Journal of
Molluscan Studies

The Malacological Society of London

Journal of Molluscan Studies (2024) **90**: eyae028. https://doi.org/10.1093/mollus/eyae028
Published online 12 August 2024

Molecular delimitation of cryptic Australian squid species of the genus *Uroteuthis* Rehder, 1945 (Cephalopoda: Loliginidae), provides a baseline of diversity to resolve classification challenges throughout the Indo-Pacific

Jess A.T. Morgan [1,2], Malcolm C. Dunning[3], Jennifer R. Ovenden [2], Damien Broderick[1,2], Raewyn Street[1,2], Julia Davies[4] and Karina C. Hall [5,6]

[1]*Department of Agriculture and Fisheries, EcoSciences Precinct, PO Box 267, Brisbane, Queensland 4001, Australia;*
[2]*Molecular Fisheries Laboratory, School of Biomedical Sciences, The University of Queensland, St Lucia, Queensland 4072, Australia;*
[3]*Queensland Museum, PO Box 3300, South Brisbane BC, Queensland 4101, Australia;*
[4]*Brisbane, Queensland 4000, Australia;*
[5]*Fisheries Research, New South Wales Department of Primary Industries, PO Box 4321, Coffs Harbour, New South Wales 2450, Australia; and*
[6]*National Marine Science Centre, School of Environment, Science and Engineering, Southern Cross University, PO Box 4321, Coffs Harbour, New South Wales 2450, Australia*

Correspondence: J.A.T. Morgan; e-mail: jessica.morgan@daf.qld.gov.au

(Received 29 November 2023; editorial decision 28 June 2024)

ABSTRACT

This study provides a comprehensive molecular phylogenetic analysis of *Uroteuthis* squid from the Indo-Pacific region. The main aim was to increase sample coverage from northern and eastern Australian waters to resolve the identity and distribution of *Uroteuthis* species taken by local fisheries. Two mitochondrial regions, cytochrome *c* oxidase subunit I (COI) and 16S ribosomal RNA (16S rRNA), were sequenced from 220 new specimens and analysed with a further 51 sequences from GenBank to create a combined phylogeny for the genus. Three nuclear regions, 18S ribosomal DNA (18S rDNA), 28S ribosomal DNA (28S rDNA) and rhodopsin, were also sequenced from representatives of each species. Based on the mitochondrial phylogeny plus distance and tree-based delimitation models, a COI species barcode gap of 4–5% is proposed for discriminating *Uroteuthis* species. Applying this gap partitioned many described species into species complexes; for example, *U. duvaucelii*, *U. noctiluca* and *U. edulis* resolved into 10 species. Although more conserved, mitochondrial 16S rRNA sequences differentiated all new species clades, whereas none of the nuclear markers resolved the closest species. Results confirm that neither *U. chinensis* nor *U. edulis* occurs in Australian waters. Five undescribed species are identified from northern and eastern Australia, of which four are consistent with earlier allozyme studies (and two align with existing DNA sequences). One is a new southeastern, deeper shelf species differentiated in this study, along with a sixth undescribed species from Indonesian waters. Results of the molecular analysis are now being used to inform complementary morphometric analyses for new species descriptions, and genetic stock structure assessments of these important fisheries resources.

INTRODUCTION

Among the soft-bodied coleoid cephalopods (cuttlefishes, squids and octopuses; Mollusca: Cephalopoda), systematics based on morphological characteristics alone has proved particularly challenging (Allcock, Lindgren & Strugnell, 2015). The wider application of molecular approaches has revealed a high prevalence of cryptic species and species complexes among cephalopods formerly described as a single widespread species (Yeatman & Benzie, 1994; Anderson, 2000a, b; Sin, Yau & Chu, 2009; Dai *et al.*, 2012; Anderson *et al.*, 2014; Sales *et al.*, 2014; Krishnan *et al.*, 2022). In addition, global fisheries for cephalopods have expanded rapidly as traditional finfish stocks have been depleted (Arkhipkin *et al.*, 2015) and taxonomic research has not kept pace. About a quarter of the world's squid catches remain unidentified, particularly among coastal loliginid squids (Rodhouse, 2005). Many of the unidentified species arise from the tropical and subtropical waters of the Indo-Pacific region, which span the tropical waters of the Indian Ocean, the western and central Pacific Ocean, and the seas connecting the two in the general area of Indonesia. In this region, species diversity is high and cephalopod taxonomy is particularly poorly defined. This is especially true for loliginid squids of the genus *Uroteuthis* Rehder, 1945, which are widely distributed throughout the region and support valuable jig and trawl fisheries (Dunning, Norman & Reid, 1998).

The generic classification of loliginid squids has changed multiple times since the original description and is still not fully resolved. Most authors recognize the genus *Uroteuthis*; however, some classification systems also include subgenera (e.g. *Photololigo* Natsukari, 1984) (Vecchione *et al.*, 1998, 2005). In this paper, we have omitted using *Uroteuthis* subgenera as the genetic literature does not support *Photololigo* as a monophyletic assemblage (Dai *et al.*, 2012; Sales *et al.*, 2013; Anderson *et al.*, 2014). Members of the genus *Uroteuthis* are the only squids of the family Loliginidae that possess paired photophores (light-emitting organs), which are positioned either side of their intestine, on the ventral surface of the ink sac (Rehder, 1945). Although morphologically distinct (i.e. not possessing photophores), the genus *Loliolus* (*Ls.*) (Steenstrup, 1856) also nests within the genus *Uroteuthis* and likely renders the latter genus paraphyletic (Anderson *et al.*, 2014; Jiang *et al.*, 2018).

