JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

Accepted on 12 February 2010
J Zool Syst Evol Res doi: 10.1111/j.1439-0469.2010.00571.x

Department of Zoology, University of Gothenburg, Göteborg, Sweden

Genetic variation and phylogeny of Scandinavian species of *Grania* (Annelida: Clitellata: Enchytraeidae), with the discovery of a cryptic species

Pierre De Wit and Christer Erséus

Abstract

Individuals of five nominal species of *Grania* (Annelida: Clitellata: Enchytraeidae) were collected from locations in Sweden, Norway and France, for studies on the intraspecific variation at the Cytochrome Oxidase I (COI) locus of mitochondrial DNA and internal transcribed spacer (ITS) region of nuclear DNA. It was found that the previously described morphospecies in general contain low variation compared to the between-species variation in both loci. In one instance, however, an individual morphologically indistinguishable from *G. ovitheca* was found to be deviant and instead cluster with *G. postclitellochaeta* both by COI and ITS. We describe this individual as a new species: *G. occulta* sp.n. Furthermore, phylogenetic analyses were conducted, showing a close relationship between *G. variochaeta*, *G. occulta*, *G. ovitheca* and *G. postclitellochaeta*, as well as between *G. pusilla* and *G. maricola*. Using the results from the phylogenetic analyses, we discuss the evolution of morphological characters in Scandinavian species of *Grania*.

Key words: cryptic species – cytochrome oxidase I – DNA barcoding – *Grania* – internal transcribed spacer – intraspecific variation – phylogeny – species delimitation

Introduction

In the ocean floor, great diversity is to be found, both in habitats and in the number of species inhabiting it (Gray 1997). Many animals inhabiting this environment are small, however, and a large proportion of these are poor in external characters, making them hard to morphologically identify to species level without in-detail studies in microscopes. To distinguish these small species, molecular methods have been shown to be helpful.

In particular, the concept of DNA barcoding, where one standard region of the genome is sequenced to identify organisms, has been shown to be valuable (Hebert and Gregory 2005; Savolainen et al. 2005). A critical feature of any region used as a 'barcode' in species identification is that the intraspecific variation must be significantly lower than the interspecific variation; a so-called 'barcoding gap' must exist (Hebert et al. 2003b; Moritz and Cicero 2004). The mitochondrial Cytochrome Oxidase I (COI) locus of about 650 base pairs has been proposed as this standard region (Hebert et al. 2003b). Although not universally applicable (Duran et al. 2004; Wörheide 2006; Shearer and Coffroth 2008), COI has been shown to distinguish morphological species well in most animal taxa (Hebert et al. 2003a,b; Moritz and Cicero 2004; Hebert and Gregory 2005; Savolainen et al. 2005; Smith et al. 2005; Vences et al. 2005; Huang et al. 2007; Gustafsson et al. 2009). Critics, however, claim that the barcoding gap vanishes as more extensive sampling is conducted over a broader geographic region (Meyer and Paulay 2005). Moreover, DNA barcoding could be difficult to apply to populations which have recently undergone speciation events and have not yet had the time to diverge genetically to the point where species distinction is possible (Moritz and Cicero 2004; Will et al. 2005; Hickerson et al. 2006).

Despite these issues, variation in the COI gene seems promising for the study of small animals for a number of reasons. First of all, many of these organisms have complicated life histories, with larval phases that are indistinguishable at species-level. To be able to recognize the species, identity of larvae would greatly facilitate the study of biodiversity. Similarly, young adult individuals that have not yet developed taxonomically useful characters, such as species-specific sexual structures, could be identified by DNA sequencing. Furthermore, we are beginning to realize that many small, inconspicuous species cannot be separated morphologically at all from each other, yet their DNA shows that they clearly are separately evolving lineages (e.g. King et al. 2008; Gustafsson et al. 2009; Nygren et al. 2009; Wiklund et al. 2009). This has serious implications for many aspects of biology. It has recently been shown that several common model organisms actually are composed of several species, which makes it hard to draw specific conclusions from studies using these organisms (Siddall et al. 2007; Erséus and Gustafsson 2009). In some cases, the lineages have been shown to differ significantly in their ecology, e.g. transmission of parasites (Beauchamp et al. 2002; Hallett et al. 2009) and physiology, e.g. heavy metal resistance (Sturmbauer et al. 1999).

The species concept is still under considerable debate, and there are a number of proposed species definitions, i.e. biological (Wright 1940; Mayr 1942), ecological (Van Valen 1976), evolutionary (Simpson 1951) as well as phylogenetic (Hennig 1966; Rosen 1979) ones. Most of these definitions, however, can be seen as different aspects of the same natural phenomenon: the fact that populations of organisms diverge from each other with time. Thus, it has recently been proposed that a single unifying concept of a species could be that each species constitutes a separately evolving metapopulation (De Queiroz 2007). Species hypotheses could then be corroborated by different criteria, e.g. by morphological, reproductive, ecological or genetic distinction (De Queiroz 2007). That a group of populations is found divergent using any of these criteria is evidence for its unique evolutionary history, and thus it could be seen as a distinct species even if it does not meet the criteria of all historically proposed species concepts. This is

Corresponding author: Pierre De Wit (pierre.de_wit@zool.gu.se)
Contributing author: Christer Erséus (christer.erseus@zool.gu.se)

particularly important for lineages which have only recently been separated, where DNA or morphological characters have not yet had the time to diverge to the point where species distinction is possible.

Grania Southern, 1913 is a genus of Clitellata (Annelida) inhabiting marine sediments, from the littoral zone to the deep sea. Species of the genus are small, whitish or transparent worms about 1–2 cm long and only about 0.1 mm thick. *Grania* is morphologically well delimited (Erséus and Lasserre 1976): In all species, segments I–IV are fused into a 'head', and chaetae are absent at least in I-III (often in more segments). In segments where they are present, chaetae are rather stout and occur one by one, as opposed to most clitellates which have bundles of two or more chaetae. Morphological identification of individual species, however, is complicated: it is necessary to stain, clear and mount specimens on microscope slides, as the commonly used characters are the details of the internal sexual organs and of the morphology and arrangement of the chaetae. Based on this, six species have been recognized from the coasts of Scandinavia: *G. maricola* Southern, 1913; *G. pusilla* Erséus, 1974; *G. postclitellochaeta* (Knöllner, 1935), *G. variochaeta* Erséus and Lasserre, 1976; *G. ovitheca* Erséus, 1977 and *G. vikinga* Rota and Erséus, 2003 (the last-mentioned taxon was initially identified as *G. roscoffensis* Lasserre, 1967; see Erséus 1977). Most of these species have distributions that range from the west coast of Norway in the north, to outside of and sometimes inside the Mediterranean Sea in the south (Rota and Erséus 2003). The exception is *G. vikinga*, which so far has only been found on the west coast of Sweden (Rota and Erséus 2003).

