MOLECULAR PHYLOGENY OF NORTH SEA SEPIOLINAE (CEPHALOPODA: SEPIOLIDAE) REVEALS AN OVERLOOKED *SEPIOLA* SPECIES

DICK S.J. GROENENBERG[1], JEROEN GOUD[1], ATE DE HEIJ[1] AND EDMUND GITTENBERGER[1,2]

[1]*National Museum of Natural History Naturalis, PO Box 9517, NL 2300 RA Leiden, The Netherlands; and*
[2]*Institute of Biology, Leiden University, PO Box 9516, NL 2300 RA Leiden, The Netherlands*

(Received 16 December 2008; accepted 21 April 2009)

ABSTRACT

Recent bottom trawl surveys in the North Sea have broadened our understanding of an obscure group of bobtail squids (Sepiolidae: Sepiolinae). There are at least two *Sepietta* species in this region, viz. *Sepietta oweniana* and *Sepietta neglecta*, and three *Sepiola* species, one of which is undescribed, viz. *Sepiola atlantica*, *Sepiola pfefferi* and *Sepiola* sp. nov. *Sepiola pfefferi* is distinct from *Sepiola aurantiaca* because of differences in the hectocotylus; the type localities of both species are far apart and data in GenBank indicate a substantial genetic difference between them. It is unclear how to distinguish *Sepiola* sp. nov. morphologically from its sister species *S. atlantica*, but molecular phylogenetic analyses and distributional data readily set them apart. The occurrence of *Sepiola rondeleti* in the northeastern Atlantic could not be confirmed. Several obviously incorrect or dubious identifications regarding GenBank data are listed.

INTRODUCTION

The Sepiolinae are small, poorly known, so-called bobtail squids, with a mantle length of up to 5 cm for the largest species used in this study. The subfamily encompasses 5 of the 15 known genera of Sepiolidae, viz. *Euprymna*, *Inioteuthis*, *Rondeletiola*, *Sepietta* and *Sepiola*. Of these, only *Sepiola* and *Sepietta* are known from the North Sea, where they are represented by seven species (Reid & Jereb, 2005). These are *Sepiola atlantica* d'Orbigny, 1839–1842 (it is unknown in what year the plate with legends which validate this name were published; see Tillier & Boucher-Rodoni, 1993), *Sepiola intermedia* Naef, 1912, *Sepiola rondeleti* Leach, 1817, *Sepiola aurantiaca* Jatta, 1896, *Sepiola pfefferi* Grimpe, 1921, *Sepietta neglecta* Naef, 1916 and *Sepietta oweniana* (d'Orbigny, 1839–1841) (the plate with legends, which could validate this name, was published in 1839–1842, whereas the text was published in 1841).

Although Reid & Jereb (2005) include the northeastern Atlantic in the geographical distribution of *S. intermedia*, there is a query about its distribution west of the Strait of Gibraltar (Reid & Jereb, 2005: 163, fig. 236). The occurrence of *S. rondeleti* in the North Sea is doubtful as well, since the recent records (Reid & Jereb, 2005: 168, fig. 244) all refer directly or indirectly to relatively old data in the literature, i.e. Joubin (1895, 1902) and Naef (1912). Apart from that, Reid & Jereb (2005) did not cite a critical paper by Grimpe (1925), in which he attributed the records by Joubin and Naef to *S. oweniana*. Later, Grieg (1933) came to the same conclusion for records of so-called *S. rondeleti* from the Scandinavian part of the North Sea. *Sepiola rondeleti* is also missing from comprehensive overviews of the German, Dutch and British areas of the North Sea (Jaeckel, 1958; Janssen, 1975; Hayward & Ryland, 1990).

Sepiola aurantiaca was described from the Tyrrhenian Sea by Jatta (1896). At the beginning of the twentieth century, Russell collected twenty specimens of sepiolids from the North Sea, off the northeast coast of Scotland (*c.* 56°N, 3°W). Despite some details of the hectocotylus, i.e. the lack of the peculiar lobe at the base of the first arm right (or right dorsal arm) (Fig. 1; illustration in Russell, 1922), Russell (1909) attributed these bobtail squids to *S. aurantiaca*. Grimpe (1921) introduced *S. pfefferi* as a new species for sepiolids from the east coast of England (*c.* 53°N, 1°E). *Sepiola pfefferi* could allegedly be discerned from *S. aurantiaca* by a deep V-shaped (instead of U-shaped) mantle outline (Fig. 2) and a basal lobe only on the left dorsal arm of the hectocotylus. According to Naef (1923) however, the differences between *S. pfefferi* and *S. aurantiaca* are so minimal that he considered *S. pfefferi* to be a form of *S. aurantiaca*.

Clearly, the Sepiolinae represent a controversial group in which species are often difficult to identify morphologically (especially juvenile and female specimens). Sibling species are perhaps most easily distinguished genetically. In such cases a database like GenBank will be most valuable. However, the utility of such databases depends on how well the taxonomic group under study is represented and whether reliably identified voucher specimens are available for more detailed morphological studies. While trying to characterize the North Sea Sepiolinae by DNA analyses ('barcoding') and to unravel their phylogenetic relationships, our initial views, largely based on data in the literature, had to be modified.

MATERIAL AND METHODS

Sampling

Sepiolinae in general are small, obscure, bottom-dwelling organisms, that are most commonly known from the by-catch of fisheries. Except for SCUBA diving, which is not a feasible means to survey larger areas, there does not seem to be an easy way specifically to catch these small cuttlefish. We therefore made use of trawl surveys in this study. For a number of years, IBTS (International Bottom Trawl Survey) and BTS (Beam Trawl Survey) surveys have been conducted in the North Sea

Correspondence: D.S.J. Groenenberg;
e-mail: groenenberg@naturalis.nl

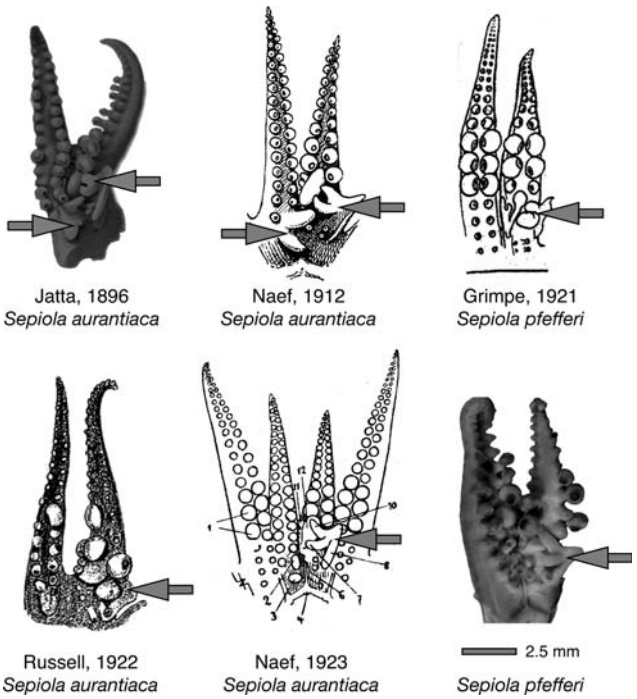

Figure 1. Ventral view of dorsal arms of *Sepiola aurantiaca* and *Sepiola pfefferi*. Arrows indicate basal lobes on the hectocotylus. The scale refers only to the photo in the lower right corner of a hectocotylus of *S. pfefferi*, collected in 1931 from Cove Bay, off Aberdeen (57°06′N 2°03′W).