DNA sequencing studies focused on phylogenetic relationships of loliginid squids, including Anderson (2000a, b) and Anderson *et al.* (2014), who used two mitochondrial DNA regions, the cytochrome *c* oxidase subunit I (COI = cox1) gene and the ribosomal 16S ribosomal RNA (16S rRNA) to construct phylogenies of the group, flagged the existence of species complexes, particularly for *U. duvaucelii* (d'Orbigny, 1835). The same markers were used by Sin *et al.* (2009) to investigate genetic differentiation among Asian samples of *U. chinensis* (Gray, 1849) and *U. edulis* (Hoyle, 1885). Their phylogeny highlighted that GenBank (National Centre for Biotechnology Information) contains a number of sequences that have been derived from misidentified specimens. Dai *et al.* (2012) confirmed the utility of the COI barcode (Hebert, Cywinska & Ball, 2003) to identify *Uroteuthis* and *Loliolus* species in Chinese waters and noted *Ls. beka* (Sasaki, 1929) was likely a species complex. A phylogeny constructed using an alignment of complete mitochondrial genomes of six loliginid species (Jiang *et al.*, 2018) displays a consistent topology to those constructed using smaller partial gene regions that is consistent with Dai *et al.* (2012) and with only a minor change in the placement of *Ls. japonica* (Hoyle, 1885) by Anderson *et al.* (2014). Only two Australian specimens have been included in the above studies, one referred to as *U.* sp. and the other as *U. etheridgei* (Berry, 1918) (Anderson, 2000a, b; Anderson *et al.*, 2014). In addition, few studies have included more slowly evolving nuclear markers that circumvent the saturation problem (more than one mutation at the same site) of rapidly evolving genes (reviewed by Allcock *et al.*, 2015). Although less reliable for resolving species-level relationships, conserved nuclear markers such as 18S and 28S ribosomal DNA (18S and 28S rDNA) can provide better resolution of deeper nodes in phylogenies (Cracraft & Donoghue, 2004).

In the current study, we aimed to extensively sample northern and eastern Australian waters and include nuclear DNA markers along with standard mitochondrial DNA barcoding to produce a more comprehensive phylogenetic analysis. In Australian waters, *Uroteuthis* species are primarily taken incidentally in commercial trawl fisheries or targeted by recreational fishers using squid jigs. The bulk of squids caught are seldom identified to species but are assumed to belong to one of two morphologically similar species, *U. edulis* or *U. chinensis*. Taxonomic consensus currently recognizes *U. edulis* and *U. chinensis* as widely distributed, extending throughout the Indo-Pacific from Japan to northern and eastern Australia, respectively (Natsukari & Okutani, 1975; Lu, 2001; Jereb & Roper, 2010). Considerable morphological variation is reported among individuals from across both species' distributions (Okutani, 2005; Sin *et al.*, 2009; Takemoto & Yamashita, 2012). An Australian endemic species *U. etheridgei* was described from specimens collected by the *HMS Endeavour*, likely somewhere off the southeast coast, but was later synonymized with *U. chinensis* based on morphological similarities (Natsukari & Okutani, 1975).

An allozyme analyses of specimens from northern Australia in the early 1990s suggested that *U. edulis* and *U. chinensis* (then assigned to the genus *Photololigo*) might be complexes containing several undescribed cryptic species in Australian waters with more limited geographical distributions (Yeatman, 1993; Yeatman & Benzie, 1993, 1994). Two *U. edulis*-like species were found off northern Australia, *U.* sp. 1 in deeper shelf waters and *U.* sp. 2 in shallower coastal waters; and two *U. chinensis*-like species were found off both northern and eastern Australia, *U.* sp. 3 in inshore waters and *U.* sp. 4 (referred to in the 1994 study as *U. chinensis*) in deeper shelf waters (Yeatman, 1993; Yeatman & Benzie, 1993, 1994). Extensive, depth-stratified research trawl sampling in southern Queensland shelf waters by Dunning, Mckinnon & Yeomans (2000) also found shallow depth preferences (<40 m) for *U.* sp. 3 (referred to as *Photololigo etheridgei*) and deeper shelf preferences (>40 m) for the offshore species (referred to as *U.* sp. 4 of Yeatman, 1993).

Recent efforts to prepare stock and risk assessments for loliginid squids in eastern Australian waters have been hampered by ongoing species identification issues and their confused taxonomy. The aim of this research was to clarify the taxonomy of *U. chinensis* and *U. edulis* in Australian waters using more modern molecular methods and delimit any undescribed cryptic species, and their likely distributional ranges.

