sources. Single-gene trees were estimated for all three markers, using coalescence analysis and they were in many parts incongruent with each other. Only one of the clusters was supported by all trees; it was also morphologically differentiated from the other clusters by the shape of its modified penial chaetae. This group consists of two specimens from the Crotot Cave in south-eastern France, and morphologically they fit a previously described but invalid variety, ‘*pigueti*’, which is here described as a new species, *Rhyacodrilus pigueti* Achurra & Martinsson **sp. n.** The study highlights the fact that a single data source (e.g., COI-barcodes) seldom provides a sufficient basis for taxonomic decisions such as species delimitation.

Svante Martinsson, Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, SE-405 30 GÖTEBORG, Sweden. E-mail: svante.martinsson@bioenv.gu.se

Ainara Achurra, (1) Station d’Ecologie Expérimentale du CNRS à Moulis, USR 2936, 09200 Moulis, France. (2) Department of Zoology and Animal Cell Biology, Faculty of Science and Technology, University of the Basque Country, Box 644, 48080 Bilbao, Spain. E-mail: ainara.achurra@ehu.es

Marcus Svensson, Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, SE-405 30 GÖTEBORG, Sweden. E-mail:

boo_marcus@hotmail.com

Christer Erséus, Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, SE-405 30 GÖTEBORG, Sweden. E-mail:

christer.erseus@bioenv.gu.se

Introduction

Cryptic species are species that are morphologically indistinguishable or so similar that they have been classified under the same species name (Bickford *et al.* 2007). This seems to be a common phenomenon among several organismal groups, not the least among clitellate worms (see review by Erséus & Gustafsson 2009). Due to the lack of externally discernible characters, especially in immature specimens, many clitellate species have proved hard to distinguish without the aid of molecular markers and the diversity has in many cases been shown to be underestimated when based on morphology alone (e.g. Gustafsson *et al.* 2009; Envall *et al.* 2012; Matamoros *et al.* 2012; Novo *et al.* 2012).

The mitochondrial (mt) genome is haploid and almost exclusively transmitted maternally reduces the effective population size (N_e) to generally one fourth of that of nuclear markers (Birky *et al.* 1989), increasing the genetic drift fourfold, resulting in faster lineage sorting and shorter time to monophyly (Neigel & Avise 1986). This together with the fact that mt genes normally evolve several times faster than nuclear genes (Brown *et al.* 1979; Brown *et al.* 1982; Lin & Danforth 2004) has made such markers widely used in studies of recent divergence and species delimitation in several animal groups including clitellates (e.g. Heethoff *et al.* 2004; James *et al.* 2010; Dózsa-Farkas *et al.* 2012). The cytochrome *c* oxidase subunit I (COI) gene is the most often used marker in such studies, and a fragment of COI has been proposed as the standard “DNA-barcode” for identifying animal taxa (Hebert *et al.* 2003). The identification is often based on the assumption that intraspecific divergence is distinctly lower than interspecific divergence (the so-called barcoding gap) and is done by comparing the sequence of an unidentified individual with known sequences in a database (see review by Taylor & Harris 2012). However, it has been shown that the use of single mt-markers for species delimitation can overestimate the numbers of species if used alone, i.e., without nuclear markers (e.g. King *et al.* 2008; Dasmahapatra *et al.* 2010; Torres-Leguizamon *et al.* 2012; Achurra & Erséus 2013).

Species delimitation and species conceptualization have long been debated among biologists, mostly regarding the species criteria. Different schools within the field have put forward their own different species concepts, based on e.g. reproductive isolation (Mayr 1942), specific mate recognition (Paterson 1985), different ecological niches (Van Valen 1976), monophyly (Rosen 1979) and diagnosability (Nelson & Platnick

1981). De Queiroz (2007) proposed a *unified species concept*, where being a separately evolving metapopulation lineage is the sole requirement of a species, and where all other criteria (secondary species criteria) are used for assessing lineage separation (species delimitation). The more secondary species criteria supporting a divergence, the stronger is the case for speciation. However, one piece of evidence, if properly examined, may be enough to establish lineage separation. De Queiroz's view, which will be shared throughout this paper, is useful when dealing with cryptic species.

Rhyacodrilus falciformis Bretscher, 1901 (Clitellata: Naididae *sensu* Erséus *et al.* 2008) is widespread in the Holarctic region. It is found in freshwater bodies, but also in wet soil and seems to be associated with ground-water (Timm *et al.* 1996; Erséus *et al.* 2005; Dumnicka 2006; Achurra & Rodriguez 2008), and it may therefore be regarded as a stygophilic species (Giani *et al.* 2011). The worm is about 5 mm long and can be identified by the sickle-shaped penial chaetae, which have given the species its name. These chaetae are located in close proximity to the male pores in sexually mature specimens. A variety, '*piguetti*', with straight penial chaetae was described by Juget (1967) from Lake Léman on the border between France and Switzerland. However, as the description of this variety was published after 1960 the name is not available and has no nomenclatorial status (ICZN 1999: §15.2). In the course of the study of the diversity of European freshwater clitellates using DNA barcoding large genetic variation was observed in *R. falciformis*.

This study's aim is to explore the genetic and morphological variation within *Rhyacodrilus falciformis* in order to investigate whether this taxon comprises a complex of cryptic species or not.

Material and methods

Taxa and specimens

Thirty-six specimens of *Rhyacodrilus falciformis* were collected in various locations in France, Great Britain, Spain and Sweden, between 2003 and 2012 (Table 1). Wet soil samples and sandy sediments were taken in aquatic habitats (e.g., groundwater springs, caves, streams, ponds, lakes etc.) for sieving and sorting under a dissecting microscope in the laboratory. After preliminary identification, specimens were divided into two

parts; the anterior parts were stained with paracarmine and mounted in Canada balsam as outlined by Erséus (1994), as hologenophore vouchers (*sensu* Pleijel *et al.* 2008), and for morphological examinations, whereas the rear ends were transferred into 95% ethanol for subsequent DNA extraction and sequencing. For specimen CE14049 only a fragment was available and used for DNA extraction.

DNA sequencing and assembly

The posterior parts of all 36 individuals were selected for DNA extraction, using Qiagen's DNeasy Blood & Tissue Kit or Epicentre's QuickExtract DNA Extraction Solution 1.0, following the manufacturer's instructions. One specimen was handled by Canadian Centre for DNA Barcoding (CCDB) (Guelph, Canada), with data stored at the Barcode of Life Datasystems (BOLD). Parts of the mitochondrial cytochrome *c* oxidase subunit I (COI) and the nuclear Histone 3 (H3) and complete Internal Transcribed Spacer region (ITS; ITS1, 5.8S rDNA and ITS2) were amplified using primers and PCR programs listed in Table S1. After amplification by PCR, the existence of the target genes was tested using 1% agarose gel electrophoresis. PCR products were purified using Exonuclease I (Fermentas) and FastAP Thermosensitive Alkaline Phosphatase (Fermentas). Sequencing was carried out by Macrogen Inc. (Seoul, Korea) and Eurofins mwg operon (Ebersberg, Germany). Sequences were assembled in Geneious Pro v. 5.6.3 (Biomatters Ltd.) and aligned using *MAFFT v6.814b* (Kato *et al.* 2002) as implemented in Geneious using the auto algorithm.

Distance analyses

Pairwise genetic distances were calculated for the COI-data set in MEGA 5.1 using both uncorrected p-distances and the Tamura-Nei, 1993 (TN93) model, using Pairwise deletion for missing data. Models were chosen after model testing conducted in MEGA 5.1 (Tamura *et al.* 2011) using Bayesian information criterion (BIC). Histograms of distances were drawn in Microsoft Excel.

The TN93 distances were analysed with the online version of ABGD (Automatic Barcode Gap Discovery) (Puillandre *et al.* 2012) (available at <http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html>) with default settings, to delimit mitochondrial clusters. The ABGD method detects the first significant gap (the

barcoding gap) between the two modes of pairwise distance distributions, where the lower distance values presumably reflect the intraspecific differences and the higher the interspecific differences. The delimited COI clusters were used as hypothetical species that were further tested (see below).