Genetically, however, little is known about *Grania*. It is not known if the morphologically distinguishable species correspond to phylogenetic species, or if any of the nominal species actually consists of several separately evolving lineages.

Here, we use Bayesian inference and maximum likelihood of mitochondrial and nuclear genetic data to statistically test the monophyletic status of five of the morphospecies that inhabit the waters of Scandinavia. To assess how DNA barcoding would work on Scandinavian species of *Grania*, we study the intraspecific and interspecific variation at the COI locus, to determine whether there are barcoding gaps present between the current taxa. A new evolutionary lineage is found based both on COI distances and on divergence at the internal transcribed spacer (ITS) region of the nuclear genome. We describe this lineage as *Grania occulta* sp.n., although it is morphologically indistinguishable from *G. ovitheca*. We also examine the phylogenetic relationships and character evolution among the studied species.

Material and Methods

Collection and identification

Grania specimens were sampled at the west coast of Sweden on numerous occasions in 2000–2008 (Table S1), in the Koster archipelago in the north, the area around the Gullmar fjord some 150 km to the south (both in the Skagerrak) and also around the shallow sand banks of Fladen and Lilla Middelgrund further south in Kattegatt. The area near Bergen, Norway, was sampled by the first author in November 2008. Sediment samples were collected by dredging from a research vessel, and then stirred with sea water followed by decantation through a 0.25-mm mesh-sized sieve. The sieved fractions were sorted using a stereo microscope. Individual worms were killed in 80% ethanol, after which they were cut in half, with the front end fixed in either Bouin's fluid, formalin, or 80% ethanol, stained in alcoholic paracarmine and mounted whole in Canada balsam on slides for identification purposes, whereas the rear end was stored in 95% ethanol for subsequent DNA extraction and sequencing. Additional material of *G. variochaeta* was found in sediment samples collected by hand by the first author in the littoral zone of Brittany, France, in 2007, and treated in the same fashion. The front ends of all *Grania* specimens were identified to the species level by the first author using light transmission and interference contrast microscopy. Type specimens and other materials of *G. ovitheca* from the Swedish Museum of Natural History (SMNH) in Stockholm were also studied in the same fashion. All new specimens were deposited as voucher specimens at SMNH. Drawings of *Grania occulta* sp.n. were made using a camera lucida.

DNA analysis

The DNA of the rear ends of a total of 87 specimens, identified as belonging to five morphospecies (Table S2), was extracted using a Qiagen DNeasy® Blood and Tissue kit, after which PCR was performed using standard COI barcoding primers LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) (forward) and HCO2198 (5′-TAACTTCAGGGTGACCAAAAAATCA-3′) (reverse) (Folmer et al. 1994). The PCR reactions consisted of an initial step of 5 min at 95°C, followed by 35 cycles, each of 40 s at 95°C, 45 s at 45°C and 60 s at 72°C. This was followed by a final step of 8 min at 72°C. The resulting extracts were purified using an Omega E.Z.N.A. cycle-pure kit, and then sent to Macrogen corp., South Korea, for ABI sequencing.

In addition, 39 individuals, of which 35 among the 87 mentioned above (see Table S2) were sequenced at the nuclear ITS region, including ITS1, 5.8S rDNA and ITS2. PCR reactions were performed using primers ITS 5 (5′-GGAAGTAAAAGTCGTAACAAGG-3′) (forward) and ITS 4 (5′-TCCTCCGCTTATTGATATGC-3′) (reverse) (White et al. 1990), with an initial step of 5 min at 95°C, followed by 35 cycles of 30 sec at 95°C, 30 sec at 50°C and 90 s at 72°C, after which a final step of 8 min at 72°C was performed. In the sequencing PCR reaction (performed by Macrogen corp.), primers 5.8 SF (5′-CGCAGCCAGCTGCGTGAATTAATGT-3′) and 5.8 SR (5′-GATGTCGATGTTCAATGTGTCCTGC-3′) (Källersjö et al. 2005), which bind to 5.8S and amplify ITS2 (forward) and ITS1 (reverse), respectively, were also used.

For outgroup comparison, the COI and ITS regions of one individual each of the species *G. laxartus* Locke and Coates, 1999 and *G. monospermatheca* Erséus and Lasserre, 1976; from Australia, Bahamas and Florida, respectively, which are likely to be close relatives of the Scandinavian species (De Wit et al., in prep), were also amplified and sequenced (Table S2).

All new sequences were assembled using Geneious Pro 4.6.4 (Rozen and Skaletzky 2000) from Biomatters Ltd and visually examined for ambiguous readings.

For sequence alignment, MAFFT version 6 (Katoh and Toh 2008) was used, applying the L-INS-i setting (slow-accurate). The genetic variation was assessed in PAUP*4.0b10 (Swofford 2002) using the 'pairwise distance' application. For each sequence, the Kimura 2 parameter (K2P) distance to each of the other sequences was calculated, ignoring sites with missing data for the pairwise comparison. Mean intraspecific distances and standard deviations, as well as between-species distances, were calculated using Microsoft Excel 2003. To study possible substitution saturation in COI, the transition- transversion ratio (Ti/Tv) was plotted against the K2P distance using DAMBE (Xia 2009) and Microsoft Excel. A third matrix was also created by separately aligning COI and ITS sequences for those individuals from which both regions were available, then concatenating the two alignments into one.

Bayesian inference analyses were performed on the three alignments. To do this, the COI alignment was first partitioned according to codon position, and the model of best fit was chosen using the Akaike Information Criterion (AIC) implemented by MrModeltest 2.2 (Nylander 2004). The models were determined to be SYM + I for the first codon position, F81 + I for the second, and GTR + G for the third. For the ITS alignment, MrModeltest determined that GTR + I + G was the most appropriate model. In the combined

analysis, the alignment was partitioned both after locus and codon position (in the COI region of the alignment). The two loci were also unlinked from each other in all parameters.

In the parallel version of MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003), two separate MCMCMC analyses were run for each alignment, each with 4 Markov chains (one cold and three hot), for 50 million generations, sampling once every 1000 generations. The resulting tree files were examined for convergence using the AWTY online software (Wilgenbusch et al. 2004; Nylander et al. 2008), and were subsequently summarized using a burn-in of 10 million generations for COI and 30 million generations for the ITS analysis, to get statistical support values for the clades. Support values were plotted on the estimate (ML-estimate) trees, which were extracted from the tree files. Separately, the two alignments were submitted to the CIPRES RAxML web portal (http://8ball.sdsc.edu:8889/cipres-web/Home.do) for ML bootstrap analyses (Stamatakis 2006). Bootstrap values were also plotted on the ML-estimate trees.

The trees were tested for congruence against the datasets (the COI tree versus the ITS dataset and vice versa) with the SH-test function in PAUP*4.0b (Shimodaira and Hasegawa 1999; Swofford 2002), using the GTR + G model of base substitution (with an empirically determined α for the gamma distribution of site rate variation) and empirical base frequencies.