MATERIAL AND METHODS

Sample collection

Fresh samples were collected between July 2015 and March 2018 directly from commercial fishers at various locations along the eastern and northern Australian coasts. Samples were also sourced in 2015 from local fishers and markets in Taiwan and Hong Kong. Historical tissue samples from Australia, China and Indonesia collected during earlier studies (Yeatman, 1993; Dunning *et al.*, 2000) and subsequent opportunistic sampling from bycatch of commercial, demersal prawn and fish trawl catches (M. Dunning, unpubl.) were also extracted and sequenced to supplement samples collected during the current study. Specimens were either stored frozen at −20 °C or preserved in 200 proof molecular grade ethanol, then stored at −20 °C. An additional 51 DNA sequences were sourced from GenBank for phylogenetic comparison. Details of the identity and origin of the 220 new samples extracted in the current study are provided in Table 1. More extensive information relating to all samples used in the phylogenetic study are provided in Supplementary Material Table S1. This table also includes the Australian Museum accession numbers for voucher specimens lodged for several of the putative species.

DNA extraction

Total genomic DNA was isolated from approximately 25 mg of mantle tissue from each specimen. Ethanol-preserved tissue was soaked in 1 ml of milli-Q water for 1 h prior to extraction to remove the preservative. DNA was extracted either using a DNeasy Blood and Tissue Kit (Qiagen, Chadstone Victoria) following the manufacturer's instructions for 'Purification of Total DNA from Animal Tissues' or using a salting-out protocol designed to retain high-molecular-weight DNA (Miller, Dykes & Polesky, 1988). Briefly, for the salting-out protocol, samples were soaked overnight at 37 °C in lysis buffer (50 mM Tris-HCl, 400 mM NaCl, 5mM EDTA pH 8.2 and 2% SDS) with 60 µg proteinase K solution. After lysis, the samples were treated with 30 µg RNase A and then proteins were precipitated with the addition of one-third volume of saturated salt (5M NaCl). Genomic DNA was isopropanol precipitated, pelleted and then washed with 70% ethanol prior to resuspension in 100 µl of TE (pH 8.0) buffer. Once extracted, DNA concentration in all samples was quantified using a Nanodrop 100 spectrophotometer (Thermo Scientific, Australia).

Table 1. Samples of *Uroteuthis* species collected from the Indo-Pacific, northern and eastern Australian waters and sequenced for this study.

Species and collection location	Number of samples sequenced	
	Total	Breakdown
***Uroteuthis* sp. 1, Australia**	**16**	
Arafura Sea, NT		12
North West Shelf, WA		4
***Uroteuthis* sp. 2, Australia**	**10**	
Gulf of Carpentaria, NT		8
Timor Sea, NT		2
***Uroteuthis* sp. 3**		
Indonesia	**2**	
Australia	**60**	
Gulf of Carpentaria, NT		21
Townsville, QLD		4
Bundaberg, QLD		4
Moreton Bay, QLD		13
Wallis Lake, NSW		1
Port Stephens, NSW		5
Hawkesbury River, NSW		10
Port Jackson, NSW		2
***Uroteuthis* sp. 4**		
Indonesia	**3**	
Australia	**38**	
North West Shelf, WA		4
Gulf of Carpentaria, NT		14
Torres Strait, QLD		3
Princess Charlotte Bay, QLD		10
Cairns, QLD		2
Townsville QLD		5
***Uroteuthis* sp. 5, Australia**	**43**	
Townsville, QLD		1
Swains Reef, QLD		6
Bundaberg, QLD		3
Southport, QLD		11
Iluka, NSW		10
Coffs Harbour, NSW		6
Port Stephens, NSW		6
***Uroteuthis* sp. 6, Indonesia**	**2**	
***Uroteuthis chinensis*, Asia**	**11**	
Taiwan		5
China		4
Indonesia		2
***Uroteuthis duvaucelii*, Asia**	**5**	
Taiwan		4
Indonesia		1
***Uroteuthis sibogae* unknown, market import**	**1**	
***Uroteuthis edulis*, Taiwan**	**19**	
***Uroteuthis noctiluca*, Australia**	**5**	
Gulf of Carpentaria, NT		3
Princess Charlotte Bay, QLD		1
Moreton Bay, QLD		1
Sepioteuthis australis*, Australia	**5**	
Western Australia		1
Port Stephens, NSW		4
Grand total	**220**	

**Sepioteuthis australis* samples were initially included to root the phylogenetic trees but were later replaced with closer genetic relatives.

PCR amplification and sequencing

Two mitochondrial DNA regions, the COI gene and 16S rRNA were sequenced from new specimens for species diagnostics and phylogenetic comparisons alongside *Uroteuthis* samples available on GenBank. Three partial nuclear regions, the rhodopsin gene and nuclear 18S and 28S rDNA were sequenced for a representative of each species to determine the resolving power of these markers for species-level comparisons among the Australian *Uroteuthis* species of interest. Amplification reactions were conducted in 10 μl volumes containing 1 μM of each primer pair (Table 2), 10–100 ng of extracted DNA, 10× PCR buffer (Qiagen, Chadstone Victoria, containing 25 mM magnesium), 1 mM dNTP and 1 unit of Taq DNA polymerase (Qiagen, Chadstone Victoria). Thermal cycling conditions consisted of an initial denaturation (95 °C for 2 min) followed by 35 cycles of 95 °C for 30 s, annealing at 48 °C or 50 °C (see Table 2) for 30 s and extension at 72 °C for 1 min 30 s, with a final extension step of 72 °C for 7 min. Cycling was performed in a Bio-Rad thermal cycler (DNA Engine Peltier, Bio-Rad, Gladesville, New South Wales, Australia). PCR products were viewed on a 3% agarose TBE gel stained with GelRed (Biotium, Hayward, CA, USA). Prior to sequencing, unwanted dNTPs and primers were removed from PCR products using ExoSAP-it® (USB Corporation distributed by GE Healthcare Bio-Sciences, Rydalmere New South Wales, Australia). Approximately 20 ng of the PCR product was used in standard ABI Dye Terminator sequencing reactions using Big-Dye Vers. 3.1 technology (Thermo Fisher Scientific, Life Technologies Australia) and was run on an ABI 3130xl Genetic Analyser (Applied Biosystems, CA, USA, now Thermo Fisher Scientific at the Australian Equine Genetics Research Centre, now Genetic Research Services, at the University of Queensland, Brisbane, Australia). Forward and reverse sequences, amplified using the same primers as the initial PCR, were edited and aligned using Sequencher v. 5.4.6 (Gene Codes Corporation, Ann Arbor, MI, USA) and then exported as a single Fasta contig. Where sequence quality was poor, samples were re-amplified and sequenced again.