Coalescent analyses

Genealogies were estimated for the COI, ITS and H3 datasets under the null hypothesis that all the included specimens of *R. falciformis* constitute a single species. Estimations were performed using Bayesian coalescence analysis as implemented in the BEAST package (Drummond & Rambaut 2007; Drummond *et al.* 2012). All .xml input files were created using BEAUti v1.7.4, with the TN93 + I model for COI, Jukes-Cantor, 1969 (JC69) + Γ model for ITS and Hasegawa-Kishino-Yano (HKY) + Γ model for H3; the model testing were conducted in MEGA 5.1 using Bayesian information criterion (BIC). For all analyses the following settings were used; base frequencies “estimated”; clock model “lognormal relaxed clock (Uncorrelated)”; tree prior “Coalescent: Constant Size”; constant.popSize “lognormal: Log(Mean) = 0, Log(Stdev) = 1, offset = 0”. For other priors default settings were used. The analyses were run in BEASTMC3 v. 1.7.4 (Drummond & Rambaut 2007; Drummond *et al.* 2012), using three chains, a delta value of 1.00 and swapping chains every 100 generations. The COI and H3 analyses were run for 10 million generations and the ITS analysis for 20 million generations, saving trees every 1000th generation. The number of generations was chosen to obtain sufficient effective sample size (ESS). Tracer v1.5 (Rambaut & Drummond 2007) was used for examining ESS for parameters and determining the burn-in. Trees were summarized using TreeAnnotator v1.7.4., discarding the first 10 % as burn-in.

Molecular species delimitation

The gene-trees were imported to Geneious and in the cases where the previously delimited clusters were found monophyletic the Species delimitation plugin (Masters *et al.* 2011) was used to calculate, (1) P (*Randomly distinct*), i.e., the probability for clades to have the observed degree of distinctiveness due to random coalescent processes (Rodrigo *et al.* 2008), and (2) *Rosenberg's* P_{AB} which expresses the probability of

reciprocal monophyly of the clade of interest and its nearest defined group, under random branching (Rosenberg 2007). Both statistics were used to test whether or not the null hypothesis of random coalescence can be rejected as the explanation for phylogenetic structure.

Morphological study

The microscopically studied material included 24 mature specimens and 7 immatures or fragments. For the detailed morphological analysis of the reproductive organs only completely mature specimens (i.e. with eggs, or with sperm in the spermathecal ampulla) were selected, while for external characters we included all available specimens. This selection guarantees that differences in measurements are not due to incompletely developed organs. Only anterior parts (approximately segments I – XV) were examined since posterior parts were used for the molecular study. External characters are related to the number and size of somatic chaetae, internal characters are related to the reproductive organs: spermathecae, male ducts and penial chaetae. The width of segment XI was included to represent the size of each specimen.

Concatenated phylogenetic analysis

In order to test whether the two species found (see Result) were reciprocally monophyletic or not was a phylogenetic analysis performed on a concatenated dataset.

The three single-gene matrixes were concatenated into one data matrix using Geneious Pro v. 5.6.3, one specimen each of *Rhyacodrilus coccineus* and *R. subterraneus* were added as out-groups (see Table 1). A phylogenetic tree was estimated using Maximum Likelihood; the analysis was performed with PHYML 3.0 (Guindon & Gascuel, 2003; S. Guindon *et al.*, 2010) as implemented at the ATGC Montpellier bioinformatics platform (<http://www.atgc-montpellier.fr/>). The TN93 model with Γ parameter estimated from the data was used after model testing in MEGA 5.2, SPR+NNI was used for tree improvement. Branch support was calculated with the Chi²-based approximate Likelihood Ratio Test (aLRT) (Anisimova & Gascuel, 2006) in PHYML.

Data deposition

All new sequences generated in this study are deposited in GenBank, submission numbers are listed in Table 1; all xml files used in the BEAST analyses, as well as log files from the analyses, are deposited in the Dryad Data Repository (<http://www.datadryad.org/>) at DOI:10.5061/dryad.1fb3d; trees and matrices are deposited in TreeBASE (<http://treebase.org/>), submission TB2:S14391 . Vouchers are deposited in the Swedish Museum of Natural History, Stockholm; accession numbers are given in Table 1.

Results

DNA sequencing

DNA from 36 samples of *Rhyacodrilus falciformis* s.l. were successfully amplified and sequenced with regard to COI, whereas in the cases of ITS, 22 specimens, and H3, 35 specimens were successfully sequenced (Tables 1 and S2). More details on the alignments are given in Table S2.

Distance analysis

The maximum pairwise distance values within *R. falciformis* s.l. for COI were 12.5 % using uncorrected p-distances and 14.0 % using TN93 distances. A barcoding gap was observed between 2.9 and 4.6% pairwise differences using uncorrected p-distances and between 3.0 and 4.9 % using TN93 distances (Fig. 1A). The ABGD analysis of the COI dataset yielded seven mitochondrial clusters, further used as hypothetical species, assuming maximal intraspecific variation between 1.3% and 6.0%. One cluster contains the specimens from France (cluster A), one the Spanish specimens (cluster B), one the English specimens (cluster C), and four groups includes specimens from Sweden (clusters D-G).

Coalescent analyses

In the COI tree (Fig. 1B) each of the seven groups recognized by the ABGD analysis was found monophyletic with maximum support. The French specimens (cluster A) are the sister-group to all other *R. falciformis* (posterior probability (pp) 0.94), whereas the clade with the British specimens (cluster C) is found as sister group to one of the clades from Sweden (cluster D) (pp 1), and the Spanish cluster B is sister-group to them with low support (pp 0.82).

In the ITS tree (Fig. 1C) there are five well supported clades (pp >0.99). One of them is identical to the French cluster (A) while the other clades either consist of specimens from several clusters or contain only some of the specimens from one cluster. Unfortunately, none of the specimens from Spain and only one specimen from England was successfully sequenced; the latter specimen (CE8714) having the same ITS haplotype as one Swedish specimen (CE11406). The French cluster A is nested among the other clades, but the relationships between most of the clades are unsupported.

In the H3 tree (Fig. 1D) clusters A, B and C are monophyletic with maximum support. None of the other mt-clusters are recovered in the analysis and the relationships between groups are generally without support.

Molecular species delimitation

In the Geneious Species delimitation analyses on the COI genealogy, the null-hypothesis of observing reciprocal monophyly by chance could be rejected for all groups [*Rosenberg's* $P_{AB} \leq 0.05$], and the null-hypothesis that the observed degree of distinctiveness is due to random coalescence processes could be rejected [P (*Randomly Distinct*) <0.05] for all groups except cluster G [P (*Randomly Distinct*) = 0.10]. In the ITS genealogy, all five well supported clades (including cluster A) had a *Rosenberg's* $P_{AB} \leq 0.05$, whereas cluster A had a P (*Randomly Distinct*) of 0.12 and P (*Randomly Distinct*) varied between 0.11 and 0.45 for the other clades. In the H3 genealogy, only clades A-C were monophyletic and could be tested. For these groups *Rosenberg's* $P_{AB} \leq 0.05$, whereas P (*Randomly Distinct*) was between 0.47 and 0.73.

Morphological study

All examined mature specimens fit with previous descriptions of *Rhyacodrilus falciformis* Bretscher, 1901 (Piguet 1906; Hrabě 1935; Juget 1967; Kasprzak 1979;

Chekanovskaya 1981; Timm *et al.* 1996) except for individual CE14048, which shows different genital chaetae (Fig. 2).

We compared the ranges of the values of morphological characters across the COI clusters to search for congruence with molecular results (Table S3). There is no evidence for morphological differentiation in reproductive characters among the clusters, except the different genital chaetae in specimen CE14048 (cluster A). Slight differences in number and size of somatic chaetae (external characters) are frequent among populations and are thus not considered species specific. At specimen level, such differences may be due to the use and replacement of chaetae. The number of genital chaetae as well as their shape and size are considered to be diagnostic characters in species of the genus *Rhyacodrilus* Bretscher, 1901 (Timm 2009; Rodriguez & Fend 2013).

Concatenated phylogenetic analysis

The Maximum Likelihood tree based on combined COI, H3 and ITS (Supplementary Fig. S1) shows a coherent, maximally supported, group of all specimens of *Rhyacodrilus falciformis* s.l.. All clusters except cluster E are found with maximal support, cluster E is paraphyletic with respect to a clade consisting of clusters B-D. Cluster A, *Rhyacodrilus pigueti* is found as sister-group to all other clusters, *Rhyacodrilus falciformis* s.s. with a support of 0.99.

Therefore, in the light of molecular and morphological data (results summarised in Fig 3) and further discussed in Discussion, we consider cluster A as a separate species from *R. falciformis* s.s.. The shape of the penial chaetae of CE14048 is in accordance with the description of *R. falciformis* var. *pigueti* Juget, 1967. However, a new name published after 1960 expressly as a "variety" is considered to be infrasubspecific and as such does not formally exist (ICZN 1999: §15.2). Therefore, a new species is proposed below for this form, including the two specimens in cluster A.

Taxonomy

Genus *Rhyacodrilus* Bretscher, 1901

Rhyacodrilus pigueti Achurra & Martinsson, **sp. n.**

(Figure 2A, C)

Rhyacodrilus falciformis, var. *pigueti* Juget, 1967: Figure 2B, Table.