Nucleotide and amino acid sequences were deposited into GenBank and the Barcode of Life Database (BoLD) along with information on voucher specimens (Table S2).

Results

Genetic variation and phylogenetic analysis

In the ML-estimate tree from the Bayesian inference analysis of the COI alignment (Fig. 1), all specimens of *G. variochaeta*, *G. maricola* and *G. pusilla* cluster together with their morphological conspecifics with minimal intraspecific variation; the same pattern within *G. variochaeta*, the variation is 0.13% ± 0.18 (0–0.68%) (12 individuals), within *G. maricola*, it is 0.03% ± 0.06 (0–0.31%) (21 individuals) and within *G. pusilla,* it is 0.16% ± 0.22 (0–1.08%) (24 individuals). The same pattern is seen in the ML-estimate tree of the analysis of the ITS region (Fig. 2). Both of the trees support each of these three species with a posterior probability (PP) of 1.

The *G. postclitellochaeta* clade is also supported by COI with a PP of 1 (Fig. 1), but there is slightly more variation (2.05% ± 3.04, 13 individuals), mainly because of one individual with many unique base substitutions (PDW151). The mean K2P distance between this individual and the remaining *G. postclitellochaeta* clade is 8.45%. If PDW151 is excluded, the intraspecific variation drops to 0.77% ± 0.39 (0–1.58%) (12 individuals). In the ITS ML-estimate tree, however, there are no differences between this individual and the other specimens of *G. postclitellochaeta* (Fig. 2); also, here *G. postclitellochaeta* is supported by a PP of 1.

Within the morphospecies *G. ovitheca*, 15 of the 17 individuals cluster together in a clade with an intraspecific variation of 0.13% ± 0.23 (0–0.85%) in COI. Two specimens are clearly different, however. First, CE699 is separated from the main *G. ovitheca* clade by a mean K2P distance of 6.27% in COI, but in the ITS region, this worm is identical to several of the other *G. ovitheca* individuals. The second worm, PDW52 (hereafter referred to as *Grania occulta* sp.n.), which was originally morphologically identified as *G. ovitheca*, clusters instead with the *G. postclitellochaeta* clade in the COI ML-estimate tree with a PP of 1 (Fig. 1), although with a mean K2P distance of 19.6%. The mean distance between PDW52 and other *G. ovitheca* individuals is 22.8%. The same pattern is visible in the ITS region, where PDW52 (*G. occulta*) also clusters with *G. postclitellochaeta* with a PP of 1 (Fig. 2) and a mean K2P distance to the latter clade of 5.35%. In COI, the between-species (n = 6, including *G. occulta*) mean K2P distance is 21.07% ± 1.29.

All three phylogenetic analyses produced congruent results. In both single-gene analyses, the Scandinavian *Grania* form a monophyletic clade with respect to the North American outgroups. The clade *G. postclitellochaeta* + *G. occulta* is supported by PP = 1. Other than this, the COI analysis finds no statistical support for any inter-species relationships (Fig. 1), indicating divergence saturation. In the analysis of the more slow-evolving ITS region, however (Fig. 2), *G. ovitheca* and *G. postclitellochaeta* form a clade with PP = 0.99. Furthermore, this clade groups with *G. variochaeta* with PP = 0.96. *Grania maricola* and *G. pusilla* group together with PP = 0.95. In the third, combined, analysis (result not shown), statistical support is found only for clades also supported in the COI analysis, further supporting that COI provides no information concerning between-species relationships.

Taxonomy

Grania occulta sp.n. (Fig. 3).

Holotype

Swedish Museum of Natural History type coll. TYPE-7844, posteriorly amputated whole-mounted specimen from just outside the Gullmar fjord: 58°12.7′N; 011°19.0′E, 10-25 m depth, shell sand with some mud; collected by P. De Wit, 26 April 2006; COI barcode sequence, GenBank acc. no. GU473645; ITS sequence, GenBank acc. no. GU473705.

Description

Body > 5.2 mm long, > 33 segments (posterior end used for DNA extraction); 157 μm wide at III, 81 μm at clitellum. Prostomium rounded, 78 μm wide, 67 μm long; epidermis 20 μm thick dorsally and ventrally, 13 μm anteriorly. Peristomium 131 μm wide at 1/2. Ventral chaetae commencing in XIII, lateral chaetae absent. Chaetae 65–80 μm long, longer posteriorly than anteriorly; chaetae somewhat thicker entally than ectally, L-shaped, entally bent into a 'foot' (20–25 μm long) with an angle of about 100 degrees between shaft and foot; foot with broad instep, prominent heel and curved sole, chaetal index (see Rota and Erséus 2003) 3.43 ± 0.352 (n = 4) (Fig. 3a). Epidermal gland cells inconspicuous. Clitellum 10 μm thick, starting in anterior of XII and extending to mid XIII, with transverse rows of granular gland cells interspersed with hyaline cells at a frequency of about 1 : 1, except near male pores where hyaline cells are absent and midventrally where gland cells are absent. Midventral copulatory gland not observed in XIV. Spermathecal pores lateral, located at a short distance from 4/5. Male pores located ventrolaterally in mid XII.

Brain posteriorly indented. Head organ absent. Pharyngeal glands in IV–VI; dorsal lobes present in IV–VI, ventral lobes absent in IV, present in V (2 pairs) and VI (2 pairs); not connected dorsally. First pair of nephridia at 7/8. Dorsal blood vessel appears to be commencing in XIX (but specimen crushed in XX-XXI). Chloragogen cells small (5–7 μm tall). Coelomocytes oval, about 8 × 15 μm, granular with unstained nucleus. Sperm sac extending posteriorly from clitellum as far back as XVII. Sperm funnels of uniform width, 30 μm wide,