Sequence alignment and phylogenetic tree construction

A single summary sequence for the 18S rDNA, 28S rDNA, rhodopsin, COI and 16S rRNA of each sample was exported from Sequencher into Geneious R11.1.2 (https://www.geneious.com) where sample alignments were constructed using the ClustalW v. 2.1 algorithm (Larkin *et al.*, 2007) with default settings (cost matrix IUB, gap open cost 15, gap extend cost 6.66). Standard nucleotide BLAST searches of all sequences were conducted in GenBank. If identical sequences were found on GenBank, a single representative of the sequence was included in phylogenetic analyses. Sequences from other closely related species, and unique sequences from related isolates of the same species, were also downloaded from GenBank to include in phylogenetic analyses (see Supplementary Material Table S1).

The ClustalW output was manually edited by eye prior to phylogenetic analysis to ensure codon reading frames were maintained across the alignment (COI and rhodopsin), and gaps minimized. The invertebrate mitochondrial code (GenBank translation table 5) was used for COI translations. Initial phylogenetic analyses of the mitochondrial genes were rooted with *Sepioteuthis australis* (Quoy & Gaimard, 1832) following Dai *et al.* (2012) and Anderson *et al.* (2014). However, this outgroup was difficult to align, particularly through the hypervariable mitochondrial DNA 16S rRNA, and following reviewer feedback, the data were re-analysed using closer relatives, *Loligo* (*Lg.*) *forbesii* Steenstrup, 1856, and *Lg. vulgaris* (Lamarck, 1798). Even using the closer outgroups, 87 bp of 16S rRNA hypervariable sequence remained difficult to align, so positions 11–16, 209–239, 275–292, 309–314 and 339–364 in the 16S rRNA alignment were excluded prior to phylogenetic analysis. The sequence alignments used in this study are publicly available on Dryad (doi.org/10.5061/dryad.gf1vhhmvr).

Table 2. Primers and their annealing temperatures for the amplification of mitochondrial and nuclear DNA products with the maximum product size amplified and the final alignment length used for phylogenetic analyses.

Primer	Sequence 5′–3′	Ta (°C)	Max (bp)	Length in final alignment (bp)
Mitochondrial DNA				
COI				
COIF[1]	TITCIACIAAYCAYAARGAYATTGG	48	664	592
sqCOIR[1]*	TAIACYTCDGGGTGWCCAAARAATCA			
16S rRNA				
16S-S1[2]	CGCCTGTTTATCAAAAACAT	50	534	510
16S-S2[2]*	CCGGTCTGAACTCARATCAYGT			
Nuclear DNA				
18S rDNA				
18S-1250F[3]*	TGGCCGTTCTTAGTTGGTGGAG	50	856	769
18S-1800R[3]	GATCCTTCCGCAGGTTCACCTACG			
28S rDNA				
28S-22F[4]	GWRRYWACCCGCTGAAYTWAAGC	50	1726	
SQ28S-750F†	CACGGTCGGCACCGGACGCA	50	938	668
28S-1300R[4]	CTGGCGATCGATTTGCACGTCAG			
Rhodopsin				
Sq-Rho-F2†	GTCCTYTGCAATTGCTCYTTTGATTA	50	526	509
Sq-Rho-R2†	CATTTCYTTCAYYTGGGCRGCATC			

Sources: 1, Geller *et al.* (2013); 2, Palumbi (1996) (referred to as S1 and S2); 3, Raupach *et al.* (2009); and 4, Barton & Morgan (2016).
*Modified from reference to be more specific to squid DNA.
†Authors.

Before conducting phylogenetic analyses, optimal substitution models were determined in ModelTest-NG v. 0.1.7 (Flouri *et al.*, 2015; Darriba *et al.*, 2020) using Bayesian information criterion (BIC) scores to rank 88 different substitution models. A maximum likelihood (ML) statistical method was used based on an initial maximum parsimony tree using all sites. The substitution model with the lowest BIC score was considered optimal.

A homogeneity partition test (PAUP) of the three nuclear regions found no significant difference ($P = 0.26$), so they were concatenated for phylogenetic analysis. Gap columns were stripped from the alignment prior to analysis reducing 18S rDNA to 692 bp and 28S rDNA to 620 bp. The bestfit model for the nuclear markers was a Tamura–Nei model (Tamura & Nei, 1993) with a discrete gamma distribution to describe among-site heterogeneity with an alpha shape parameter, G4 = 0.05 (summarized as TrN + G4).