Holotype. CE14048, whole-mounted worm, anterior part (segments I-XVI only), stained in paracarmine and mounted in Canada balsam. Collected by Michel C. des Chatelliers, 26 Jan 2011. Deposited in Swedish Museum of Natural History (SMNH), Stockholm, accession no. SMNH000000. COI barcode sequence, GenBank acc. no. XX000000; ITS sequence, GenBank acc. no. XX000000; H3 sequence, GenBank acc. no. XX000000.

Type locality. Crotot Cave, northeast of Besançon, Département du Doubs, France (47° 26' 44.52'' N, 6° 24' 31.64'' E, DMS datum WGS84; elevation: 375 m).

Other material, genetically examined. CE14049, only molecular data, no voucher available. COI barcode sequence, GenBank acc. no. xxxxxx; ITS sequence, GenBank acc. no. xxxxxx; H3 sequence, GenBank acc. no. xxxxxx.

Etymology. Named after Émile Piguet, a specialist on aquatic oligochaetes of the early 20th century. The new material fits the description of *R. falciformis* var. *pigueti* Juget, 1967, but as varieties published after 1961 are invalid, Juget's name *pigueti* is not available (§15.2 and §45.5 ICZN 1999). Therefore the name *pigueti* is here proposed as new for this taxon, now regarded as a species. According to Juget 1967, Émile Piguet had once found specimens with the characters of Juget's variety, but Piguet only identified them as *Rhyacodrilus* sp. and never formally described them taxonomically.

Diagnosis. This species can be distinguished from other species in *Rhyacodrilus* by the shape of the penial chaetae (Fig 2) and by genetic data (Fig 3). See also Remarks and Discussion below.

Description.

Incomplete specimen, with 16 anteriormost segments 1.93 mm long, body width 265 µm in segment V, 380 µm in segment XI. Prostomium 120 µm long, 132 µm wide. Coelomocytes abundant within coelomic cavity of all segments, round, 27 µm diameter, granulated, when nuclei stained with paracarmine. Clitellum indistinguishable (poorly developed?). Spermathecal pores mid-lateral, located anteriorly in segment X. One pair of male pores open ventrally in segment XI.

Dorsal and ventral chaetae of identical shape, bifid, with distal tooth somewhat thinner and longer than proximal. Anteriorly 3-5 chaetae per bundle, 1-2 in segments XI-XV, and 2-3 in segment XVI; maximum length 92 μm , 2-3 μm thick. Modified penial chaetae in segment XI, a single chaeta per “bundle”, straight, somewhat spoon-shaped (Fig. 2A, C); chaeta 150 μm long, 15 μm thick in proximal and middle parts, narrowing to 10 μm before widening to 15 μm in the distal portion. Proximal end of penial chaeta associated with conspicuous chaetal gland (diameter 75 μm) and strong muscular strands. Distal end located within male pore. Penial chaetae orientated with their distal ends towards segment X.

Atria paired in segment XI, close to septum XI/XII, elongate, pear-shaped, ampullae 150 μm long, maximum diameter 50 μm , with about 10 μm thick muscular layer. Proximal end of atrium barely observed due to the presence of the assembly of penial chaeta + chaetal gland + muscular strands. Abundant prostatic cells covering proximal and middle parts of atrium. Spermathecae paired, ampullae 96 μm diameter, 60 μm long, with a single-layer epithelium 5-10 μm thick; lumen filled with random mass of sperm; duct 80 μm long, 60 μm diameter.

Distribution and habitat. Crotot Cave is a large cavity (5,800 m long), in Jurassic limestone, on a plateau near the valley of the Doubs River (Departement du Doubs), eastern France. Following Juget (1967), the new species has also been found in the deep zone of Lake Léman (Lake Geneva), in Switzerland and France, approximately 100 km away from the Crotot Cave.

Remarks. *Rhyacodrilus pigueti* **sp. n.** is well distinguished from *R. falciformis* *s.s.* by the shape of the penial chaetae. The chaetae are sickle-shaped in *R. falciformis* while they are straight and somewhat spoon-shaped in the new species (Fig 2) with a generally larger ratio for the distance between the focal point of the sickle and the dorsal axis of the shaft divided with the length of the chaeta (0.03 in *R. pigueti* n=1, 0.02-0.10; mean = 0.06; n=7). However, this measurement should be interpreted with caution as it varies with the orientation of the chaeta. The distal end of the penial chaeta is more pointed and longer in *R. falciformis* than in the new species, in which the distal part is widening proximal to the apex. In addition, the penial chaetae are 150 μm long in the new species while they are between 100 and 140 μm long in our specimens of *R. falciformis*. Other measurements in the literature also show slightly shorter penial chaetae for *R. falciformis*: 100-138 μm (Piguet 1906; Hrabě 1935; Juget 1967; Kasprzak

1979; Chekanovskaya 1981; Timm *et al.* 1996), no measurements of the penial chaetae are given in the original description (Bretscher 1901).

Discussion

Based on the results from the ABGD and species delimitation tests on the COI dataset, it seems that *Rhyacodrilus falciformis* consists of 6-7 different species, one of which morphologically conforms to the “*pigueti*” form reported by Juget (1967). However, the barcoding gap is small 1.7-1.9 % depending on model, and there are no barcode gap between *Rhyacodrilus pigueti* and *R. falciformis s.s.* several other studies have found a large barcoding gap between clitellate species (e.g. Gustafsson *et al.*, 2009; Kvist *et al.*, 2010; Richard *et al.*, 2010; Envall *et al.*, 2012; Matamoros *et al.*, 2012). The difference in barcoding gap between that study and others may be explained by the age of the lineages in the studies, were the split between *Rhyacodrilus pigueti* and *R. falciformis s.s.* could be distinctly younger than the splits between the taxa in the other studies. In the nuclear gene trees, members of the hypothetical species suggested by the mtDNA analyses are in many cases found scattered and mixed in different clades. In the case of ITS, five well supported clades were found, with a *Rosenberg's* $P_{AB} \leq 0.05$, whereof one constitutes cluster A: *Rhyacodrilus pigueti*.

The Molecular operational taxonomic units (MOTUs) (Floyd *et al.* 2002) that will be defined thus differ considerably between the genes, both in numbers and constitution. This highlights the problems with using a single genetic marker for delimiting taxonomic units. Instead using an integrative approach (see review by Schlick-Steiner *et al.* 2010) combining several genetic markers as well as other data sources can result in more robust taxonomic hypotheses.

The support for the hypothesis that the delimited COI-based groups represent separately evolving lineages varies. Some of the groups are supported by COI only, others by two loci, but only cluster A, here regarded as *Rhyacodrilus pigueti* **sp. n.**, is supported by all genes as well as morphology (Fig. 3). The differences in topology between the gene-trees suggest that gene flow is present between some of the different COI-based groups, and that speciation – if at all in progress – is not yet complete between them. However, recent speciation events, where lineage sorting of nuclear genes still is incomplete and morphological differences not yet have accumulated to any

discernible degree, cannot be ruled out, especially not as some of the COI clusters show a clear geographic pattern.

The shape of the penial chaetae is the only morphological diagnostic character that we found to separate the new species from *R. falciformis* s.s., but together with the molecular data it gives us strong support for *R. pigueti* being a separately evolving lineage. The integration of independent data sources are here crucial for a proper determination of the species boundaries in *R. falciformis* s.l.; neither individual gene trees nor the morphological data alone provide enough support to clearly delimit the two species. New techniques may reveal other putative diagnostic anatomic structures that are currently not observable. For instance, Cuadrado and Martinez-Ansemil (2001), using Scanning Electron Microscopy (SEM), found several bundles of cilia near the male pores in specimens of *R. falciformis* that had not been observed with other techniques. Such ultrastructural details may deserve attention in future taxonomic work on cryptic species of aquatic clitellates. However, SEM is a rather expensive and destructive technique, and may therefore be of little help in specimen identification.

It is possibly that there are physiological and/or ecological differences between the two species as has been shown for the different lineages within *Tubifex tubifex* (Sturmbauer *et al.*, 1999; Beauchamp *et al.*, 2002).

The separation of *R. pigueti* seems to render *R. falciformis* paraphyletic in the ITS and H3 trees, but the support for the topologies is weak or lacking and in the concatenated analysis *R. falciformis* is retrieved with high support. However, even if the topologies should prove to be correct despite the low or lacking support, non-monophyly of species in single gene-trees are not uncommon (Crisp & Chandler 1996; Funk & Omland 2003), and the reason in this case is likely to be recent separation and incomplete lineage sorting. The N_e of the nuclear genes is greater than that of mitochondrial genes and it takes a longer time for such genes to reach complete lineage sorting and reciprocal monophyly (see Introduction). Evidence of reciprocal monophyly in several gene-trees is a strong indication that groups are separately evolving lineages, i.e. different species. However, reciprocal monophyly is not a necessity for species delimitation (see e.g. Doyle 1995; Helbig *et al.* 2002) except under a strict monophyletic species concept.