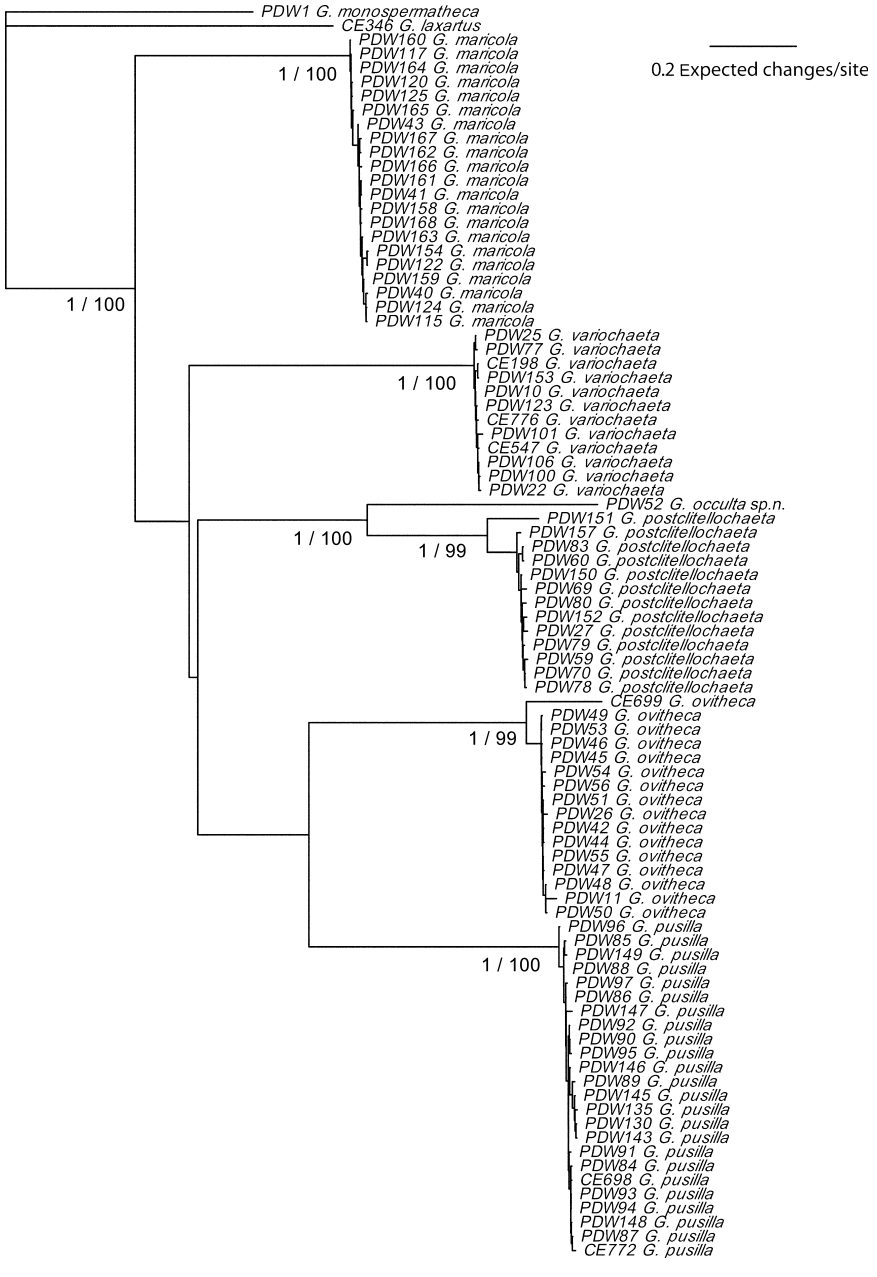

Fig. 1. ML-estimate tree from Bayesian inference analysis of the COI locus, with posterior probabilities (PP) ≥ 0.95 (first number) and maximum likelihood bootstrap values ≥ 80 (second number) noted

12–15 times as long as wide, folded with collar directed posteriorly in X–XI. Heads of spermatozoa 15 μm long. Vasa deferentia long, coiled in XII to XIV, 7.5 μm wide, internally ciliated. Penial apparatuses (Fig. 3b) with oval glandular structures, 40 μm long, 55 μm wide, next to epidermal invaginations; vasa deferentia opening into epidermal invaginations; stylets absent (penial bulb type 3 *sensu* Coates 1984). Egg sac extending at least to XIX (specimen crushed in XX–XXI). Spermathecae (Fig. 3c) attached to oesophagus in posterior of V through narrow ental ducts; ampullae egg-shaped, 55 × 30 μm, ectal ducts deeply incised into ampullae, 55 μm long, 18 μm wide near spermathecal pores, but for most part 14 μm wide; 15–20 sperm rings per spermatheca, 4–5 μm in diameter; no glands at spermathecal pores.

Remarks

As the individual PDW52 is separated from any other species used in this analysis by 5.35% in the nuclear ITS region (ITS1, 5.8S rRNA gene and ITS2) (excluding regions with gaps), it is clear that it belongs to a lineage, which has had no gene flow with any of the other lineages represented in this study in a long time. A long historical separation of PDW52 is also supported by the fact that its COI sequence is 20% or more different from the corresponding sequences of all other specimens.

Morphologically, this individual is similar to both *G. ovitheca* and *G. postclitellochaeta*. The shape of its chaetae (entally thicker with a prominent heel and broad instep) (Fig. 3a) is similar to what has previously been reported in these two species (Rota and Erséus 2003: Fig 10B; Erséus 1977). The elongate chaetal foot with a curved sole is slightly more similar to that of *G. ovitheca*. The chaetae of *G. occulta* are apparently slightly longer than the ones of *G. ovitheca*; the longest chaeta ever reported in *G. ovitheca* was 68 μm (Rota and Erséus 2003) and in the type material, the size range was 50–65 μm according to Erséus (1977).