For the mitochondrial markers, the bestfit model for the COI data alone was a general time-reversible model (Tavaré, 1986) with gamma shape, G4 = 2.161, and estimates of invariant sites, I = 0.594 (summarized as GTR + I + G4). For the 16S rRNA data alone, the bestfit model was a Hasegawa, Kishino & Yano (1985) model (Hasegawa *et al.*, 1985) with a gamma shape, G4 = 0.501, and invariant sites, I = 0.69 (summarized as HKY + I + G4). The model, selected for the concatenated COI plus 16S rRNA alignment of 1,102 bases and 116 unique sequences, was a Kimura 3-parameter model with unequal base frequencies (TPM3uf) (Kimura, 1981) plus a gamma shape, G4 = 1.22, and estimates of invariant sites, I = 0.62 (summarized as TPM3uf + I + G4).

ML phylogenetic trees were constructed using COI and 16S rRNA datasets alone, and then, because no major conflict was observed between the single-gene trees (changes in branch associations were only observed at nodes with poor branch support, <60%), on a concatenated dataset using PAUP* v. 4.0a (build 169) (Swofford, 2003). Starting trees were obtained via random stepwise addition. Other settings used were Mulpars in effect, Maxtrees set to 150 (limited by computational time), one heuristic search repetition and tree-bisection-reconnection (TBR) branch swapping. Starting branch lengths were obtained using the Rogers–Swofford approximation method. Branch length optimization was set to one-dimensional Newton–Raphson with pass limit equal to 20 and tolerance of 1e-007. Likelihood calculations were performed in single precision and vector processing was enabled. Likelihoods were computed using standard Felsenstein pruning, and the conditional-likelihood rescaling threshold was set to 1e-020.

Branch support was determined by bootstrapping a likelihood analysis using 1,000 replicates (Felsenstein, 1985) in MEGA 7.0.26 (Kumar, Stecher & Tamura, 2016) and by Bayesian analysis using the MrBayes v. 3.2.6 plugin (Huelsenbeck & Ronquist, 2001) in Geneious. For the Bayesian analysis, a GTR + G4 substitution model was the closest model available for the concatenated dataset, so was used with four gamma categories. The Markov chain Monte Carlo (MCMC) settings were chain length 1,100,000, subsampling frequency 1,000 and burnin 250,000. Four heated chains and one cold chain were run and run length was determined using a minimum effective sample size of 200. Priors for the analysis were unconstrained branch lengths where tree length was associated with a gamma distribution with mean 10, and branch length proportions were associated with a uniform Dirichlet distribution.

Reciprocal monophyly, the barcode gap and determination of putative species

Species boundaries were delimited using three methods: tree topologies, automatic barcode gap discovery (ABGD) (Puillandre *et al.*, 2012) and Bayesian implementation of the Poisson tree processes model for species delimitation (bPTP) (Zhang *et al.*, 2013). For the tree topology method, the concatenated COI plus 16S rRNA phylogenetic tree was used. Monophyletic assemblages that contained recognized, published species sequences provided the baseline species groupings. These groupings gave an indication of the sequence divergence (branch lengths) needed to differentiate new species and flagged new putative species/clades of interest. The ABGD method was applied to uncorrected pairwise COI distances

only. These sequences were interrogated for an unoccupied range or break, referred to as the barcode gap or species threshold (proposed by Hebert *et al.*, 2003, revised by Meier, Zhang & Ali, 2008). A DNA barcode gap can be quantified as a strict measure (maximum intraspecific distance to minimum interspecific distance) or as an average measure (mean intraspecific distance to mean interspecific distance) (Meier *et al.*, 2008). Here, we use the term to represent the strict measure unless otherwise stated.

The ABGD method provides a species partition hypothesis based on pairwise differences without any prior knowledge of species genealogy (Puillandre *et al.*, 2012). No standard limit or range exists for a COI barcode gap; it must be calculated for each group of organisms. It should also be noted that very recent speciation events may not have given species sufficient time to diverge enough to display a barcode gap (Puillandre *et al.*, 2012). Default settings used for the ABGD analysis were: the maximum prior intraspecific distance (P) range set between 0.001 and 0.1 divergence (i.e. the approximate indication of the area where the barcode gap should be detected); 10 steps; 20 bins; and simple distance. The relative gap width (X), a measure of sensitivity, was set to 1 for the analysis. By integrating tree topology and the barcode gap: (1) existing species boundaries were confirmed, (2) suspected and previously reported species complexes were identified and (3) new putative species (molecular operational taxonomic units) were determined. The effect of using a DNA barcode gap to define *Uroteuthis* species was created manually in GenStat v. 22.1.0.167 (VSN International, Hemel Hempstead, UK) as a stacked bar graph by plotting the relative frequency of each distance for each comparison category (population or species) against pairwise distance.

A concatenated COI plus 16S rRNA ML tree was used as the phylogenetic input tree for the bPTP analysis. The bPTP analysis was run for 500,000 generations of MCMC with thinning set to 100 and a burnin of 20%. Convergence of the MCMC chains was confirmed via visual checking the likelihood plot of each delimitation.

Ethical approval

Animals sourced for this study either were harvested by commercial fishers or were sourced from historical tissue collections from research surveys.