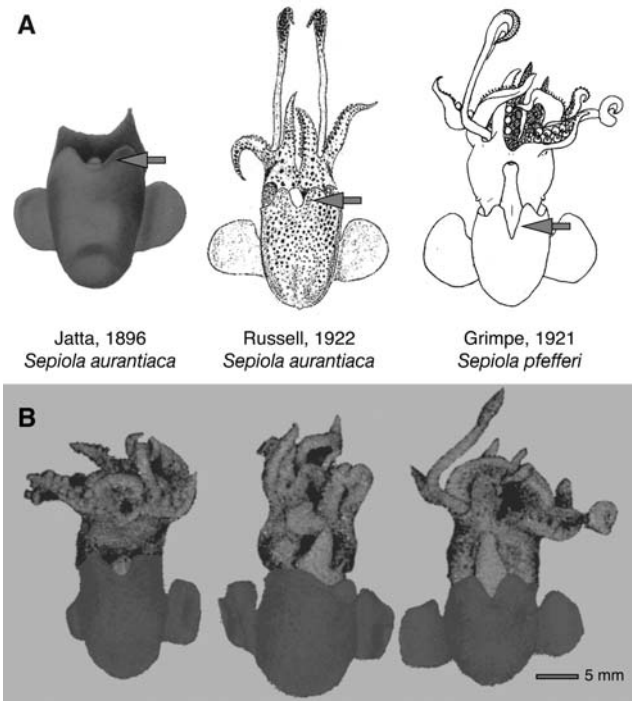

Figure 2. A. U- and V-shaped mantle outline (arrows) of *Sepiola aurantiaca* and *Sepiola pfefferi*. **B.** Illustration after Grimpe (1921) showing the type specimens of *S. pfefferi* with their mantles highlighted. The scale only refers to the photographed part of this figure.

by IMARES (the Netherlands Institute for Fisheries Research) under auspices of ICES (the International Council for the Exploration of the Sea). The sepiolids from these surveys were stored in 96% ethanol and added to the collection of the National Museum of Natural History *Naturalis* (Leiden, The Netherlands). Our identifications were primarily based on morphological data published by Grimpe (1925) and Reid & Jereb (2005), i.e. size, characters of the hectocotylus, the number and arrangement of suckers, and the presence and position of light organs. All species that were indisputably known to occur in the North Sea were represented in our material: *Sepiola atlantica* (21 specimens), *Sepiola aurantiaca*/*Sepiola pfefferi* (5), *Sepietta oweniana* (4) and *Sepietta neglecta* (8). Hence we assume that the IMARES surveys yielded a representative sample of this subfamily for this area. Specimens and associated data (depth, gender) are listed in Table 1 and sampling sites are illustrated in Figure 3.

Type specimens and vouchers

The syntypes of neither *S. aurantiaca* nor *S. pfefferi* could be located in the institutes where they reportedly had been deposited, i.e. the Stazione Zoologica di Napoli, Naples (A. Travaglini, personal communication) and the Museum für Naturkunde der Humboldt-Universität, Berlin (M. Glaubrecht, personal communication), respectively. We are not aware of any recently collected material of undisputable *S. aurantiaca* and a voucher specimen for the sequence in GenBank (AF035708) was not available (M.E. Nishiguchi, personal communication).

DNA isolation, PCR and sequencing

Total genomic DNA was extracted from up to 15 mg of muscle tissue, using a DNeasy® Tissue Kit (Qiagen), following the manufacturer's protocol. Specimen vouchers were deposited at the National Museum of Natural History Naturalis, Leiden, The Netherlands. A 658-bp fragment of cytochrome oxidase subunit I (COI), that was shown to be informative in other cephalopod studies (Nishiguchi, Ruby & McFall-Ngai, 1998; Lindgren, Giribet & Nishiguchi, 2004; Nishiguchi, Lopez & Boletzky, 2004), was amplified with primers L1490 5′-GGTCA ACAAATCATAAAGATATTGG-3′ and H2198 5′-TAAACTT CAGGGTGACCAAAAAATCA-3′ (Folmer *et al.*, 1994). PCR conditions were 1.5 mM $MgCl_2$, 0.2 mM DNTPs, 0.4 μM of each primer and 5 U of *Taq* DNA polymerase (Qiagen) per reaction. Reactions were always carried out in a volume of 25 μl using a thermocycle profile of 3 min at 94°C, followed by 40 cycles of 15 s at 94°C, 30 s at 50°C and 40 s at 72°C, and a final extension of 5 min at 72°C. PCR products were purified with Nucleospin® Extract II columns (Macherey-Nagel) and both strands were directly sequenced using the same primers. Sequencing was done on an ABI 3730 automated sequencer (Applied Biosystems) at Macrogen Corp. (Korea). Forward and reverse sequences were assembled and checked using Sequencher version 4.2 (Gene Codes Corp.). Sequences were aligned manually using MacClade 4.08 (Maddison & Maddison, 2003).

Phenetic analysis

A neighbour-joining phylogram (optimality criterion = distance, total character difference, distance measure = uncorrected *p*) was constructed with PAUP 4.0b10 (Swofford, 2003), including sequences from this study, all Sepiolinae sequences that had been previously deposited in GenBank and four outgroup sequences (representing Rossinae, Heteroteuthinae and Sepiadariidae), also from GenBank. Genetic distances (uncorrected *p*-distances; absolute number of differences, Table 2) were calculated with PAUP 4.0b10 as well. Although corrected distances [such as General Time Reversible (GTR) model, see

Table 1. Specimens, sampling information and GenBank accession numbers (GB ACCN).