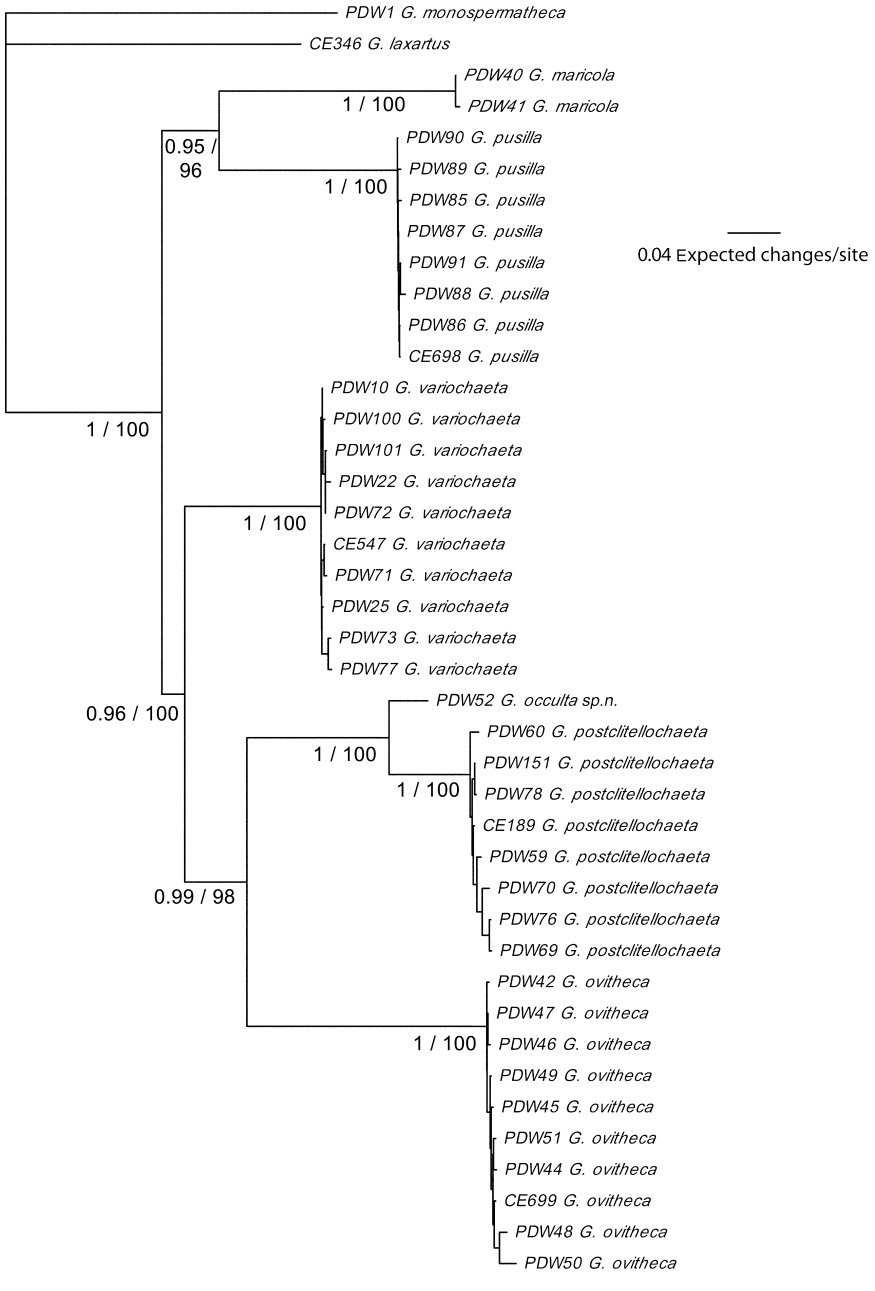

Fig. 2. ML-estimate tree from Bayesian inference analysis of the ITS region, with posterior probabilities (PP) ≥ 0.95 (first number) and maximum likelihood bootstrap values ≥ 80 (second number) noted

In *G. occulta,* the size range is from 65 to 80 µm. When re-examining the holotype of *G. ovitheca* (SMNH type coll. 3071), however, several chaetae with lengths above 70 µm were found in the posterior end of the body. This character can thus not be used for species distinction. *Grania postclitellochaeta* also possesses chaetae in the same size range (Rota and Erséus 2003). The chaetal distribution is identical to both previously described species, with ventral chaetae from segment XII, but no lateral chaetae.

A distinguishing feature of *G. ovitheca* is the absence of ventral lobes of the pharyngeal glands in IV (Rota and Erséus 2003), something which is shared by PDW52. *Grania postclitellochaeta*, however, does possess ventral lobes in IV. The spermathecae of *G. occulta* are identical to the ones of *G. ovitheca*, with egg-shaped ampullae deeply incised by the ectal ducts, containing numerous small sperm rings (Erséus 1977) and lateral spermathecal pores at a distance from 4/5 (Rota and Erséus 2003). The sperm funnels are very long, as originally described in *G. ovitheca* (Erséus 1977), and their collars are directed posteriorly. The penial apparatuses of *G. occulta* resemble those of *G. ovitheca*; they clearly lack the large glandular structures surrounding the penial bulbs of *G. postclitellochaeta* (see Coates 1984).

Thus, it is clear that *G. occulta* shares all distinguishing morphological features with *G. ovitheca*, but is clearly different from its closest relative, *G. postclitellochaeta*.

Etymology

The name *occulta* (= 'hidden') refers to the fact that this species is morphologically indistinguishable from *G. ovitheca*.

Discussion

Species delimitation

Despite the similarity to *G. ovitheca*, *G. occulta* sp.n. clusters with *G. postclitellochaeta* in all our analyses. As *G. ovitheca*

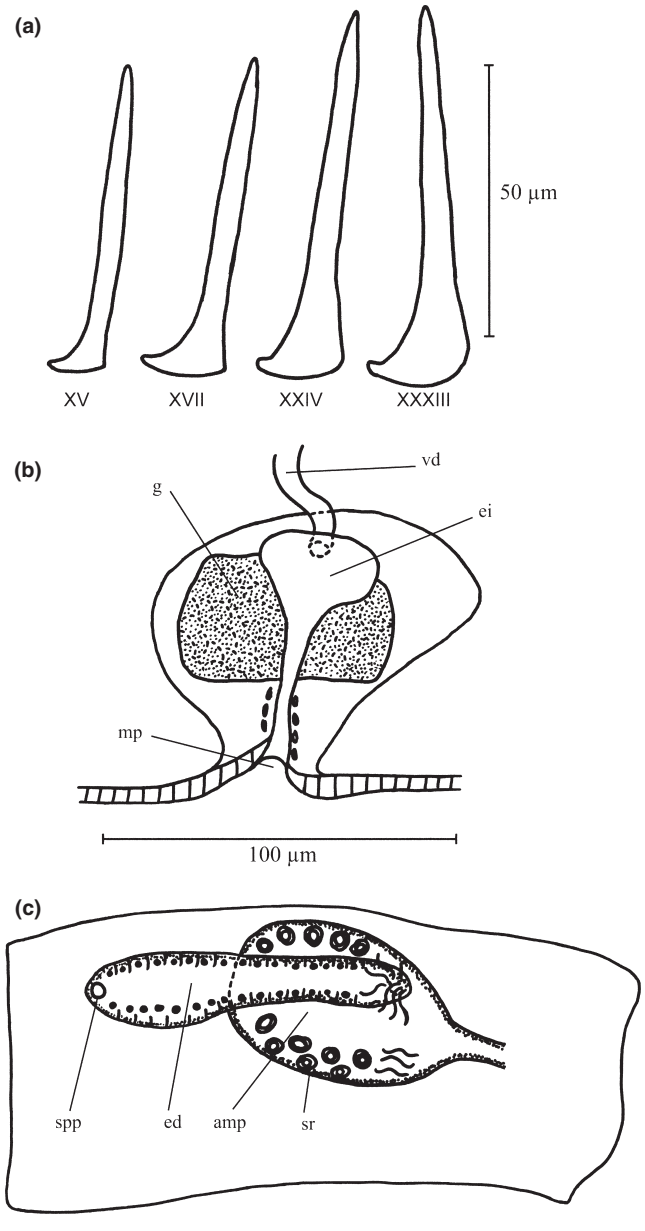

Fig. 3. *Grania occulta* sp.n. a: chaetae, b: penial apparatus, c: spermatheca. Abbreviations: mp, male pore; vd, vas deferens; g, gland; ei, epidermal invagination; spp, spermathecal pore; ed, ectal duct; amp, ampulla; sr, sperm ring. Roman numerals denote segment numbers

and *G. postclitellochaeta* + *G. occulta* come out as sister taxa in the ITS analysis, an interesting situation arises. *Grania ovitheca* (or rather, *G. postclitellochaeta*, which has nomenclatural priority) could be seen as the monophyletic taxon including all three above-mentioned lineages. This, however, is a viewpoint which would disagree with some phylogenetic species concepts (Vogler and Monaghan 2007), which define a species as the smallest resolvable separately evolving lineage. Further, as there are numerous morphological characters separating *G. ovitheca* and *G. postclitellochaeta*, we find that this would not be the best solution. It is also possible to see the morphospecies *G. ovitheca sensu lato* (i.e. including *G. occulta*) as a paraphyletic parent species of *G. postclitellochaeta*. The large difference in the nuclear ITS region of PDW52 (i.e. *G. occulta*), however, strongly suggests that the population that this individual belongs to has had no gene flow with *G. ovitheca sensu stricto* for a long time, and thus the two *G. ovitheca* lineages are separate species under both phylogenetic and biological species concepts. The third option, which we prefer, is to recognize three species in this clade, namely *G. ovitheca*, *G. postclitellochaeta* and *G. occulta*. In this way, we recognize monophyletic, separately evolving lineages as species. Evidence for the independent evolution of these species can be seen in their mitochondrial and nuclear genomes, despite the fact that they exist sympatrically (*G. ovitheca* and *G. occulta* were found in the same sample). Evidence is also found morphologically in *G. postclitellochaeta* being different from the other two.