RESULTS

The three nuclear DNA markers were too conserved to reliably differentiate among closely related species of *Uroteuthis*. For this reason, the nuclear markers were only sequenced from a small subset of the samples. Both the mitochondrial DNA COI (592 bp) and 16S rRNA (510 bp) were sequenced for all 220 extracted tissue samples. Seven of the 51 samples downloaded from GenBank were sequenced for only one of the two mitochondrial DNA markers but were included in the analyses because of their unique collection location or sequence. A further seven samples downloaded had incomplete coverage across the two markers but were again retained because of their relevance to the study. The negative effects of including taxa with missing or poor-quality sequence data are inflated divergence measures and the inclusion of possibly false signal. The value of including these additional taxa is that they may be reliably placed on the tree, provided they contain sufficient genetic signal. They can also improve the accuracy of genetic trees by subdividing long branches, which can cause systematic errors due to long-branch attraction (Wiens, 2006).

Genus-level phylogeny

The mitochondrial DNA COI sequences were the most variable (93 unique sequences with *Uroteuthis* species divergence ranging from 5% to 20%) followed by the mitochondrial DNA 16S rRNA (61 unique sequences with *Uroteuthis* species divergence ranging from 0.74% to 12%). Among the nuclear DNA markers, 28S rDNA was the next most variable (*Uroteuthis* species divergence ranging from 0.16% to 10%), followed by 18S rDNA (*Uroteuthis* species divergence ranging from 0.21% to 8%) and lastly, the rhodopsin gene (*Uroteuthis* species divergence ranging from 0% to 0.79%). Both mitochondrial DNA datasets were phylogenetically congruent, although the higher divergence of COI provided greater resolving power to differentiate among closely related *Uroteuthis* species. The concatenated COI plus 16S rRNA alignment consisted of 274 samples of which 116 had unique sequences. An ML mitochondrial DNA phylogeny of the genus, grouping and labelling species by their current classification, is shown on the left in Fig. 1. The ML nuclear DNA phylogeny, based on a reduced sample dataset, is mirrored to the right of the mitochondrial phylogeny in Fig. 1. The nuclear phylogeny supports the separation of *U. edulis* and *U.* sp. 1 from the group containing *U. chinensis* and *U.* sp. 2 through *U.* sp. 6 but lacks the power to resolve species within this latter group.

The species-level topology of the single mitochondrial DNA tree is consistent with that of the strict consensus tree of the 150 trees retained. Species grouped together in well-supported clusters on the tree. Within-species divergence, however, was extremely high for some species. Poor branch support at some nodes (Fig. 1) reflected inconsistencies between the topology of the COI and 16S rRNA trees. The relative positions of *U. noctiluca* (Lu, Roper & Tait, 1985), the *Loliolus* group and the group containing *U. sibogae* (Adam, 1954), *U. singhalensis* (Ortmann, 1891) and *U. duvaucelii* are not clear (Supplementary Material Figs S1, S2). In the 16S rRNA tree (Supplementary Material Fig. S2), *U. singhalensis* falls within the *U. duvaucelii* group with *U. sibogae* more distant. Relationships within the group containing *U. chinensis* and *U.* sp. 2 through to *U.* sp. 6 are poorly resolved in the 16S rRNA tree (Supplementary Material Fig. S2), but mutations in the 16S rRNA marker have impacted the placement of *U. chinensis* in the analysis of the concatenated dataset. In the COI tree, *U.* sp. 6 is sister to *U. chinensis*.

Species complexes and new species boundaries

Both the phylogeny and a comparison of pairwise genetic distances between and within the currently defined species highlight the presence of species complexes within the genus. A species complex represents a group of closely related species often so similar in appearance that they are mistakenly grouped into one species despite their inability to interbreed. In genetic terms, population differences (intraspecific divergence) are expected to be smaller than species differences (interspecific divergence). Using the current classification there is considerable overlap: COI intraspecific divergence is 0–9% and interspecific divergence ranges from 6% to 12% (Table 3). Pairwise divergence measures for all the genetic markers are presented in Table 3 for comparison. Species diagnostics for *Uroteuthis* are based on the mitochondrial markers only.

The ABGD analysis of the COI data differentiated 24 groups (19 *Uroteuthis* and 5 *Loliolus*) with a gap at 1–3% (Table 3). The bPTP analysis of the concatenated COI and 16S rRNA data supported 22 of the 24 groups identified in the ABGD analysis and further divided *U. noctiluca* clade 1 and *U. chinensis* into two species each (Table 3). The splitting of *U. noctiluca* clade 1 and *U. chinensis* is likely an artefact of poor-quality sequences from GenBank.

We propose a conservative COI barcode gap of 4–5% for *Uroteuthis*, for reasons detailed below. With a 4–5% COI barcode gap, the *Uroteuthis* included in this study likely represents 19 species (Table 3, Figs 2–6). The four undescribed species differentiated using allozymes (Yeatman, 1993; Yeatman & Benzie, 1993), *U.* sp. 1 to *U.* sp. 4, were supported, and an additional two species were differentiated, *U.* sp. 5 and *U.* sp. 6.