Species	Lat (°N)	Long (°E)	Depth (m)	Sex	Date	GB ACCN	RMNH_MOL
Sepietta neglecta Naef, 1916	56.33653	4.5567	62	F	2.ii.2005	FJ231301	105647
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	F	8.ii.2008.	FJ231312	110308
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	F	8.ii.2008.	FJ231324	110308.1
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	F	8.ii.2008.	FJ231325	110308.2
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	F	8.ii.2008.	FJ231326	110308.3
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	M	8.ii.2008.	FJ231327	110309.1
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	M	8.ii.2008.	FJ231328	110309.2
Sepietta neglecta Naef, 1916	56.66683	−1.61233	58	M	8.ii.2008.	FJ231329	110309.3
Sepietta oweniana (d'Orbigny, 1839–1841)	55.73613	−0.27547	73	M	9.ix.2005	FJ231297	107426
Sepietta oweniana (d'Orbigny, 1839–1841)	57.82507	−2.42627	89	M	26.viii.2005	FJ231298	108685
Sepietta oweniana (d'Orbigny, 1839–1841)	55.68268	6.40243	43	-	23.viii.2007	FJ231299	110057
Sepietta oweniana (d'Orbigny, 1839–1841)	54.64303	0.52227	70	F	12.ix.2005	FJ231300	108941
Sepiola atlantica d'Orbigny, 1839–1842	53.4611	0.9181	20	F	15.ii.2006	FJ231302	102844
Sepiola atlantica d'Orbigny, 1839–1842	53.106798	1.7576	38	M	26.i.2005	FJ231303	105641
Sepiola atlantica d'Orbigny, 1839–1842	53.106798	1.7576	20	M	26.i.2005	FJ231304	105658
Sepiola atlantica d'Orbigny, 1839–1842	51.57346	2.77096	32	M	25.i.2005	FJ231308	105638
Sepiola atlantica d'Orbigny, 1839–1842	53.62875	2.46205	25	M	11.ix.2007	FJ231309	110059
Sepiola atlantica d'Orbigny, 1839–1842	52.2099	2.99283	38	F	23.8.2004	FJ231310	99500
Sepiola atlantica d'Orbigny, 1839–1842	54.67276	2.57207	20	F	9.ii.2005	FJ231311	105655
Sepiola atlantica d'Orbigny, 1839–1842	54.70469	0.37152	68	M	z2006	FJ231314	105637
Sepiola atlantica d'Orbigny, 1839–1842	52.68033	4.04217	24	F	28.i.2008	FJ231316	110291
Sepiola atlantica d'Orbigny, 1839–1842	54.28333	8.03333	19	M	29.i.2008	FJ231317	110293
Sepiola atlantica d'Orbigny, 1839–1842	54.12033	4.90917	44	M	30.i.2008	FJ231318	110297
Sepiola atlantica d'Orbigny, 1839–1842	53.8275	4.21467	29	F	13.ii.2008	FJ231322	110316
Sepiola sp. nov. (Goud & de Heij, in prep.)	56.78483	−0.23631	79	F	7.ii.2005	FJ231305	105650
Sepiola sp. nov. (Goud & de Heij, in prep.)	57.82509	−2.42626	63	M	26.viii.2005	FJ231306	108684
Sepiola sp. nov. (Goud & de Heij, in prep.)	55.68268	6.40243	43	M	21.viii.2007	FJ231307	110058
Sepiola sp. nov. (Goud & de Heij, in prep.)	54.77194	−0.25279	66	M	z2006	FJ231313	105636
Sepiola sp. nov. (Goud & de Heij, in prep.)	55.21893	−0.50832	75	F	z2005	FJ231315	108938
Sepiola sp. nov. (Goud & de Heij, in prep.)	55.83367	3.35133	64	M	5.ii.2008	FJ231319	110298
Sepiola sp. nov. (Goud & de Heij, in prep.)	56.27517	4.45883	65	F	6.ii.2008	FJ231320	110299
Sepiola sp. nov. (Goud & de Heij, in prep.)	56.09333	1.19967	83	M	11.ii.2008	FJ231321	110311
Sepiola sp. nov. (Goud & de Heij, in prep.)	54.30417	−1.27367	94	M	7.ii.2008	FJ231323	110301
Sepiola pfefferi Grimpe, 1921	57.74243	−1.3547	90	M	26.viii.2005	FJ231292	108846.1
Sepiola pfefferi Grimpe, 1921	57.74243	−1.3547	90	-	26.viii.2005	FJ231293	108846.2
Sepiola pfefferi Grimpe, 1921	53.44067	0.96167	20	F	26.i.2005	FJ231294	105642
Sepiola pfefferi Grimpe, 1921	53.46110	0.9181	38	F	15.ii.2006	FJ231295	110056
Sepiola pfefferi Grimpe, 1921	53.11671	1.7827	38	M	26.i.2005	FJ231296	108845

RMNH_MOL refers to the voucher reference number (Molluscan collection) of the National Museum of Natural History, Naturalis (Leiden, The Netherlands). Sex determination: M, male; F, female; -, juvenile (or undetermined).

the paragraph below] will be more realistic (e.g. they compensate for multiple substitutions at the same position), we choose to use uncorrected distances, because these at least show the minimum amount of change.

Phylogenetic analysis

To reduce computational time and focus on the genera present in the area of study, we retained one representative of the sequences from GenBank for *Euprymna tasmanica* (AY293713), *Euprymna scolopes* (AF035701) and *Euprymna hyllebergi* (AY293714) for Bayesian analysis. The sequence in GenBank for *Rondeletiola minor* (AF035714) was also excluded, as we suspect that it represents a specimen of *E. scolopes* (see Discussion). For the Bayesian analysis, the best-fit model of sequence evolution was selected with MrModeltest version 2.2 (Nylander, 2004). Both hLRT and AIC selected a GTR model (Nst = 6) with gamma distribution rate variation across sites and a proportion of invariable sites (rates = invgamma; GTR + I + G). A flat Dirichlet distribution was selected as prior on state frequencies (statefreqpr = dirichlet (1,1,1,1)). Bayesian inference was done with MrBayes 3.1.2. (Ronquist & Huelsenbeck, 2003), using two independent runs of four Markov chains with 10,000,000 generations. Sampling was done every 100 generations and the first 5,000,000 generations were discarded (burnin = 50,000; standard deviation of split frequencies was 0.0176 and dropped below 0.01 after 8,775,000 generations). *Sepioloidea lineolata* (AF000064) was specified as outgroup.

RESULTS

Both phenetic and cladistic analyses (Figs 4, 5, respectively) distinguished *Sepietta neglecta* (group F) and *Sepietta oweniana* (group G). No intraspecific variation was observed in either species according to our own data, whereas the interspecific divergence between these species was 7.6% (Table 2).

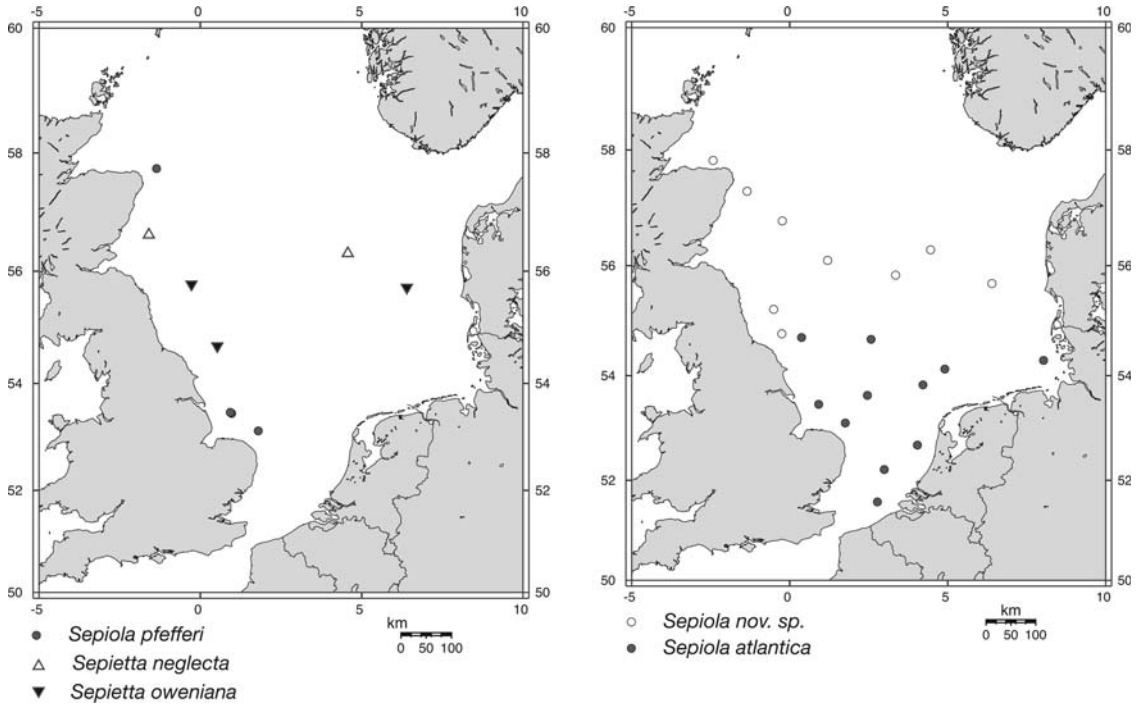

Figure 3. Sampling localities of *Sepiola* and *Sepietta* in the North Sea.