The divergence in COI sequences observed in the Swedish specimens of *R. falciformis* could be due to allopatric divergence with secondary contact. This could be the result of recolonisation from different refugia after deglaciation, as has been

discussed for other animal groups (e.g. Verovnik *et al.* 2005; De Wit & Erséus, 2010; Achurra & Erséus 2013). The combination of higher mutation rate and a faster genetic drift in mitochondrial markers compared to nuclear ones could explain why divergent haplotype clusters are formed to a higher extent in the COI-dataset compared to the nuclear datasets. These haplotype clusters will then be retained after a secondary contact, but if gene flow between them resumes the unlinked nuclear and mitochondrial haplotype clusters will be mixed in the populations the nuclear genome will also be further affected by recombination increasing the blending between haplotype clusters. Specimens from different Swedish COI-clusters are found at the same locality in several cases (see Table 1).

The western Alps and adjacent areas seem to have a rich *Rhyacodrilus* fauna both in surface and ground water (e.g. Juget 1984; 1987; Lafont & Juget 1993) including the here described *R. pigueti*, which together with *R. falciformis* is also found in Lake Léman (Juget 1967). From Lake Léman, also a seemingly undescribed species was reported and illustrated by Piguet and Bretscher (1913: Fig. 13). It has straight penial chaetae of approximately the same length as the somatic chaetae and lack hair chaetae, like both *R. falciformis* and *R. pigueti*. Future studies of the clitellate fauna in this area may reveal an even larger diversity of *Rhyacodrilus*.

Finding unexpected genetic divergence within a morphospecies can be a starting point for further investigation using an integrative approach with various data sources. However, COI alone should not be used to claim that cryptic speciation has occurred, as several studies have found deep mitochondrial divergence within populations with retained gene flow (e.g. Webb *et al.* 2011; Hogner *et al.* 2012; Torres-Leguizamon *et al.* 2012). Nevertheless, we certainly recognise the usefulness of DNA-barcoding in specimen identification, once species boundaries have been properly established.

Acknowledgements

This collaborative effort was made possible by a Zoologica Scripta Travel Fellowship, from The Norwegian Academy of Science and Letters to the second author. Further, we are grateful to Alexander Bär, Michel C. de Chatelliers, Tim Jones, Lee Knight, Emilie Lindquist, and Patrick Martin for collecting and/or otherwise making material available to us; to Anna Ansebo, Alexander Bär, Ida Envall, Magnus Johansson, Emelie Lindquist, and Maria Lindström, for assistance in the lab; and to the

Swedish Research Council, The Swedish Taxonomy Initiative (ArtDatabanken), the University of The Basque Country UPV/EHU (Contratación de Doctores Recientes hasta su Integración en Programas de Formación Postdoctoral del Vicerrectorado de Investigación) and The Basque Government (Programa Posdoctoral de perfeccionamiento de doctores), for financial support.

References

- Achurra, A., & Erséus, C. (2013). DNA barcoding and species delimitation: the *Stylogdrilus heringianus* case (Annelida : Clitellata : Lumbriculidae). *Invertebrate Systematics*, 27, 118-128.
- Achurra, A., & Rodriguez, P. (2008). Biodiversity of groundwater oligochaetes from a karst unit in northern Iberian Peninsula: ranking subterranean sites for conservation management. *Hydrobiologia*, 605, 159-171.
- Anisimova, M. & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology*, 55, 539-552.
- Beauchamp, K. A., Gay, M., Kelley, G. O., El-Matbouli, M., Kathman, R. D., Nehring, R. B., et al. (2002). Prevalence and susceptibility of infection to *Myxobolus cerebralis*, and genetic differences among populations of *Tubifex tubifex*. *Dis Aquat Organ*, 51, 113-121.
- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., et al. (2007). Cryptic species as a window on diversity and conservation. *Trends Ecol Evol*, 22, 148-155.
- Birky, C. W., Jr., Fuerst, P., & Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, 121, 613-627.
- Bretscher, K. (1901). Beobachtungen über die Oligochaeten der Schweiz. *Revue Suisse Zoologie*, 9, 189-223.
- Brown, W. M., George, M., Jr., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*, 76, 1967-1971.
- Brown, W. M., Prager, E. M., Wang, A., & Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol*, 18, 225-239.

- Chekanovskaya, O. V. (1981). *Aquatic Oligochaeta of the USSR*. New Delhi: Amerind Publishing Co. Pvt. Ltd.
- Crisp, M. D., & Chandler, G. T. (1996). Paraphyletic species. *Telopea*, 6, 813-844.
- Cuadrado, S., & Martinez-Ansemil, E. (2001). External structures used during attachment and sperm transfer in tubificids (Annelida, Oligochaeta). *Hydrobiologia*, 463, 107-113.
- Dasmahapatra, K. K., Elias, M., Hill, R. I., Hoffman, J. I., & Mallet, J. (2010). Mitochondrial DNA barcoding detects some species that are real, and some that are not. *Mol Ecol Resour*, 10, 264-273.
- De Queiroz, K. (2007). Species concepts and species delimitation. *Syst Biol*, 56, 879-886.
- De Wit, P., & Erséus, C. (2010). Genetic variation and phylogeny of Scandinavian species of *Grania* (Annelida: Clitellata: Enchytraeidae), with the discovery of a cryptic species. *Journal of Zoological Systematics and Evolutionary Research*, 48, 285-293.
- Doyle, J. J. (1995). The Irrelevance of Allele Tree Topologies for Species Delimitation, and a Non-Topological Alternative. *Systematic Botany*, 20, 574-588.
- Dózsa-Farkas, K., Porco, D., & Boros, G. (2012). Are *Bryodrilus parvus* Nurminen, 1970 and *Bryodrilus librus* (Nielsen and Christensen, 1959) (Annelida: Enchytraeidae) really different species? A revision based on DNA barcodes and morphological data. *Zootaxa*, 3276, 38-50.
- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*, 7, 214.
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*, 29, 1969-1973.
- Dumnicka, E. (2006). Composition and abundance of oligochaetes (Annelida : Oligochaeta) in springs of Krakow-Czestochowa upland (Southern Poland): Effect of spring encasing and environmental factors. *Polish Journal of Ecology*, 54, 231-242.
- Envall, I., Gustavsson, L. M., & Erseus, C. (2012). Genetic and chaetal variation in *Nais* worms (Annelida, Clitellata, Naididae). *Zoological Journal of the Linnean Society*, 165, 495-520.
- Erséus, C. 1994. The Oligochaeta. In J. A. Blake & B. Hilbig (Eds) *Taxonomic Atlas of the Bentic Fauna of the Santa Maria Basin and Western Santa Barbara Channel*

Volume 4 Oligochaeta to Polychaeta: Phyllodocida (Phyllodocidae to Paralacydoniidae) pp. 5-38). Santa Barbara, California: Santa Barbara Museum of Natural History.

- Erséus, C., & Gustafsson, D. 2009. Cryptic Speciation in Clitellate Model Organisms. *Annelids in Modern Biology* pp. 31-46): John Wiley & Sons, Inc.
- Erséus, C., Rota, E., Matamoros, L., & De Wit, P. (2010). Molecular phylogeny of Enchytraeidae (Annelida, Clitellata). *Mol Phylogenet Evol*, 57, 849-858.
- Erséus, C., Rota, E., Timm, T., Grimm, R., Healy, B., & Lundberg, S. (2005). Riverine and riparian clitellates of three drainages in southern Sweden. *International Journal of Limnology*, 41, 183-194.
- Erséus, C., Wetzel, M. J., & Gustafsson, L. (2008). ICZN rules—a farewell to Tubificidae (Annelida, Clitellata). *Zootaxa*, 1744, 66-68.
- Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, 11, 839-850.
- Funk, D. J., & Omland, K. E. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 397-423.
- Giani, N., Sambugar, B., Martínez-Ansemil, E., Martin, P., & Schmelz, R. M. (2011). The groundwater oligochaetes (Annelida, Clitellata) of Slovenia. *Subterranean Biology*, 9, 85-102.
- Guindon, S., Dufayard J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307-321.
- Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood. *Systematic Biology*, 52, 696-704.
- Gustafsson, D. R., Price, D. A., & Erséus, C. (2009). Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Mol Phylogenet Evol*, 51, 182-189.
- Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc Biol Sci*, 270, 313-321.
- Heethoff, M., Etzold, K., & Scheu, S. (2004). Mitochondrial COII sequences indicate that the parthenogenetic earthworm *Octolasion tyrtaeum* (Savigny 1826)