By applying a barcode gap to COI to define *Uroteuthis* species (Fig. 7), the overlap between populations and species is removed,

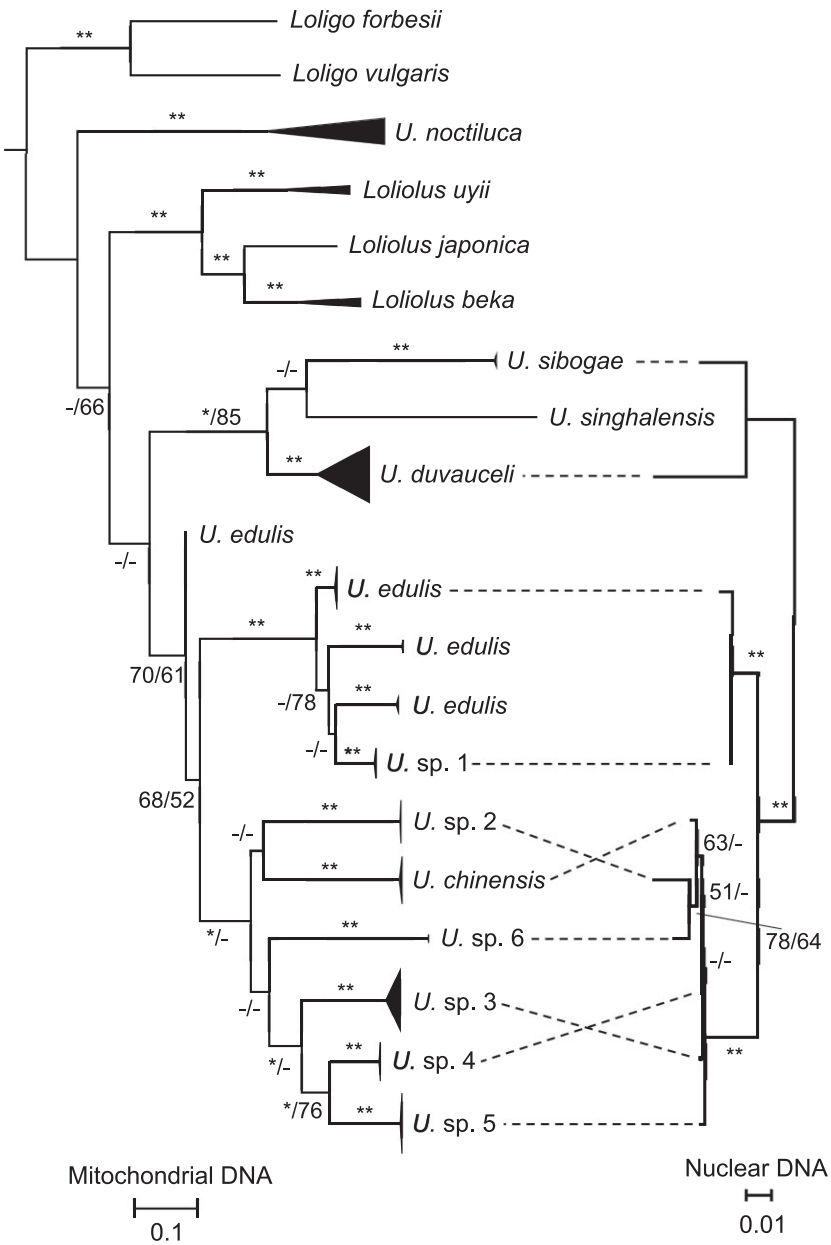

Figure 1. Summary ML phylogeny of *Uroteuthis* showing the mitochondrial DNA (COI + 16S rRNA) tree on the left (Ln = −7906.989) and the nuclear DNA (18S rDNA, 28S rDNA and rhodopsin) tree on the right (Ln = −3755.877). The mitochondrial phylogeny on the left was constructed with *Loligo forbesii* and *Lg. vulgaris* as an outgroup. For Figures 1–6 trees, branch support is shown as Bayesian then bootstrap values with * indicating greater than 90% support, a number if support was 50–90 and a dash (–) if branch support was less than 50; the trees are drawn to scale, with branch lengths measured as the estimated number of substitutions per site. Clades representing a single species are shown as collapsed branches (black triangles). The width of the triangles represents branch length, and the height reflects the number of nodes. For the trees shown in Figures 2 to 6 trees, the sequences from this study are in bold. A superscript symbol following a taxon label indicates that identical 16S rRNA and COI sequences were obtained from multiple individuals (number in brackets) from the same location. These link to individual samples in Supplementary Material Table S1, which provides detailed information on the samples, their origin and GenBank accession numbers. Where samples from the type locality have been sequenced, and they fall within a single clade, the clade is underlined to indicate this will likely become the type sequence for the species.

and for the bulk of species, intraspecific divergence (blue) is less than 1% (average maximum population difference to average minimum species difference is 0.65–7.4%). The 3% divergence recorded in *U. noctiluca* clade 1 is probably an artefact of a partial, poor-quality GenBank sequence (Table 3, Fig. 2). The one stand-out species with 3.6% COI intraspecific divergence is *U.* sp. 3 (Table 3, Fig. 5). The two *U.* sp. 3 clades could not be differentiated using 16S sequences.