Sequences of specimens provisionally named *Sepiola pfefferi* (group E) differed substantially (14.5% sequence divergence) from the only available alleged *Sepiola aurantiaca* sequence (AF035708, GenBank), while intraspecific variation within *S. pfefferi* was low (0.15% or less, Table 2). Specimens initially identified as *Sepiola atlantica*, turned out to represent two monophyletic groups: *S. atlantica* and *Sepiola* sp. nov. (groups A and B respectively; Figs 4, 5) with a minimum sequence divergence of 3.50% (Table 2). There was no intraspecific sequence divergence within *Sepiola* sp. nov. and our data show low intraspecific sequence divergence (0.15%) within *S. atlantica* (Table 2). Numerous Sepiolinae species were represented by more than one COI sequence in GenBank, but rarely did allegedly conspecific individuals form monophyletic groups (Fig. 4, Table 3).

DISCUSSION

Five species of Sepiolidae occur in the North Sea, representing two genera, viz. *Sepietta oweniana* (group G) and *Sepietta neglecta* (group F), and *Sepiola atlantica* (group B), *Sepiola pfefferi* (group E) and *Sepiola* sp. nov. (group A), with different distributional patterns (Fig. 3). For the moment, *S. atlantica* and *Sepiola* sp. nov. are considered sibling species. Based on DNA analyses of COI, a provisional phylogeny reconstruction (Fig. 5) is given, including sepiolid taxa from elsewhere using data in GenBank. Our data lead us to conclude that *Sepiola* is paraphyletic, whereas *Sepietta* is monophyletic. This study indicates that a high percentage of the species, genera and, in one case, subfamilies of Sepiolidae in GenBank are most probably misidentified. This hampered our conclusions, in particular since voucher specimens were not available for inspection.

There is a striking contrast between sequences obtained with this study, that clearly show distinct species that are well separated from each other (small intraspecific sequence divergences and large interspecific divergences) and sequences deposited in GenBank, which showed two other extremes: either sequences of the same so-called species that showed large interspecific divergences (Table 3) or sequences with limited intraspecific divergences of which the vouchers were identified as different species (group D, Figs 4, 5).

Both this incongruence between our data and the data from GenBank, as well as the high number of likely misidentifications (Fig. 4, Table 3), limited our ability to evaluate whether the sequence of *Sepiola aurantiaca* is actually distinct from group D (Fig. 4, 5). According to Nishiguchi *et al*. (1998, 2004) group D represents seven species. Among these all, except *S. atlantica* (AF035707), were collected in the Mediterranean Sea; most came from the same locality (Banyuls-sur-Mer; Nishiguchi *et al*., 1998, 2004). Comparing the topology of Figures 4, 5, and divergence among sequences (Table 2) referred to as *S. atlantica* d'Orbigny, 1839–1842, *Sepiola intermedia* Naef, 1912, *Sepiola affinis* Naef, 1912 and *Sepiola ligulata* Naef, 1912, led us to conclude that GenBank entries AF035707, AY293718, AY293716 and AF035710, respectively (group D), were probably misidentified and likely all pertain to the species *Sepiola robusta*.

Considering that *Heteroteuthis dispar* (Rüppell, 1844) (AF035713, group D) belongs to the subfamily Heteroteuthinae instead of Sepiolinae, and does not group with either its congener *Heteroteuthis hawaiiensis* (Berry, 1909) (AY293728) or with other representatives of the Heteroteuthinae, like *Stoloteuthis leucoptera* (Verrill, 1878) (AF000044), this specimen may also have been misidentified. One *S. robusta* Naef, 1912 sequence (AF035711, group D) is problematic in several ways: it has an insertion that causes a frame shift, coding region information is not in GenBank and it has four amino acid substitutions that are otherwise unique among sepiolid COI sequences in GenBank.

The sequence assigned to *S. aurantiaca* in GenBank (AF035708) differs considerably from other members of group D (7.85% sequence divergence between AF035708 and AY293718; Table 2). However, AF035708 is of unusual length and bears four insertions and two deletions, which are unique among sepiolid COI sequences (GenBank). AF035708 also contains stopcodons regardless of the chosen reading frame and even using an optimized alignment to minimize the number of

Table 2. Sequence divergence percentages (calculated as the uncorrected *p*-distance times one hundred) between sequences used in this study.