- constitutes of two lineages differing in body size and genotype. *Pedobiologia*, 48, 9-13.
- Helbig, A. J., Knox, A. G., Parkin, D. T., Sangster, G., & Collinson, M. (2002). Guidelines for assigning species rank. *Ibis*, 144, 518-525.
- Hogner, S., Laskemoen, T., Lifjeld, J. T., Porkert, J., Kleven, O., Albayrak, T., et al. (2012). Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecol Evol*, 2, 2974-2988.
- Hrabě, S. (1935). Über *Moraviodrillus pygmaeus* n. g. n. sp., *Rhyacodrillus falciformis* Br., *Ilyodrillus bavaricus* Oschm. und *Bothrioneurum vej dovskyanum* Stolc. *Publications de la Faculté des Sciences de l'Université Masaryk*, 209, 1-19.
- ICZN (1999). *International code of zoological nomenclature*. London: The International Trust for Zoological Nomenclature.
- James, S. W., Porco, D., Decaens, T., Richard, B., Rougerie, R., & Erséus, C. (2010). DNA barcoding reveals cryptic diversity in *Lumbricus terrestris* L., 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). *PLoS ONE*, 5, e15629.
- Juget, J. (1967). Quelques données nouvelles sur les oligochètes du léman : composition et origine du peuplement. *Annales de Limnologie*, 3, 217-229.
- Juget, J. (1984). Oligochaeta of the epigeal and underground fauna of the alluvial plain of the French upper Rhône (biotypological trial). *Hydrobiologia*, 115, 175-182.
- Juget, J. (1987). Contribution to the study of the Rhyacodrilinae (Tubificidae, Oligochaeta), with description of two new stygobiont species from the alluvial plain of the French upper Rhone, *Rhyacodrillus amphigenus*, sp. n. and *Rhizodriloides phreaticola*, g. n., sp. n. *Hydrobiologia*, 155, 107-118.
- Kasprzak, K. (1979). Oligochaetes (Oligochaeta) of the Pien-iny Mountains. II. Naididae, Tubificidae, Haplotaxidae, Lumbriculidae, Branchiobdellidae. *Fragmenta Faunistica*, 24, 57-80.
- Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*, 30, 3059-3066.
- King, R. A., Tibble, A. L., & Symondson, W. O. (2008). Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. *Molecular Ecology*, 17, 4684-4698.

- Kvist, S., Sarkar, I. N., & Erséus, C. (2010). Genetic variation and phylogeny of the cosmopolitan marine genus *Tubificoides* (Annelida: Clitellata: Naididae: Tubificinae). *Mol Phylogenet Evol*, 57, 687-702.
- Lafont, M., & Juget, J. (1993). Description de *Rhyacodrilus ardierae* n. sp. (Oligochaeta, Tubificidae) récoltée dans l'Ardières (affluent de la Saône). *Bulletin de la Société zoologique de France*, 118, 115-123.
- Lin, C.-P., & Danforth, B. N. (2004). How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Mol Phylogenet Evol*, 30, 686-702.
- Masters, B. C., Fan, V., & Ross, H. A. (2011). Species Delimitation--a Geneious plugin for the exploration of species boundaries. *Mol Ecol Resour*, 11, 154-157.
- Matamoros, L., Rota, E., & Erséus, C. (2012). Cryptic diversity among the achaetous *Marionina* (Annelida, Clitellata, Enchytraeidae). *Syst Biodivers*, 10, 509-525.
- Mayr, E. (1942). *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. New York: Columbia University Press.
- Neigel, J. E., & Avise, J. C. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of specialtion. In E. Nevo & S. Karlin (Eds) *Evolutionary processes and theory* pp. 515-534). New York: Academic Press.
- Nelson, G., & Platnick, N. I. (1981). *Systematics and biogeography*. New York: Columbia University Press.
- Novo, M., Fernandez, R., Marchan, D. F., Monica, G., & Cosin, D. J. (2012). Compilation of morphological and molecular data, a necessity for taxonomy: The case of *Hormogaster abbatissae* sp. n. (Annelida, Clitellata, Hormogastridae). *Zookeys*, 1-16.
- Paterson, H. E. H. 1985. The recognition concept of species. In E. S. Vrba (Ed) *Species and speciation* pp. 21-29). Pretoria: Transvaal Museum.
- Piguet, E. (1906). Oligochètes de la Suisse française. *Revue Suisse de Zoologie*, 14, 389-403.
- Piguet, E., & Bretscher, K. (1913). *Oligochètes. Catalogue des invertébrés de la Suisse*. Genève: Georg & Cie.
- Pleijel, F., Jondelius, U., Norlinder, E., Nygren, A., Oxelman, B., Schander, C., et al. (2008). Phylogenies without roots? A plea for the use of vouchers in molecular phylogenetic studies. *Mol Phylogenet Evol*, 48, 369-371.

- Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 21, 1864-1877.
- Rambaut, A. & Drummond, A.J. (2007). Tracer v1.5. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Richard, B., Decaens, T., Rougerie, R., James, S. W., Porco, D., & Hebert, P. D. (2010). Re-integrating earthworm juveniles into soil biodiversity studies: species identification through DNA barcoding. *Mol Ecol Resour*, 10, 606-614.
- Rodrigo, A., Bertels, F., Heled, J., Noder, R., Shearman, H., & Tsai, P. (2008). The perils of plenty: what are we going to do with all these genes? *Philos Trans R Soc Lond B Biol Sci*, 363, 3893-3902.
- Rodriguez, P., & Fend, S. V. (2013). New species of *Rhyacodrilus* (Annelida: Clitellata: Rhyacodrilinae) of North America, with re-description of *R. sodalis* (Eisen, 1879). *Zootaxa*, 3664, 1.
- Rosen, D. E. (1979). Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bulletin of the American Museum of Natural History*, 162, 267-376.
- Rosenberg, N. A. (2007). Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution*, 61, 317-323.
- Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., & Crozier, R. H. (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu Rev Entomol*, 55, 421-438.
- Sturmbauer, C., Opadiya, G. B., Niederstätter, H., Riedmann, A., & Dallinger, R. (1999). Mitochondrial DNA reveals cryptic oligochaete species differing in cadmium resistance. *Mol Biol Evol*, 16, 967-974.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*, 28, 2731-2739.
- Taylor, H. R., & Harris, W. E. (2012). An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. *Mol Ecol Resour*, 12, 377-388.
- Timm, T. (2009). A guide to the freshwater Oligochaeta and Polychaeta of Northern and Central Europe. *Lauterbornia*, 66, 1-235.

- Timm, T., Erséus, C., & Lundberg, S. (1996). New and Unusual Records of Freshwater Oligochaeta from the Scandinavian Peninsula. *Nordic Journal of Freshwater Research*, 72, 15-29.
- Torres-Leguizamon, M., Mathieu, J., Livet, A., Decaens, T., & Dupont, L. (2012). Isolation of polymorphic microsatellite markers in *Aporrectodea icterica* (Savigny 1826). *Soil Biology & Biochemistry*, 51, 16-19.
- Van Valen, L. (1976). Ecological species, multispecies, and oaks. *Taxon*, 25, 233-239.
- Webb, W. C., Marzluff, J. M., & Omland, K. E. (2011). Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Molecular Ecology*, 20, 2390-2402.
- Verovnik, R., Sket, B., & Trontelj, P. (2005). The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molecular Ecology*, 14, 4355-4369.

Table 1 List of specimens included in the study, collection localities and dates, GenBank accession nos. and voucher nos