Upon morphological comparison between PDW52 and the holotype of *G. ovitheca*, no distinguishing differences could be found. As no DNA is available from this holotype, it is impossible to objectively conclude which lineage in this study corresponds to *G. ovitheca sensu stricto*. The type locality of *G. ovitheca* is in the Koster archipelago on the west coast of Sweden. In this study, we have analysed one specimen of *G. ovitheca* from this location, CE 699, and it clusters with the lineage proposed to remain as *G. ovitheca* and not with *G. occulta*. This is seen as evidence for our taxonomic hypothesis, which is that the lineage which PDW52 belongs to is not that of the holotype of *G. ovitheca*. It is a problem that only one individual of the *G. occulta* population has been found to date, but more individuals will be found in the future, which might elucidate yet unnoticed morphological differences between *G. occulta* and *G. ovitheca*. It is possible, or even likely, that individuals of *G. occulta* have previously been identified as *G. ovitheca*, considering that our DNA samples do not cover the entire distributional range of this species complex. For example, *G. occulta* may prove to be a southern species, which partially overlaps with *G. ovitheca s.str.* in Scandinavia.

The two COI-divergent individuals of *G. ovitheca* and *G. postclitellochaeta*, CE699 and PDW151, respectively, are virtually identical to the remaining individuals of their respective morphospecies in the nuclear ITS region. This indicates that gene flow exists in both cases, and that these individuals are part of separate metapopulations, i.e. of *G. ovitheca* and *G. postclitellochaeta*, respectively.

What process would give rise to the observed pattern of divergence in the mitochondrial genome but not in the nuclear one? It is obvious that most marine animals inhabiting the coasts of Northern Europe have moved south during past glaciations and then immigrated back in interglacial periods. At this latter point in time, it is possible that introductions occurred from two or more refugia that were distinctly separated for long periods of time. For example, it is known that several marine refugia existed in Europe during the last glaciation, most notably the coast of the Iberian Peninsula and the Mediterranean, and also the coast of South-Western Ireland (Maggs et al. 2008). As animal populations from these two (or more) refugia merged in Scandinavia, nucleotide polymorphisms in the maternally inherited mitochondrial genome would be kept, whereas ones in the nuclear genome would tend to be smoothed over with time through recombination. *Grania variochaeta* individuals from Brittany, France (PDW100 and PDW101), are genetically so similar to the Scandinavian specimens that they clearly belong to the same original population. Unfortunately, although *G. pusilla* and *G. maricola* have been reported from Ireland, no genetic data

are available for *Grania* species from the British Isles at the moment.

Phylogeny and character evolution

According to the phylogeny inferred from the ITS data, *G. postclitellochaeta*, *G. ovitheca* and *G. occulta* form a monophyletic clade. Morphologically, the three species are similar. One diagnostic character of *G. postclitellochaeta*, however, is the presence of large glandular structures surrounding the penial bulbs (Knöllner 1935). In the light of the phylogeny inferred in this study, it seems likely that *G. postclitellochaeta*'s penial apparatus has evolved from a *G. ovitheca*-like arrangement, as both *G. ovitheca* and *G. occulta* possess identical male structures, but without the large glands (Fig. 4: character 1). Further, all three taxa lack cuticular penial stylets, while all other Scandinavian species of *Grania* do possess such stylets. The phylogeny thus suggests that the lack of stylets is a synapomorphy for *G. postclitellochaeta*, *G. ovitheca* and *G. occulta* (Fig. 4: character 2). Similarly, the spermatheca, with a pore located laterally at a distance from the intersegmental furrow 4/5 and an ectal duct which is deeply incised into an egg-shaped ampulla, is also likely to be an apomorphy for this clade (Fig. 4: character 3). *Grania vikinga*, the only Scandinavian species of the genus not included in our analysis, is similar to *G. variochaeta* in the morphology of the spermatheca and the penial apparatus (Rota and Erséus 2003), suggesting either that there is a close evolutionary relationship between the two, or that spermathecae with wide ectal ducts and oval ampullae and penial apparatuses with small glandular structures, aglandular sacs and stylets are plesiomorphic traits within the Scandinavian *Grania*-clade.

The missing ventral lobes of the pharyngeal glands in IV is a feature shared by *G. ovitheca* and *G. occulta* but not by *G. postclitellochaeta*. Therefore, the ventral lobes appear either to have re-appeared in the latter or been lost independently in *G. ovitheca* (Fig. 4: character 4a) and G. occulta (character 4b). Recent studies on other annelids have found independent character loss to be common (Struck et al. 2007; Bleidron et al. 2009), so we judge the latter scenario to be more likely.

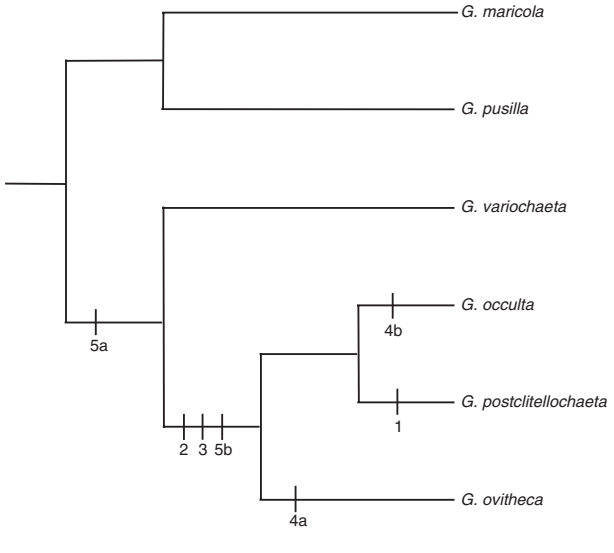

Fig. 4. Hypothesized character evolution in Scandinavian spp. of *Grania*. 1. Gain of large glandular structure surrounding penial bulb. 2. Loss of penial stylet. 3. Evolution of characteristic spermatheca (see text). 4a and b. Reduction of ventral pharyngeal glands in IV. 5a. Reduction of anterior chaetae to irregular pre-clitellar ventral chaetation and lateral chaetae of posteriormost 10 segments. 5b. Complete reduction of pre-clitellar chaetae and lateral chaetae.