Group	Taxon	Group	Taxon	Seq. divergence (%)
A	*Sepiola* sp. nov. FJ231306	A	*Sepiola* sp. nov. FJ231305	0.00
B	*Sepiola atlantica* FJ231317	B	Sepiola atlantica FJ231302	0.15
B	*Sepiola atlantica* AY293721	B	*Sepiola atlantica* FJ231317	0.77
A	*Sepiola* sp. nov. FJ231305	B	*Sepiola atlantica* FJ231302	3.50
B	*Sepiola atlantica* FJ231317	A	*Sepiola* sp. nov. FJ231305	3.65
C	*Sepiola rondeleti* AY293720	C	*Sepiola affinis* AF035706	0.16
C	*Sepiola rondeleti* AY293720	C	*Sepiola intermedia* AF035709	0.99
C	*Sepiola intermedia* AF035709	C	*Sepiola affinis* AF035706	0.83
C	*Sepiola rondeleti*AY293720	B	*Sepiola atlantica* FJ231317	1.67
C	*Sepiola rondeleti*AY293720	A	*Sepiola* sp. nov. FJ231305	3.96
C	*Sepiola affinis* AF035706	B	*Sepiola atlantica* FJ231317	1.81
C	*Sepiola affinis* AF035706	A	*Sepiola* sp. nov. FJ231305	3.96
C	*Sepiola intermedia* AF035709	B	*Sepiola atlantica* FJ231317	2.65
C	*Sepiola intermedia* AF035709	A	*Sepiola* sp. nov. FJ231305	4.78
–	*Sepiola rondeleti* AF035712	C	*Sepiola rondeleti*AY293720	4.00
–	*Sepiola rondeleti* AF035712	C	*Sepiola affinis* AF035706	3.84
–	*Sepiola rondeleti* AF035712	C	*Sepiola intermedia* AF035709	3.67
–	*Sepiola rondeleti* AF035712	B	*Sepiola atlantica* FJ231318	5.50
–	*Sepiola rondeleti* AF035712	A	*Sepiola* sp. nov. FJ231305	7.50
D	*Sepiola robusta* AY293719	B	*Sepiola atlantica* AF035707	0.00
D	*Sepiola robusta* AY293719	D	*Sepiola intermedia* AY293718	0.46
D	*Sepiola robusta* AY293719	D	*Heteroteuthis dispar* AF035713	0.00
D	*Sepiola robusta* AF035711	D	*Sepiola ligulata* AF035710	2.66
D	*Sepiola ligulata* AF035710	D	*Sepiola intermedia* AY293718	1.06
D	*Sepiola ligulata* AF035710	D	*Sepiola affinis* AY293716	1.77
D	*Sepiola robusta* AY293719	D	*Sepiola aurantiaca* AF035708	7.85
E	*Sepiola pfefferi* FJ231296	E	*Sepiola pfefferi* FJ231295	0.00
E	*Sepiola pfefferi* FJ231296	E	*Sepiola pfefferi* FJ231294	0.15
F	*Sepietta neglecta* FJ231301	E	*Sepiola pfefferi* FJ231294	10.03
F	*Sepietta neglecta* FJ231329	F	*Sepietta neglecta* FJ231301	0.00
–	*Rondeletiola minor* AY293725	F	*Sepietta neglecta* FJ231301	11.57
–	*Sepiola ligulata* AY293717	E	*Sepiola pfefferi* FJ231294	9.29
D	*Sepiola aurantiaca* AF035708	E	*Sepiola pfefferi* FJ231294	14.48
G	*Sepietta oweniana* FJ231300	G	*Sepietta oweniana* FJ231297	0.00
F	*Sepietta neglecta* AY293722	G	*Sepietta oweniana* FJ231297	0.31
F	*Sepietta neglecta* AY293722	–	*Euprymna stenodactyla* AF035704	3.22
–	*Euprymna stenodactyla* AF035704	G	*Sepietta oweniana* FJ231297	2.92
G	*Sepietta obscura* AF036912	F	*Sepietta neglecta* AY293722	0.46
F	*Sepietta neglecta* FJ231329	G	*Sepietta oweniana* FJ231297	7.60
D	*Sepiola robusta* AF035711	D	*Sepiola aurantiaca* AF035708	7.13
D	*Sepiola intermedia* AY293718	D	*Sepiola aurantiaca* AF035708	7.85
–	*Sepietta obscura* AY293723	F	*Sepietta neglecta* FJ231301	9.30
–	*Sepietta obscura* AY293723	G	*Sepietta oweniana* FJ231300	9.45
B	*Euprymna tasmanica* DQ646730	B	*Sepiola atlantica* FJ1231302	1.52
B	*Euprymna tasmanica* DQ646730	–	*Euprymna tasmanica*AY293713	12.94

amino acid substitutions, it has 14 unique amino acids at positions that do not differ through sequences in the family. A BLAST search showed that this sequence is cephalopod in origin. In our opinion the large number of point mutations, indels and stopcodons could indicate a nuclear mitochondrial pseudogene (NUMT), though no bias in nucleotide composition was observed. NUMTs can also be detected by unexpected phylogenetic placement of the taxon under study (Bensasson *et al*., 2001), but the unclear phylogenetic relationships within the Sepiolinae, the availability of only a single sequence (AY035708) and the absence of a voucher (to check the morphological identification of the specimen in question) preclude use of these methods. Since syntypes of *S. aurantiaca* are missing, whereas according to Jereb, Mazzola & Di Stefano (1997) the records of the species are limited to those published by Jatta (1896) and Naef (1912), we consider *S. aurantiaca* a nomen dubium, referring to a species described from the Tyrrhenian Sea, most probably not occurring in the North Sea.

After exclusion of the dubious sequences AY035708 and AF035711, the maximum sequence divergence within group D is 1.77% (between AF035710 and AY293716, Table 2). Accepting a threshold of 3% sequence divergence for species status, as suggested by Hebert *et al*. (2003), this supports our assumption that group D represents only a single species, instead of seven according to Nishiguchi *et al*. (1998, 2004).

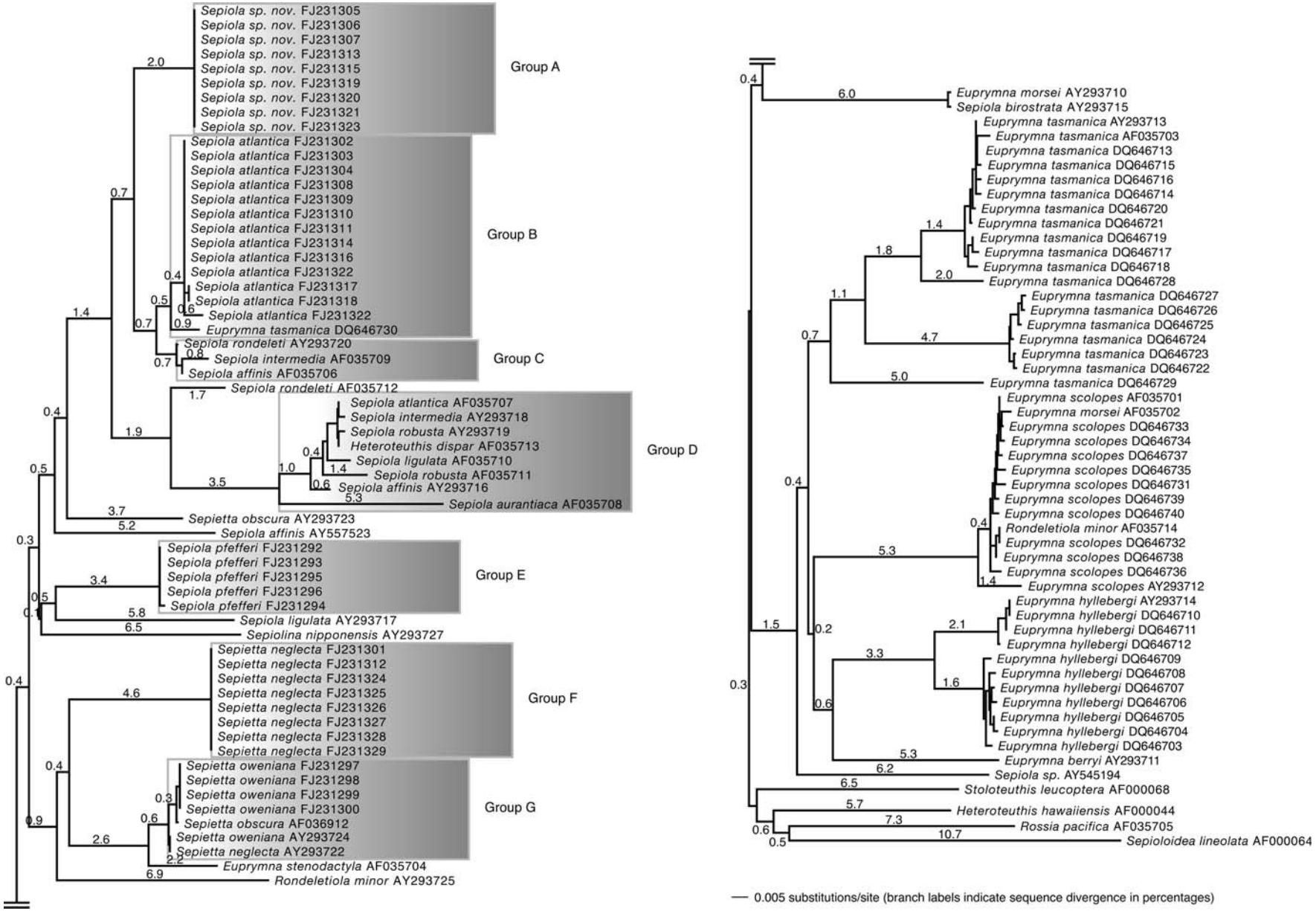

Figure 4. Neighbour-joining phylogram (branch lengths indicate sequence divergence percentages). Grey shading indicates groups of sequences of which the voucher specimens are considered conspecific in this study.