| Species | ID no. | COI cluster | Collection locality | Collection date | GenBank accession no. | | | Voucher no. |
|------------------------------------|---------|-------------|---|-----------------|-----------------------|----------|----------|-----------------------------|
| | | | | | COI | ITS | H3 | |
| <i>Rhyacodrilus falciformis</i> | CE631-1 | E | Vitärkällan Spring, Gotland, SE | September 2003 | KF267920 | KF267973 | KF267939 | SMNH130504 |
| <i>Rhyacodrilus falciformis</i> | CE637 | E | Vitärkällan Spring, Gotland, SE | September 2003 | KF267935 | KF267974 | KF267940 | SMNH130505 |
| <i>Rhyacodrilus falciformis</i> | CE795-1 | F | Wet soil, Hälleklis, Götene, SE | May 2004 | KF267926 | KF267990 | KF267941 | SMNH130506 |
| <i>Rhyacodrilus falciformis</i> | CE795-2 | F | Wet soil, Hälleklis, Götene, SE | May 2004 | KF267927 | KF267991 | KF267942 | SMNH130507 |
| <i>Rhyacodrilus falciformis</i> | CE795-3 | F | Wet soil, Hälleklis, Götene, SE | May 2004 | KF267928 | – | KF267943 | SMNH130508 |
| <i>Rhyacodrilus falciformis</i> | CE836 | F | Säveån River, Lerum, SE | June 2004 | KF267929 | KF267988 | KF267944 | SMNH130509 |
| <i>Rhyacodrilus falciformis</i> | CE1203 | D | Swamp, Rya skog, Hisingen, SE | November 2005 | KF267910 | – | KF267945 | SMNH130510 |
| <i>Rhyacodrilus falciformis</i> | CE3868 | F | Aspen Lake, Lerum, SE | April 2008 | KF267913 | KF267975 | KF267956 | SMNH130511 |
| <i>Rhyacodrilus falciformis</i> | CE3869 | F | Aspen Lake, Lerum, SE | April 2008 | KF267914 | – | KF267946 | SMNH130512 |
| <i>Rhyacodrilus falciformis</i> | CE3870 | F | Aspen Lake, Lerum, SE | April 2008 | KF267915 | KF267976 | KF267947 | SMNH130513 |
| <i>Rhyacodrilus falciformis</i> | CE4387 | G | Stream, Omberg, Stocklycke, Ödeshög, SE | July 2008 | KF267916 | KF267977 | KF267948 | SMNH130514 |
| <i>Rhyacodrilus falciformis</i> | CE4388 | G | Stream, Omberg, Stocklycke, Ödeshög, SE | July 2008 | KF267917 | KF267978 | KF267949 | SMNH130515 |
| <i>Rhyacodrilus falciformis</i> | CE4563 | F | Stream, Frösön, Östersund, SE | July 2008 | KF267918 | KF267989 | KF267950 | SMNH130516 |
| <i>Rhyacodrilus falciformis</i> | CE4564 | G | Stream, Frösön, Östersund, SE | July 2008 | KF267919 | KF267980 | KF267951 | SMNH130517 |
| <i>Rhyacodrilus falciformis</i> | CE6611 | F | Wet soil, Uppsala Botanical Garden, SE | June 2009 | IN260080 ^a | – | – | SMNH130518 |
| <i>Rhyacodrilus falciformis</i> | CE6612 | F | Wet soil, Uppsala Botanical Garden, SE | June 2009 | KF267921 | KF267979 | KF267957 | SMNH130519 |
| <i>Rhyacodrilus falciformis</i> | CE6615 | D | Wet soil, Uppsala Botanical Garden, SE | June 2009 | KF267922 | – | KF267958 | SMNH130520 |
| <i>Rhyacodrilus falciformis</i> | CE6631 | E | Prästjärden Lake, Åkersberga, SE | June 2009 | KF267923 | KF267987 | KF267952 | SMNH130521 |
| <i>Rhyacodrilus falciformis</i> | CE6632 | E | Prästjärden Lake, Åkersberga, SE | June 2009 | KF267924 | – | KF267953 | SMNH130522 |
| <i>Rhyacodrilus falciformis</i> | CE7737 | F | Stream, Guldheden, Göteborg, SE | May 2010 | KF267925 | – | KF267954 | SMNH130523 |
| <i>Rhyacodrilus falciformis</i> | CE8714 | C | Borehole, Abbotshood Farm, Devon, GB | January 2010 | KF267930 | KF267983 | KF267955 | SMNH130524 |
| <i>Rhyacodrilus falciformis</i> | CE10955 | C | Spring, Greatcombe Head Farm, Devon, GB | January 2010 | KF267901 | – | KF267959 | SMNH130525 |
| <i>Rhyacodrilus falciformis</i> | CE11311 | G | Ögonakällan Spring, Östansjö, Hallsberg, SE | April 2011 | KF267902 | KF267981 | KF267960 | SMNH130526 |
| <i>Rhyacodrilus falciformis</i> | CE11339 | G | Spring, Öglunda, Skara, SE | April 2011 | KF267903 | KF267982 | KF267961 | SMNH130527 |
| <i>Rhyacodrilus falciformis</i> | CE11351 | F | Stream, Guldheden, Göteborg, SE | April 2011 | KF267904 | – | KF267962 | SMNH130528 |
| <i>Rhyacodrilus falciformis</i> | CE11366 | F | Ditch, Guldheden, Göteborg, SE | April 2011 | KF267900 | – | KF267963 | SMNH130529 |
| <i>Rhyacodrilus falciformis</i> | CE11405 | E | Stream, Vitsippsdalen, Göteborg, SE | March 2011 | KF267905 | – | KF267970 | SMNH130530 |
| <i>Rhyacodrilus falciformis</i> | CE11406 | E | Stream, Vitsippsdalen, Göteborg, SE | March 2011 | KF267906 | KF267984 | KF267964 | SMNH130531 |
| <i>Rhyacodrilus falciformis</i> | CE11451 | F | Ditch, Guldheden, Göteborg, SE | March 2011 | KF267907 | KF267992 | KF267965 | SMNH130532 |
| <i>Rhyacodrilus falciformis</i> | CE11452 | D | Ditch, Guldheden, Göteborg, SE | March 2011 | KF267908 | KF267985 | KF267966 | SMNH130533 |
| <i>Rhyacodrilus falciformis</i> | CE11489 | E | Ground water seepage, Brännäs, Lerum, SE | April 2011 | KF267909 | KF267986 | KF267967 | SMNH130534 |
| <i>Rhyacodrilus pigueti</i> sp. n. | CE14048 | A | Crotot Cave, Besançon, Dep. Doubs FR | January 2011 | KF267911 | KF267993 | KF267968 | SMNH Type-8475 ^b |
| <i>Rhyacodrilus pigueti</i> sp. n. | CE14049 | A | Crotot Cave, Besançon, Dep. Doubs FR | January 2011 | KF267912 | KF267994 | KF267969 | No voucher |
| <i>Rhyacodrilus falciformis</i> | CE16690 | B | Jivero Cave, Matienzo, Cantabria, ES | October 2012 | KF267932 | – | KF267938 | SMNH130535 |
| <i>Rhyacodrilus falciformis</i> | CE16691 | B | Jivero Cave, Matienzo, Cantabria, ES | October 2012 | KF267931 | – | KF267937 | SMNH130536 |
| <i>Rhyacodrilus falciformis</i> | CE16692 | B | Jivero Cave, Matienzo, Cantabria, ES | October 2012 | KF267933 | – | KF267936 | SMNH130537 |
| <i>Rhyacodrilus coccineus</i> | CE623 | – | Stream, Alingsås, SE | July 2003 | GU902110 ^c | KF267996 | KF267971 | No voucher |
| <i>Rhyacodrilus subterraneus</i> | CE10572 | – | Spring, Willeleigh Farm, Devon, GB | January 2010 | KF267934 | KF267995 | KF267972 | SMNH130538 |

SE, Sweden; GB, Great Britain; FR, France; ES, Spain.

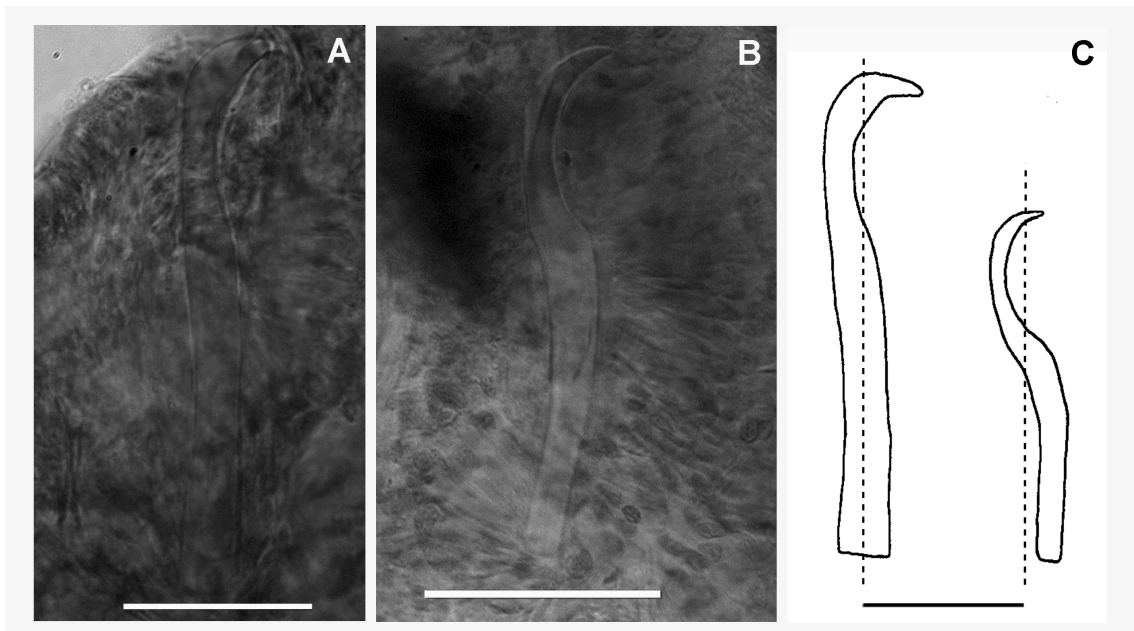
^aBOLD accession no: ENSWD081-11, ^bHolotype, ^cFrom Erséus et al. (2010).