Within *Grania*, the most common chaetal distribution is with ventral chaetae in every segment from IV to VI posteriorly, and lateral chaetae in every segment from XII to XX (posterior to the clitellum). In our analysis, *G. ovitheca*, *G. postclitellochaeta* and *G. occulta* deviate from this condition in that the ventral chaetae are further reduced anteriorly, and the lateral ones are completely reduced, another feature supporting their monophyly. *Grania variochaeta*, which has an intermediate chaetal distribution with irregular occurrences of pre-clitellar chaetae and lateral chaetae in only the posteriormost segments, is the sister taxon to this clade. Thus, the phylogeny seems to indicate that within this North-East Atlantic group, evolution has progressed via a reduction in the anterior ventral and lateral chaetae to a total loss of lateral chaetae (Fig. 4: characters 5a, b). This is also supported by the fact that the sister clade (*G. maricola* + *G. pusilla*) has a 'normal' distribution, i.e. ventral chaetae from VI (or sometimes V in *G. pusilla*) and lateral ones from XX to XXV. *Grania vikinga* possesses ventral chaetae from IV and laterals from XIX to XX (Rota and Erséus 2003), somewhat supporting the alternative that this species is not the closest relative of *G. variochaeta*, but rather is derived from a more basal position in the *Grania* tree.

It is worth noticing that species of *Grania* in other parts of the world also have reduced anterior chaetae. For example, *G. eurystila* Coates and Stacey, 1997; so far only known from Darwin harbour, Northern Australia, has the same chaetal distribution as *G. postclitellochaeta* (Coates and Stacey 1997), and, similarly, the North-West Atlantic *G. monospermatheca* Erséus and Lasserre, 1976 (used here as outgroup), lacks all pre-clitellar ventral chaetae, as well as lateral chaetae altogether (Erséus and Lasserre, 1976). With respect to both of these species, the species found in Scandinavia are a monophyletic group (De Wit et al., in prep), so it is clearly possible that this character has evolved convergently in several different lineages of *Grania*.

COI barcoding

The analysis of this study shows that there are clear barcoding gaps present between the Scandinavian species of *Grania*. Using COI barcodes, we would be able to identify all individuals in this study to the correct species using a 99% PP cut-off (although the finding of *G. occulta* suggests that more cryptic species may be present). Previous studies on clitellates have shown similar results. Chang et al. (2009) concluded that almost all species of Chinese and Taiwanese earthworms could be identified through barcodes. Similarly, Erséus and Kvist (2007) found that species of *Tubificoides* (Naididae) were suitable for identification using DNA barcodes, at least when applied to material from a limited geographical area. In this study, however, it is troublesome to see that occasional individuals show divergences from conspecific individuals of 6–8%. In such cases, additional data from nuclear markers are essential to verify species status, a fact which warns us about using only mtDNA for identification. It is probable that with more extensive geographical sampling, the intraspecific variation within COI could increase further as individuals from more populations are sampled. In the future,

it would seem that ITS is a good additional marker to use to test species hypotheses when mitochondrial data indicates historical evolutionary separation events between populations.

In any case, the deposition of these sequences into the Barcode of Life Database will facilitate future studies of these taxa. It remains no less important, however, to critically examine the morphology of new specimens in the future, to further test the taxonomy suggested by DNA data (both mitochondrial and nuclear). The best taxonomy is without doubt the product of a holistic approach where morphological, ecological and genetic evidence is used together to delimitate species.

Acknowledgements

We thank the staffs at the Sven Lovén Center for Marine Research (Kristineberg and Tjärnö) in Sweden, and at Espegrend Marine Station in Norway, for all help with specimen collection. We also thank the Swedish Taxonomy Initiative for allowing us to join collecting cruises along the west coast of Sweden. The first author further thanks Christopher Jensen for assistance with field work in France. We are also indebted to the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS), the Swedish Taxonomy Initiative, Wilhelm and Martina Lundgren's Science Fund, and Helge Ax:son Johnson's Foundation, for financial support.

Zusammenfassung

Genetische Variation und Phylogenie der skandinavischen Grania-*Arten (Annelida: Clitellata: Enchytraeidae) und Entdeckung einer kryptischen Art*

Wir beschreiben hier die intraspezifische Variation am COI Locus (mitochondriale DNA) und der ITS Region (nukleäre DNA) von Individuen von 5 Arten von *Grania* (Annelida: Clitellata: Enchytraeidae) aus Schweden, Norwegen und Frankreich. Unsere Studie zeigt eine geringe Variation innerhalb der früher beschriebenen Morphospezies, verglichen mit einer höheren Variation an beiden Loci zwischen den Arten. Eine einzige Ausnahme war ein Individuum von *Grania*, welches auf morphologischer Basis nicht von *G. ovitheca* unterschieden werden konnte, jedoch an beiden Loci, COI und ITS, abweicht von *G. ovitheca* und stattdessen nahe Verwandtschaft zu *G. postclitellochaeta* zeigt. Mittels dieses Individuums beschreiben wir hier eine neue Art, *G. occulta* sp.n. Zusätzliche Analysen zeigen eine nahe Verwandtschaft zwischen *G. variochaeta*, *G. occulta*, *G. ovitheca* und *G. postclitellochaeta* sowie auch zwischen *G. pusilla* und *G. maricola*. Auf der Basis der phylogenetischen Analysen diskutieren wir die Evolution morphologischer Merkmale innerhalb der skandinavischen *Grania*-Arten.

References

Beauchamp KA, Gay M, Kelley GO, El-Matbouli M, Kathman RD, Nehring RB, Hedrick RP 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.

Bleidorn C, Hill N, Erséus C, Tiedemann R (2009) On the role of character loss in orbiniid phylogeny (Annelida): Molecules vs. morphology. Mol Phyl Evol **52:**57–69.

Chang C-H, Rougerie R, Chen J-H (2009) Identifying earthworms through DNA barcodes: pitfalls and promises. Pedobiologia **52:**171–180.

Coates KA (1984) Specific criteria in *Grania* (Oligochaeta, Enchytraeidae). Hydrobiologia **115:**45–50.

Coates KA, Stacey D (1997) Enchytraeids (Oligochaeta: Annelida) of the lower shore and shallow subtidal of Darwin Harbour, Northern Territory, Australia. In: Hanley J, Caswell G, Megirian D, Larson H (eds), Proceedings of the Sixth International Marine Biological Workshop: The Marine Flora and Fauna of Darwin Harbour, Northern Territory, Australia. Museums and Art Galleries of the Northern Territory and the Australian Marine Sciences Association, Darwin, Australia, pp 67–79.

De Queiroz K (2007) Species concepts and species delimitation. Syst Biol **56:**879–886.