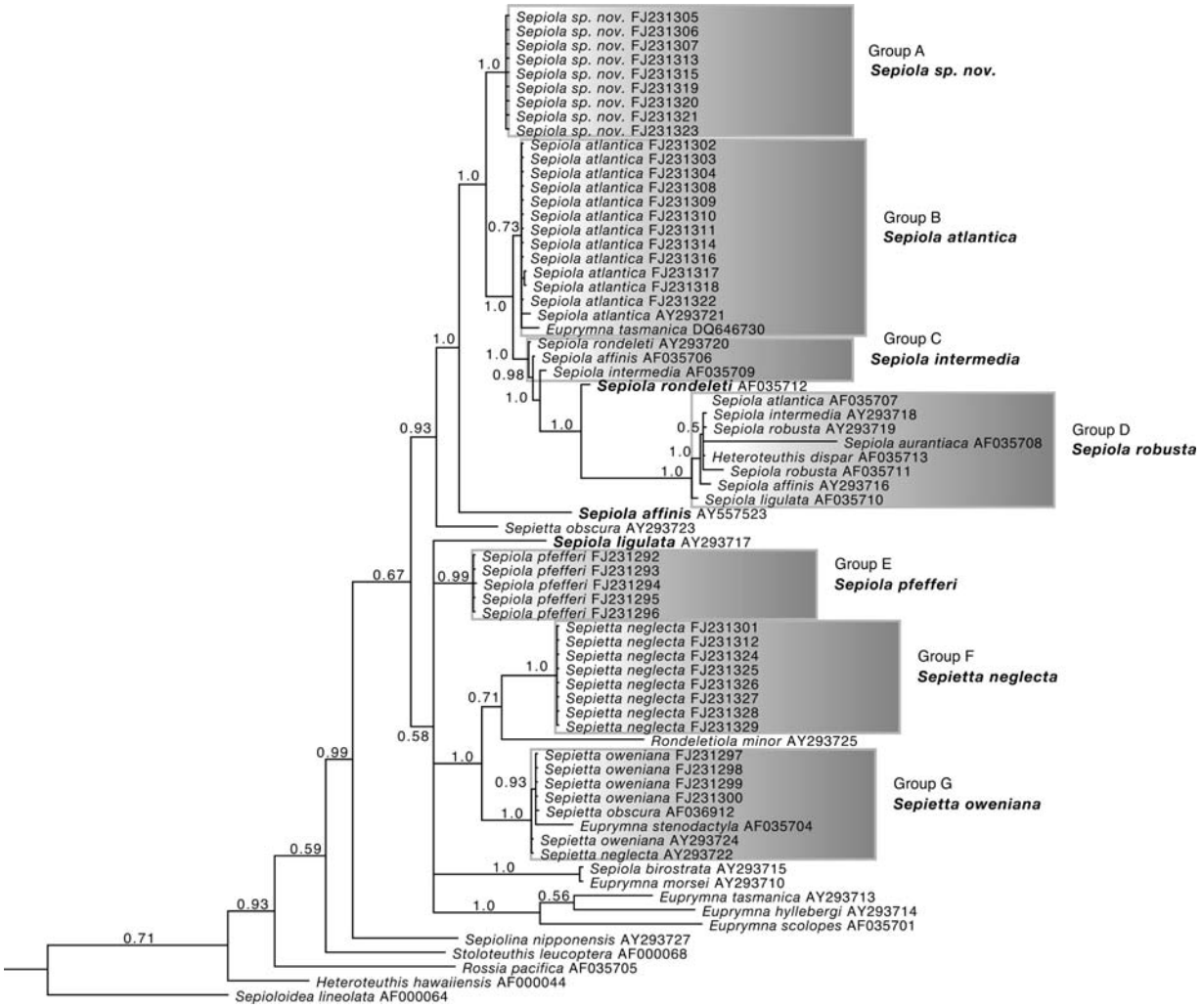

Figure 5. Bayesian phylogram with posterior probabilities. Grey shading indicates groups of sequences of which the voucher specimens are considered conspecific in this study.

Table 3. Accession numbers of conspecific specimens (according to GenBank registration) that do not form monophyletic groups.

Sepiola atlantica	AY293721 and AF035707
Sepila rondeleti	AY293720 and AF035712
Sepiola affinis	AF035706, AY557523 and AY293716
Sepiola intermedia	AF035709 and AY293718
Sepiola ligulata	AF035710 and AY293717
Sepietta obscura	AY293723 and AF036912
Euprymna tasmanica	DQ646730 and all other conspecifics (Fig. 4)
Euprymna morsei	AY293710 and AF035702
Rondeletiola minor	AY293725 and AF035714

Assuming that incorrect names were assigned to GenBank entries AF035707, AY293718, AY293716, AF035710 and AF035713, group D only encompasses *S. robusta* (AY293719), which is one of the most frequently captured Mediterranean *Sepiola* species (Jereb *et al.*, 1997; Reid & Jereb, 2005) and is suggested to be the sister species (Fig. 5) of *Sepiola rondeleti* according to Bello (1998).

In our view, Grimpe's (1921) description of *S. pfefferi* was inaccurate. The photograph of the type specimen (Fig. 2B; Grimpe, 1925) clearly showed a V-shaped mantle outline (specimen on the right), but the two syntypes on the left of it, seemed to have that structure more U- than V-shaped (as indicated for *S. aurantiaca*). Our specimens did not show the typical V-shaped mantle outline cited by Grimpe (1921). The drawing that was added (illustration on the right of Fig. 2A) is further misleading in that the size of the V-shaped incision extends up to one-third of the mantle length, whereas on the photograph the length is about one-fifth (Grimpe, 1925: specimen on the right of Fig. 2B). Even though the shape of the mantle outline is a dubious character, a clear-cut difference exists in the hectocotylus between *S. pfefferi* and *S. aurantiaca*. There are no lobes on the base of the right dorsal arm of the hectocotylus in *S. pfefferi* (Grimpe, 1921), while both Jatta (1896) and Naef (1912) depict lobes on the right dorsal arm of *S. aurantiaca* (Fig. 1). The geographical distance between the type localities of *S. aurantiaca* (Tyrrhenean Sea) and *S. pfefferi* (North Sea, East coast of England) is noteworthy as well. Despite our doubts on the correctness of the alleged *S. aurantiaca* sequence (AF035708), we have included it in our table of sequence divergences (Table 2). The *S. pfefferi* sequences (group E) differ from the sequences of all other *Sepiola* species included in this study, in particular from the alleged *S. aurantiaca* sequence. With 9.29% minimum sequence divergence (Table 2) with its sister species (Fig. 4) *S. ligulata* (AY293717), *S. pfefferi* warrants species status if the same

threshold of 3% sequence divergence (Hebert *et al*., 2003) is imposed.