Figure legends

Figure 1. Distribution of the genetic variation within *Rhyacodrilus falciformis* s.l. **A** Histogram of pairwise COI distances given in both uncorrected p-distances and corrected TN93 distances. **B-C** Gene-trees estimated with Bayesian coalescent analysis in BEAST. Letters right of specimen codes represent COI cluster, if clusters are found monophyletic is a line followed by the letter shown right of the specimens. Letter A represent *Rhyacodrilus pigueti* sp.n. and B-G *R. falciformis* s.s. Numbers above branches are posterior probability only values above 0.50 are given. Scale shows expected numbers of substitution per site. **B** COI gene tree. **C** ITS gene tree. **D** H3 gene tree.

Figure 2. Differences in the penial chaetae between *Rhyacodrilus pigueti* sp.n. and *R. falciformis*. **A** *Rhyacodrilus pigueti*, Holotype. **B** *Rhyacodrilus falciformis*, **C** Drawings of the penial chaetae in *R. pigueti* (straight chaeta) and *R. falciformis* (sickle-shaped chaeta). Scale bars = 50 μ m.

Figure 3. Summary of the results, indicating the amount of support for each COI cluster. For COI, ITS and H3 clusters that were monophyletic are indicated with a grey bar, pale grey indicates a *Rosenberg's* $P_{AB} \leq 0.05$ and dark grey both a *Rosenberg's* $P_{AB} \leq 0.05$ and a $P(\text{Randomly Distinct}) \leq 0.05$. Cluster A represent *Rhyacodrilus pigueti* sp.n. and clusters B-G *R. falciformis* s.s.



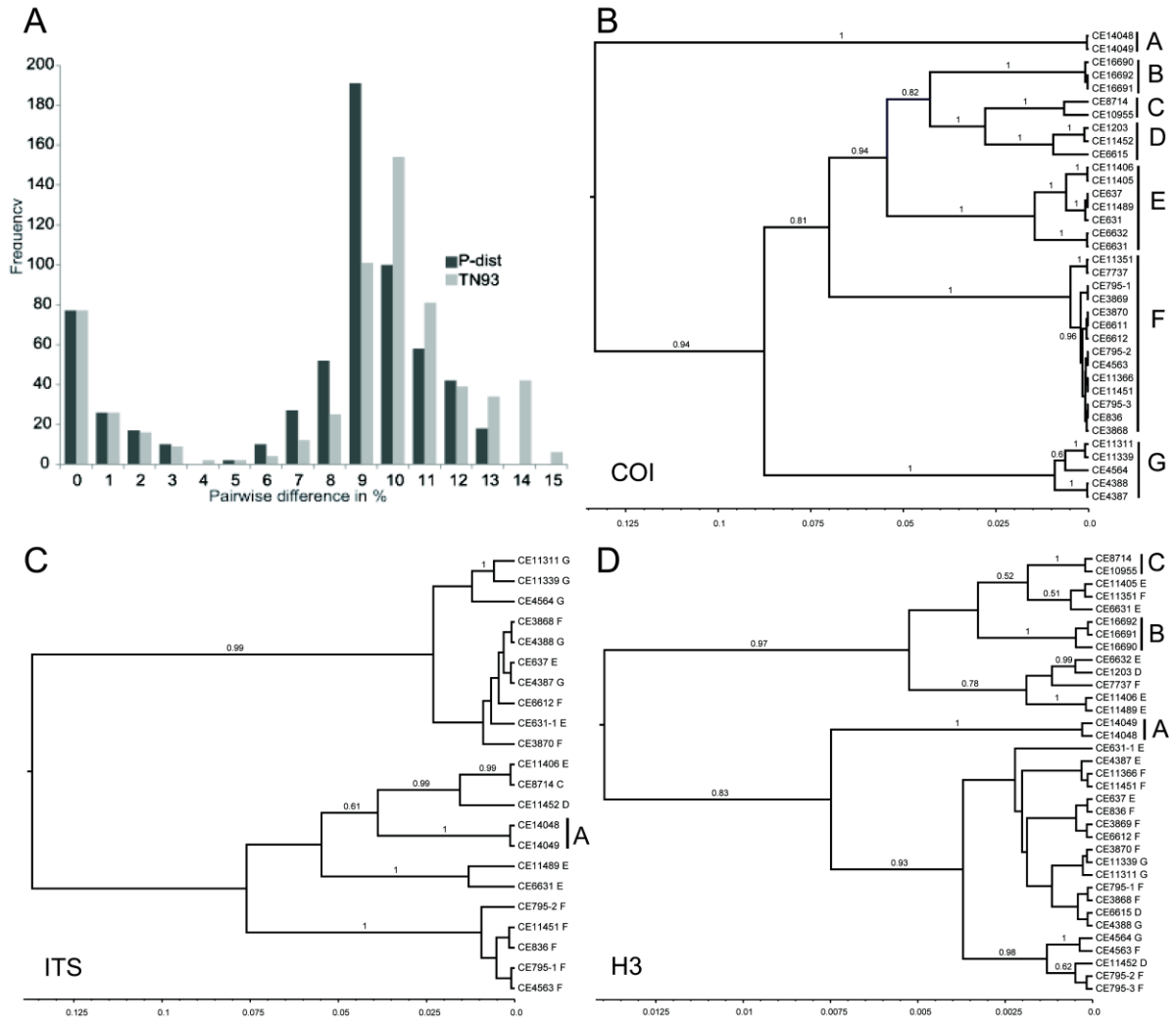


Figure 1. Distribution of the genetic variation within *Rhyacodrilus falciformis* s.l. **A** Histogram of pairwise COI distances given in both uncorrected p-distances and corrected TN93 distances. **B-C** Gene-trees estimated with Bayesian coalescent analysis in BEAST. Letters right of specimen codes represent COI cluster, if clusters are found monophyletic a line followed by the letter shown right of the specimens. Letter A represent *Rhyacodrilus aemilii* sp. n. and B-G *R. falciformis*. Numbers above branches are posterior probability only values above 0.50 are given. Scale shows expected numbers of substitution per site. **B** COI gene tree. **C** ITS gene tree. **D** H3 gene tree.

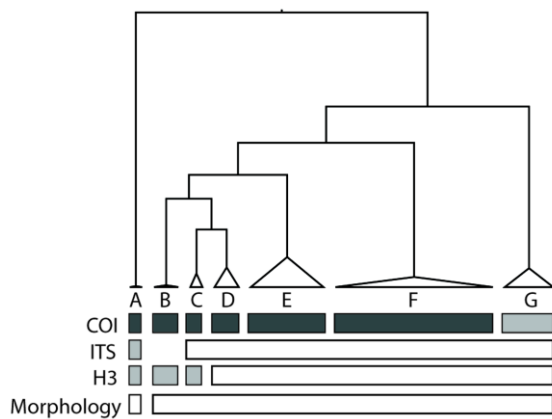


Figure 2. Summary of the results, indicating the amount of support for each COI cluster. For COI, ITS and H3 clusters that were monophyletic are indicated with a grey bar, pale grey indicates a *Rosenberg's* $P_{AB} \leq 0.05$ and dark grey both a *Rosenberg's* $P_{AB} \leq 0.05$ and a $P(\text{Randomly Distinct}) \leq 0.05$. Cluster A represent *Rhyacodrilus aemilii* sp. n. and clusters B-G *R. falciformis*.

Supplementary Tables

Table S1. Primers, sequences and programs used for amplification of the mitochondrial COI and nuclear ITS and H3.

| Gene | Primer | Sequence | Reference | Amplification program |
|------|----------|----------------------------|--------------------------------|--|
| COI | LCO1490 | GGTCAACAAATCATAAAGATATTGG | Folmer <i>et al.</i> (1994) | 95°C for 5 min, 35 cycles each of 95°C for 40 sec, 45°C for 45 sec and 72°C for 60 sec, finally, 72°C for 8 min. |
| | HCO2198 | TAAACTTCAGGGTGACCAAAAATCA | Folmer <i>et al.</i> (1994) | |
| | COI-E | TATACTTCTGGGTGTCCGAAGAATCA | Bely and Wray (2004) | |
| ITS | ITS-5 | GGAAGTAAAAGTCGTAACAAGG | White <i>et al.</i> (1990) | 95°C for 5 min, 35 cycles each of 95°C for 30 sec, 50°C for 30 sec and 72°C for 90 sec finally, 72°C for 8 min. |
| | ITS-4 | TCCTCCGCTTATTGATATGC | White <i>et al.</i> (1990) | |
| | 5.8mussF | CGCAGCCAGCTGCGTGAATTAATGT | Källersjö <i>et al.</i> (2005) | |
| | 5.8mussR | GATGTTCGATGTTCAATGTGTCCTGC | Källersjö <i>et al.</i> (2005) | |
| H3 | H3F | ATGGCTCGTACCAAGCAGACVGC | Brown <i>et al.</i> (1999) | 95°C for 5 min, 35 cycles each of 95°C for 30 sec, 50°C for 30 sec and 72°C for 90 sec, finally, 72°C for 8 min. |
| | H3R | ATATCCTTRGGCATKATRGTGAC | Brown <i>et al.</i> (1999) | |

Table S2. Details of alignments of the mitochondrial COI and the nuclear H3 and ITS genes.