Duran S, Pascual M, Turon X (2004) Low levels of genetic variation in mtDNA sequences over the western Mediterranean and Atlantic range of the sponge *Crambe crambe* (Poecilosclerida). Mar Biol **144:**31–35.

Erséus C (1974) *Grania pusilla* sp.n. (Oligochaeta, Enchytraeidae) from the west coasts of Norway and Sweden with some taxonomic notes on the genus Grania. Sarsia **56:**87–94.

Erséus C (1977) Marine Oligochaeta from the Koster area, west coast of Sweden, with descriptions of two new enchytraeid species. Zool Scr **6:**293–298.

Erséus C, Gustafsson DR (2009) Chapter 3. Cryptic speciation in clitellate model organisms. in: Shain DH (ed), Annelids in Modern Biology. Wiley-Blackwell, Hoboken, New Jersey, pp 31–46.

Erséus C, Kvist S (2007) COI variation in Scandinavian marine species of *Tubificoides* (Annelida: Clitellata: Tubificidae). J Mar Biol Assoc UK **87:**1121–1126.

Erséus C, Lasserre P (1976) Taxonomic status and geographic variation of the marine enchytraeid genus *Grania* Southern (Oligochaeta). Zool Scr **5:**121–132.

Folmer O, Black M, Hoen 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.

Gray JS (1997) Gradients in marine biodiversity. In: Ormond RFG, Gage JD, Angel MV (eds), Marine Biodiversity: Patterns and Processes. University Press, Cambridge, pp 18–34.

Gustafsson DR, Price DA, Erséus C (2009) Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. Mol Phylogenet Evol **51:**182–189.

Hallett SL, Lorz HV, Atkinson SD, Rasmussen C, Xue L, Bartholomew JL (2009) Propagation of the myxozoan parasite *Myxobolus cerebralis* by different geographic and genetic populations of *Tubifex tubifex*: an Oregon perspective. J Invertebr Pathol **102:**57–68.

Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol **54:**852–859.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003a) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci **270:**313–321.

Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B Biol Sci **270 (Suppl.):** 96–99.

Hennig W (1966) Phylogenetic Systematics. University of Illinois Press, Urbana.

Hickerson MJ, Meyer CP, Moritz C (2006) DNA barcoding will often fail to discover new animal species over broad parameter space. Syst Biol **55:**729–739.

Huang J, Xu Q, Jun Sun Z, Lan Tang G, You Su Z (2007) Identifying earthworms through DNA barcodes. Pedobiologia **51:**301–309.

Källersjö M, von Proschwitz T, Lundberg S, Eldenas 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. Zool Scr **34:**415–424.

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform **9:**286–298.

King RA, Tibble AL, Symondson WOC (2008) Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. Mol Ecol **17:**4684–4698.

Knöllner FH (1935) Ökologische und systematische Untersuchungen über litorale und marine Oligochäten der Kieler Bucht. Zoologische Jahrbücher (Systematik) **66:**425–512.

Lasserre P (1967) Oligochétes marins des côtes de France. II. Roscoff, Penpoull, Étangs saumâtres de Concarneau: Systématique, Écologie. Cah Biol Mar **8:**273–293.

Locke JM, Coates KA (1999) Redescriptions of *Grania americana, G. bermudensis* and descriptions of two new species of *Grania* (Annelida: Clitellata: Enchytraeidae) from Bermuda. Proc Biol Soc Wash **112:**598–623.

Maggs CA, Castilho R, Foltz D et al. (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology **89:**S108–S122.

Mayr E (1942) Systematics and the Origin of Species. Columbia University Press, New York.

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol **3:**2229–2238.

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. PLoS Biol **2:**1529–1531.

Nygren A, Eklöf J, Pleijel F (2009) Arctic-boreal sibling species of *Paranaitis* (Polychaeta, Phyllodocidae). Mar Biol Res **5:**315–327.

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics **24:**581–583.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572–1574.

Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. Bull Am Mus Nat Hist **162:**267–376.

Rota E, Erséus C (2003) New records of *Grania* (Clitellata, Enchytraeidae) in the Northeast Atlantic (from Tromso to the Canary Islands), with descriptions of seven new species. Sarsia **88:**210–243.

Rozen S, Skaletzky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds), Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Tottowa, NJ, pp 365–386.

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopaedia of life: an introduction to DNA barcoding. Philos Trans R Soc Lond B Biol Sci **360:**1805–1811.

Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence, not intraspecific variation. Mol Ecol Resour **8:**247–255.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol **16:**1114–1116.

Siddall ME, Trontelj P, Utevsky SY, Nkamany M, Macdonald KS III (2007) Diverse molecular data demonstrate that commercially available medicinal leeches are not *Hirudo medicinalis*. Proc Biol Sci **274:**1481–1487.

Simpson GG (1951) The species concept. Evolution **5:**285–298.

Smith MA, Fisher BL, Hebert PDN (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos Trans R Soc Lond B Biol Sci **360:**1825–1834.

Southern R (1913) Part 48: Oligochaeta. Clare Island Survey. Proceedings of the Royal Irish Academy (B), pp 1–14.

Stamatakis A (2006) RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinform **22:**2688–2690.

Struck THU, Schult N, Kusen T, Hickman E, Bleidorn C, McHugh D, Halanych KM (2007) Annelid phylogeny and the status of Sipuncula and Echiura. BMC Evol Biol **7:**57, doi: 10.1186/1471-2148-7-57.

Sturmbauer C, Opadiya GB, Niederstätter H, Riedmann A, Dallinger R (1999) Mitochondrial DNA reveals cryptic oligochaete species differing in cadmium resistance. Mol Biol Evol **16:**967–974.

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

Van Valen L (1976) Ecological species, multispecies, and oaks. Taxon **25:**233–239.

Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philos Trans R Soc Lond B Biol Sci **360:**1859–1868.

Vogler AP, Monaghan MT (2007) Recent advances in DNA taxonomy. J Zoolog Syst Evol Res **45:**1–10.

White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds), PCR Protocols: A Guide to Methods and Applications. Academic Press, Inc., New York, NY, pp 315–322.

Wiklund H, Glover AG, Johannessen PJ, Dahlgren TG (2009) Cryptic speciation at organic-rich marine habitats: a new bacteriovore annelid from whale-fall and fish farms in the North-East Atlantic. Zool J Linn Soc **155:**774–785.

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference (available at http://ceb.csit.fsu.edu/awty).

Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. Syst Biol **54:** 844–851.

Wörheide G (2006) Low variation in partial cytochrome oxidase subunit I (COI) mitochondrial sequences in the coralline demosponge *Astrosclera willeyana* across the Indo-Pacific. Mar Biol **148:**907–912.

Wright S (1940) The statistical consequences of Mendelian heredity in relation to speciation. In: Huxley J (ed), The New Systematics. University Press, London, pp 161–183.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Station list, collecting sites of studied specimens.

Table S2. Specimens, loci used for analyses with GenBank accession numbers. For location details, see Table S1.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.