Only two *Sepietta* species, *S. neglecta* and *S. oweniana*, are known from the North Sea. Reid & Jereb (2005) indicated that these species can be difficult to distinguish morphologically; both are easily separated by their COI sequences (group F and G, Figs 4, 5, Table 2), however. Therefore, we accept *S. neglecta* and *S. oweniana* as separate taxa (Bayesian posterior probability of 1.0, Fig. 5). Our data indicated no intraspecific variation in the small samples of *S. neglecta* ($n = 8$) and *S. oweniana* ($n = 4$, Table 1, Fig. 3). Addition of the GenBank *Sepietta* sequences (AY293722, AY293724 and AF036912) once more brought misidentifications to light. These sequences form a monophyletic group with our sequences of *S. oweniana* (group G), only slightly increasing the sequence divergence within that clade to 0.31% (Table 2). If we again impose a divergence threshold of 3%, group G represents a single species. This implies that the specimens belonging to GenBank sequences AF036912 (Nishiguchi *et al*., 1998) and AY293722 (Nishiguchi *et al*., 2004) identified as *Sepietta obscura* and *S. neglecta*, most likely are *S. oweniana*.

A second sequence (AY293723) is referred to as *S. obscura* in GenBank. This is the only *Sepietta* sequence that indicates that this genus is not monophyletic. Except for *Rondeletiola minor* (AY293725) and *Euprymna stenodactyla* (AF035704) AY293723, the other 15 *Sepietta* sequences used in this study, form a monophyletic group (Figs 4, 5). Consequently, we are sceptical about the identification of AY293723 as *S. obscura* by Nishiguchi *et al*. (2004); we suggest it might belong to a *Sepiola* species. More sequences of reliably identified *S. obscura* are required to clarify this point.

The phylogenetic analysis (Fig. 5) and genetic distances (Table 2) show that individuals which we initially identified as *S. atlantica* represent two groups, viz. *S. atlantica* (group B) and a cryptic sister species, *Sepiola* sp. nov. (group A). Although we cannot (yet) distinguish the latter species from *S. atlantica* on morphological characters such as size, the number and arrangement of suckers and comparison of the hectocotyli, we consider it justified to report it as a separate taxon. Even within the relatively small area of the North Sea, there seems to be a sharp geographic delineation (Fig. 3) between *S. atlantica* (group B, the southern species) and *Sepiola* sp. nov. (group A, the northern species). Further investigation has to show to what extent the sharp delineation between these groups should be attributed to either different geographical distributions of both taxa indeed (southern and northern North Sea, respectively; Fig. 3) or to a preference for different depths (on average 31 *vs* 70 m, respectively; Table 1).The separation between groups A and B is supported by high Bayesian posterior probabilities (Fig. 5). Both analyses (Figs 4, 5) show that not group A, but group C is the sister taxon of group B. Due to the inclusion of groups C and D (Fig. 5), group B is paraphyletic. According to our data ($n = 21$) no intermediates were detected between groups A and B. The minimum sequence divergence between these clades (3.5%, Table 2) exceeds the species threshold of 3% used here, while intraspecific divergences were low (0.15% and 0% within *S. atlantica* and *Sepiola* sp. nov., respectively). Due to a transversion (T to A), all individuals of *Sepiola* sp. nov. have a methionine residue at amino acid position 160, which apparently is typical for this species (all other sepiolid COI sequences in GenBank show a leucine residue at that position).

In both analyses two GenBank sequences (AY293721 and DQ646730) are placed in group B. The specimen belonging to sequence AY293721 was collected near Vigo, Spain, and (unlike the voucher for AF035707, Table 3) seems to be correctly identified as *S. atlantica*. Inclusion of AY293721 slightly increases the maximum divergence within group B to 0.77% (Table 2). According to Jones *et al*. (2006), the voucher for sequence DQ646730 was collected near Adelaide, Australia and represents *Euprymna tasmanica* (Pfeffer, 1884). This is surprising, since the other nineteen *E. tasmanica* sequences in GenBank form a monophyletic, if heterogeneous, group (Fig. 4). Possibly an invading specimen of *S. atlantica* (e.g. traveling in the ballast water of a ship from the North Sea) was caught and not recognized as such, this far outside its normal geographic distribution. Including DQ646730 in group B increases its maximum divergence to 1.52%. For the moment, we conclude that the sequence DQ646730 is actually from *S. atlantica* (Table 3).

The position of group C (Figs 4, 5) is unclear. The maximum sequence divergence between groups B and C (2.65%, Table 2) is inflated as both AF035709 and AF035706 have an insertion that causes a frame shift and coding region information is not provided to GenBank. As groups B and C cluster on the neighbour-joining (NJ) tree (Fig. 4), it is tempting to regard them as a single species, *S. atlantica*. The Bayesian analysis (Fig. 5), on the other hand, shows that groups C and D are monophyletic. According to the species identifications now registered in GenBank, there is an allegedly conspecific specimen elsewhere in the tree (Table 3) for each so-called species in group C. Among the three so-called *S. affinis* sequences in GenBank, AY557523 is basal to the groups A, B, C and D and AF035706 and AY293716 belong to groups C and D, respectively. The minimum sequence divergence between groups C and D (between AY293720 and AY293716) is 9.74% (Table 2), which (with a 3% threshold for species status) suggests that these groups are not conspecific. As shown (see the paragraph on group D), sequence AY293716 most likely belongs to *S. robusta*. Sequence AY557523 appears to be the sole representative of a separate group (sequence divergence >3%), maybe *S. affinis*, indicating that sequence AF035706 probably does not belong to this species.

Sepiola rondeleti and *S. intermedia* are each represented by two sequences, viz. AY293720 and AF035712, and AY293718 and AF035709. Of the former, AY293720 belongs to group C, whereas sequence AF035712 probably represents a separate group (minimum distance to group C, *S. rondeleti* AY293720 is 3.84% and to group D, *S. affinis* AY293716 is 7.97%, Table 2), maybe *S. rondeleti* indeed. Of the latter species, sequence AY293718 likely belongs to *S. robusta* (see the paragraph on group D). Consequently, the only sequence of group C that does not refer to a conspecific, more reliably identified specimen elsewhere in the tree (AF035709) was identified as *S. intermedia* by Nishiguchi *et al*. (1998). Hence we tentatively label group C as *S. intermedia*.

There is a discrepancy, which cannot yet be resolved, between the relationships shown in the NJ (Fig. 4) and the Bayesian (Fig. 5) figures, in particular regarding the position of group D relative to *S. atlantica*. We propose this can be attributed to a difference between phenetic and cladistic methods. The topology of a minimum evolution analysis (results not shown) is identical to Figure 4, whereas the topology of a maximum parsimony analysis (results not shown) is similar to that obtained by Bayesian inference (Fig. 5). As to the sequences identified as *R. minor* in GenBank, we can only conclude that they are not conspecific (Table 3) and that AF035714 is from *E. tasmanica* (Fig. 4). Given the consistency between the molecular phylogeny and our morphological determinations, and the observation that the intraspecific sequence divergences of these specimens are generally small (Table 2), we are convinced that other explanations for the observed discrepancies between the molecular phylogeny and the species identifications registered in GenBank, like incomplete lineage sorting or mitochondrial introgression, are all far less probable than our suggestion of misidentification.