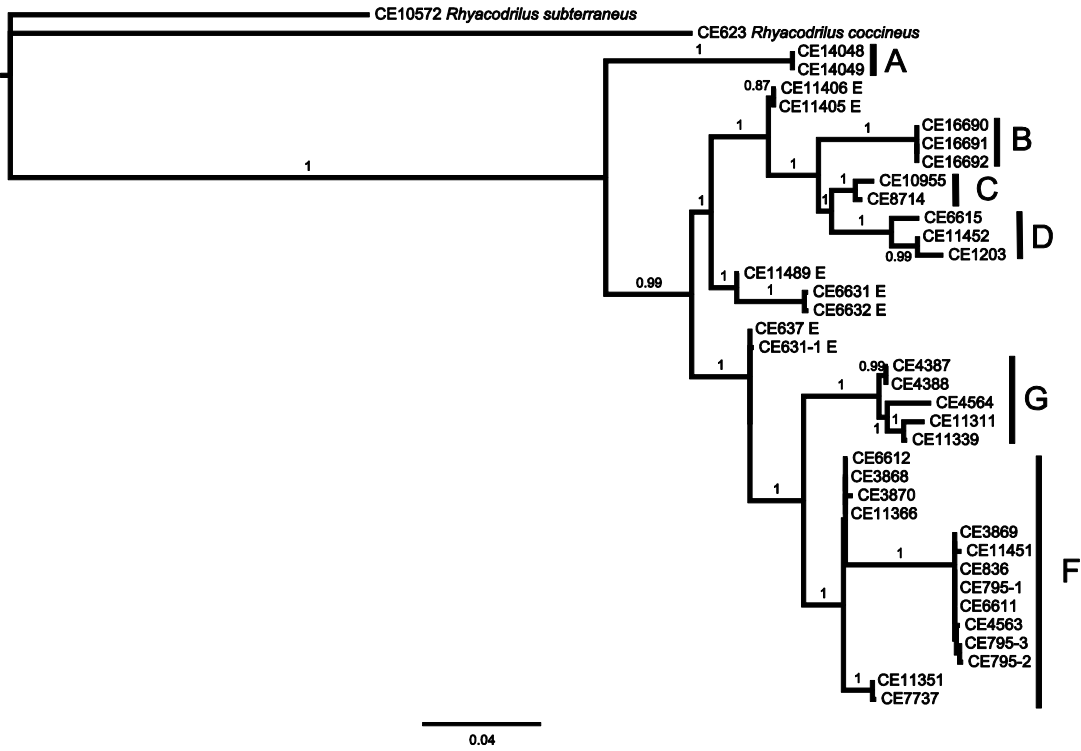
| Gene | No. of specimens | Length of alignment | AT content | Variable positions |
|------|------------------|---------------------|------------|--------------------|
| COI | 33 | 658 | 65.8 % | 147 |
| H3 | 32 | 328 | 37.4 % | 19 |
| ITS | 22 | 1115 | 47.8% | 489 |

Table S3. Ranges (μm) for morphological characters of the COI-clusters (A-G). Cluster A corresponds to *Rhyacodrilus pigueti* sp. n., clusters B-G to *R. falciformis* s.s.. n = number of measured specimens. Only those specimens that were completely mature were selected for the morphological analysis. Ranges among different groups are compared without any correction for the body size due to the overlap in range values of the width of segment XI.

| COI cluster | | A | | B | | C | | D | | E | | F | | G | |
|----------------------------|--|-----------------------|---|-------|---|-------|---|-------|----|---------|-------|---------|----|-------|---|
| | | range | n | range | n | range | n | range | n | range | n | range | n | range | n |
| internal characters | Length of atrium | 150 (ampulla only) | 1 | - | - | - | - | - | - | 180 | 1 | 192-240 | 3 | 260 | 1 |
| | Maximum diameter of atrium | 50 | 1 | - | - | 88 | 1 | - | - | 52 | 1 | 56-92 | 7 | 44-76 | 2 |
| | Thickness of muscular layer of atrial ampulla | 10 | 1 | - | - | 8 | 1 | - | - | 8 | 1 | 6-12 | 5 | 12 | 2 |
| | length of spermathecal ampulla | 60 | 1 | - | - | - | - | - | - | 56 | 1 | 56-68 | 5 | 64 | 1 |
| | Maximum diameter of spermathecal ampulla | 96 | 1 | - | - | - | - | - | - | 60 | 1 | 76-112 | 8 | 116 | 1 |
| | length of spermathecal canal | 80 | 1 | - | - | - | - | - | - | 32 | 1 | 36-72 | 5 | 80 | 1 |
| | Maximum diameter of spermathecal canal | 60 | 1 | - | - | - | - | - | - | 52 | 1 | 36-72 | 5 | 72 | 1 |
| | Length of penial chaeta | 150 | 1 | 120 | 1 | 112 | 1 | - | - | 100-110 | 2 | 118-140 | 9 | - | - |
| | maximum diameter of Penial chaeta | 15 | 1 | 8 | 1 | 12 | 1 | - | - | 10 | 2 | 16-okt | 9 | - | - |
| Diameter of coelomocytes | 27 | 1 | - | - | - | - | - | - | 28 | 1 | 24-32 | 4 | 20 | 1 | |
| somatic chaetae | Number of preclitellar ventral chaetae per bundle | 3-4 | 1 | 3-4 | 1 | 6 | 1 | 4-5 | 1 | 3-5 | 6 | 2-5 | 14 | 3-5 | 3 |
| | Maximum length of preclitellar ventral chaetae | 88 | 1 | 80 | 1 | 76 | 1 | 64-80 | 1 | 56-80 | 4 | 60-84 | 9 | 76-84 | 3 |
| | Number of preclitellar dorsal chaetae per bundle | 3-5 | 1 | 2-4 | 1 | 4 | 1 | 2-5 | 1 | 2-6 | 6 | 2-5 | 13 | 2-5 | 3 |
| | Maximum length of preclitellar dorsal chaetae | 92 | 1 | 88 | 1 | 76 | 1 | 68-80 | 1 | 60-84 | 4 | 56-92 | 10 | 68-88 | 3 |
| | Number of postclitellar ventral chaetae per bundle | - | - | - | - | - | - | 2-3 | 1 | - | - | 2-3 | 2 | 3-4 | 2 |
| | Number of postclitellar dorsal chaetae per bundle | 1-3 | 1 | 1 | 1 | - | - | 2-3 | 1 | 1-2 | 1 | 2-3 | 3 | - | - |
| | Maximum length of postclitellar chaetae | - | - | - | - | - | - | - | - | - | - | 56-72 | 2 | - | - |
| size of specimens | Width of segment XI | 380 | 1 | 264 | 1 | 336 | 1 | - | - | 260-372 | 3 | 256-480 | 9 | 288 | 1 |

References

- Bely, A. E., & Wray, G. A. (2004). Molecular phylogeny of naidid worms (Annelida: Clitellata) based on cytochrome oxidase I. *Mol Phylogenet Evol*, 30, 50-63.
- Brown, S., Rouse, G., Hutchings, P., & Colgan, D. (1999). Assessing the usefulness of histoneH3, U2 snRNA and 28S rDNA in analyses of polychaete relationships. *Australian Journal of Zoology*, 47, 499-516.
- 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. *Mol Mar Biol Biotechnol*, 3, 294-299.
- Källersjö, M., von Proschwitz, T., Lundberg, S., Eldenäs, P., & Erséus, C. (2005). Evaluation of ITS rDNA as a complement to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using Unionidae from north-western Europe. *Zoologica Scripta*, 34, 415-424.
- White, T., Bruns, T., Lee, S., & Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. A. Innis, D. H. Gelfand, J. J. Sninsky & T. J. White (Eds) *PCR protocols: a guide to methods and applications* pp. 315-322. San Diego, Academic Press.



Supplementary Fig. S1. Maximum Likelihood tree based on combined COI, H3 and ITS data from *Rhyacodrilus falciformis s.l.* obtained with PhyML, numbers at branches are aLRT branch support. Letters right of specimen codes represent COI cluster, if clusters are found monophyletic is a line followed by the letter shown right of the specimens. Letter A represent *Rhyacodrilus pigueti sp.n.* and B-G *R. falciformis s.s.* Scale shows expected numbers of mutations per site.