In addition to contributing to our knowledge of the systematics and phylogeny of the Sepiolidae, in particular of the genera *Sepiola* and *Sepietta*, this study should serve as a warning against indiscriminant use of GenBank data, as well as being a plea for the deposition of voucher specimens in easily accessible institutional collections. DNA barcoding can be advantageous when dealing with cryptic species or for an initial assessment of biodiversity. However, to avoid confusion and error, it should be preceded or combined with a critical taxonomic analysis.

ACKNOWLEDGEMENTS

We are grateful to the Institute for Marine Resources and Ecosystem Studies (IMARES, The Netherlands), especially to the crew and researchers of the Tridens, who provided all of the North Sea Sepiolinae used in this study. Thanks also to Dr A.B. Baird, Dr D.R. Uit de Weerd, Dr J.R. Voight and two anonymous referees for their constructive comments and suggestions.

REFERENCES

BELLO, G. 1998. Cladistic analysis of the Atlanto-Mediterranean sepiolines (Cephalopoda: Sepiolidae) based on morphological characters. *Memorie del Museo Civico di Storia Naturale die Verona (2nd series), Sezione Scienze della Vita*, **13**: 81–85.

BENSASSON, D., ZHANG, D., HARTL, D.L. & HEWITT, G.M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**: 314–321.

FOLMER, O., BLACK, M., HOEH, W., LUTZ, R. & VRIJENHOEK, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**: 294–299.

GRIEG, J.A. 1933. Cephalopods from the west coast of Norway. Communication No. 79 from the Bergen Museum, Biological Station.

GRIMPE, G. 1921. 1. Teuthologische Mitteillungen. Die Sepiolinen der Nordsee. *Zoologischer Anzeiger*, **53**: 1–12.

GRIMPE, G. 1925. Zur Kenntnis der Cephalopodenfauna der Nordsee. *Wissenschaftliche Meeresuntersuchungen*, **16**: 1–122.

HAYWARD, P.J. & RYLAND, J.S. 1990. In: *The marine fauna of the British Isles and North-West Europe*. Vol. 2: *Molluscs to Chordates*, pp. 787–793. Clarendon Press, Oxford.

HEBERT, P.D.N., CYWINSKA, A., BALL, L.S. & DEWAARD, J.R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **270**: 313–321.

JAECKEL, S.G.A. 1958. Cephalopoda. In: *Die Tierwelt der Nord- und Ostsee*. Vol. 37 (G. Grimpe & E. Wagner, eds), pp. 479–723. Akademische Verlagsgesellschaft Geest & Portig K.G., Leipzig.

JANSSEN, A.W. 1975. Systematische lijst van Nederlandse recente en fossiele mollusken. *Mededelingen van de Werkgroep Tertiaire en Kwartaire Geologie*, **12**: 115–170.

JATTA, G. 1896. Cefalopodi viventi nel Golfo di Napoli (Sistematica). *Fauna und Flora des Golfes von Neapel*, **23**: 1–268.

JEREB, P., MAZZOLA, A. & DI STEFANO, M. 1997. Sepiolinae (Mollusca: Cephalopoda) from the Strait of Sicily. *Scientia Marina*, **61**: 459–470.

JONES, B.W., LOPEZ, E.J, HUTTENBURG, J. & NISHIGUCHI, M.K. 2006. Population structure between environmentally transmitted vibrios and bobtail squids using nested clade analysis. *Molecular Ecology*, **15**, 4317–4329.

JOUBIN, L. 1895. Cephalopods from the scientific expedition of Prince Albert I of Monaco. Contribution to the study of cephalopods of the North Atlantic (English translation). Smithsonian Institution Libraries and the National Science Foundation, Washington, DC.

JOUBIN, L. 1902. Révision des Sepiolidae. Mémoires. *Société Zoologique de France*, **15**: 80–145.

LINDGREN, A.R., GIRIBET, G. & NISHIGUCHI, M.K. 2004. A combined approach to the phylogeny of Cephalopoda (Mollusca). *Cladistics*, **20**: 454–486.

MADDISON, D.R. & MADDISON, W.P. 2003. *MacClade 4: analysis of phylogeny and character evolution*. Version 4.08. Sinauer Associates, Sunderland, Massachusetts.

NAEF, A. 1912. Teuthologische Notizen, 3. Die Arten der Gattungen *Sepiola* und *Sepietta*. *Zoologischer Anzeiger*, **39**: 262–271.

NAEF, A. 1923. Die Cephalopoden. *Fauna e Flora del Golfo di Napoli, Monograph*, **35**, 1(Part 2): 149–863.

NISHIGUCHI, M.K., RUBY, E.G. & McFALL-NGAI, M.J. 1998. Competitive dominance among strains of luminous bacteria provides an unusual form of evidence for parallel evolution in sepiolid squid-vibrio symbioses. *Applied and Environmental Microbiology*, **64**: 3209–3213.

NISHIGUCHI, M.K., LOPEZ, J.E. & BOLETZKY, S. 2004. Enlightenment of old ideas from new investigations: more questions regarding the evolution of bacteriogenic light organs in squids. *Evolution and Development*, **6**: 141–149.

NYLANDER, J.A.A. 2004. *MrModeltest v. 2*. Evolutionary Biology Centre, Uppsala University.

PFEFFER, G. 1884. Die Cephalopoden des Hamburger Naturhistorischen Museums. *Abhandlungen aus dem Gebiete der Naturwissenschaften, Hamburg*, **8**: 1–30.

RONQUIST, F. & HUELSENBECK, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**: 1572–1574.

REID, A. & JEREB, P. 2005. Family Sepiolidae. In: *Cephalopods of the world. An annotated and illustrated catalogue of species known to date*. Vol. 1: *Chambered nautiluses and sepioids (Nautilidae, Sepiidae, Sepiolidae, Sepiadariidae, Idiosepiidae and Spirulidae)* (P. Jereb & C.F.E. Roper, eds), pp. 153–203. FAO Species Catalogue for Fishery Purposes, no. 4, vol. 1. FAO, Rome.

RUSSELL, E.S. 1909. Preliminary notice of the Cephalopoda collected by the fishery cruiser 'Goldseeker', 1903–1908. *Annals and Magazine of Natural History, Series 8*, **3**: 446–455.

RUSSELL, E.S. 1922. Report on the Cephalopoda collected by the research steamer 'Goldseeker', during the years, 1903–1908. *Fishery Board for Scotland, Scientific Investigations*, **3**: 1–45.

SWOFFORD, D. L. 2003. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4.0 beta 10. Sinauer Associates, Sunderland, Massachusetts.

TILLIER, A. & BOUCHER-RODONI, R. 1993. Férussac and d'Orbigny's 'Histoire naturelle générale et particulière des Céphalopodes acétabulifères': dates of publication of plates and text. *Nautilus*, **107**: 97–103.