Uroteuthis duvaucelii, U. noctiluca, Ls. beka *and* Ls. uyii *(Wakiya & Ishikawa, 1921) species complexes*

Uroteuthis duvaucelii, *U. noctiluca*, *Ls. beka* and *Ls. uyii* (Wakiya & Ishikawa, 1921) all formed species complexes, with four clades (putative species represented by two or more sequences) resolved for *U. duvaucelii* and two clades or variants (putative species represented by a single sequence) for each of the other species (Fig. 2). The clade representing the type species of *U. duvaucelii* could not be

Table 3. Percentage divergence of mitochondrial and nuclear markers for within and between species comparisons (including putative species clades from Figs 2–6 and their support based on ABGD and bPTP analyses) of *Uroteuthis and Loliolus*.

Species	Maximum intraspecific COI divergence	Minimum interspecific COI divergence	Minimum interspecific 16S rRNA (28S rDNA, 18S rDNA and rhodopsin) divergence	Species supported by ABGD analysis of COI data	Species supported by Bayesian bPTP analysis of COI + 16S rRNA data
Loligo vulgaris	Na	10%	9%		
Lg. forbesii	Na	10%	9%		
Loliolus beka	6%	9%	2.95%		
Ls. beka variant 1	Na	6%	2.95%	Yes	1
Ls. beka variant 2	Na	6%	1.72%	Yes	1
Ls. japonica	Na	9%	1.47%	Yes	1
Ls. uyii	7%	10%	1.47%		
Ls. uyii variant 1	Na	7%	1.47%	Yes	1
Ls. uyii variant 2	Na	7%	1.47%	Yes	1
Uroteuthis noctiluca	8%	12%	6%		
U. noctiluca clade 1	2.68%*	8%	4%	Yes	2 species†
U. noctiluca clade 2	0.17%	8%	4%	Yes	0.73†
U. sibogae	0.34%	8%	7% (10%, 5%, 0.79%)	Yes	1
U. singhalensis	0.34%‡	9%	5%	Yes	1
U. duvaucelii	9%	9%	4%		
U. duvaucelii clade 1	0.34%	7%	1.72%	Yes	0.99
U. duvaucelii clade 2	0.34%	5%	1.23%	Yes	0.87
U. duvaucelii clade 3	0.17%	5%	0.25% (10%, 8%, 0.79%)	Yes	0.99
U. duvaucelii clade 4	0.17%	5%	0.25%	Yes	0.86
U. sp. 1	0.51%	6%	0.74% (0.61%, 1.15%, 0.2%)	Yes	0.91
U. edulis	7%	6%	0.74%		
U. edulis clade 1	0.34%	6%	0.74%	Yes	0.99
U. edulis clade 2	0.85%	6%	0.74% (0.61%, 1.15%, 0.2%)	Yes	0.75†
U. edulis variant 3	Na	Na	0.74%	Na	1
U. edulis clade 4	0.17%	6%	Na	Yes	0.99
U. sp. 6	0.17%	10%	3.93% (2.82%, 1.21%, 0.39%)	Yes	0.93
U. sp. 2	0.68%	11%	2.45% (2.93%, 1.45%, 0.39%)	Yes	0.93
U. chinensis	1%	10%	2.45% (0.16%, 1.21%, 0%)	Yes	2 species†
U. sp. 3	3.60%	8%	1.47%		
U. sp. 3 north	0.74%	3%	0.25% (1.03%, 0.21%, 0%)	Yes	0.92
U. sp. 3 south	0.17%	3%	0.25%	Yes	0.93
U. sp. 4	0.68%	6%	0.74% (0.16%, 0.89%, 0%)	Yes	0.87
U. sp. 5	0.85%	6%	0.74% (0.47%, 0.21%, 0.2%)	Yes	0.86

Abbreviation: Na, comparison not possible due to just one sequence or no sequences being available.
*Partial poor-quality GenBank sequence likely inflating value.
†Incomplete or poor-quality GenBank sequence likely deflating value or splitting group.
‡Determined during manuscript revision with Krishnan *et al.* (2022) publication of *U. singhalensis* sequences.

specified, as three of the four *U. duvaucelii* clades (1, 3 and 4) have been sampled from the Indian coast (Krishnan *et al.*, 2022), the species type locality. *Uroteuthis noctiluca* clade 2 likely represents the type species by Lu *et al.* (1985). These clade 2 squid are restricted to shallow coastal bays and estuaries and are less than 90 mm in length. *Uroteuthis noctiluca* clade 1 squid are larger, can be distinguished morphologically, and are commonly caught incidentally during continental shelf trawling (Dunning, 1998).

The U. edulis *species complex and* U. *sp. 1*

The *U.* sp. 1 samples collected from northern Australia form a well-supported clade within the *U. edulis* species complex (Fig. 3). The three distinct *U. edulis* clades in the complex (1, 2 and 4) originate from various eastern and southern Asian localities. The type specimen of *U. edulis* (Hoyle, 1886) was collected from Yokohama Market, Japan, suggesting *U. edulis* clade 2 will likely represent the type species. Samples from India and Iran group together to form *U. edulis* clade 1, but the identity of these specimens requires further investigation as currently only COI sequences are available for the Indian samples. Originating from China, *U. edulis* clade 4 is represented by only COI data. Another sample of Chinese origin, *U. edulis* variant 3 does not fall within the *U. edulis* complex. The strange phylogenetic placement of this animal may reflect that this variant is represented by only a 16S rRNA sequence. The sample was included in the phylogeny because the 16S sequence was quite different from the other *U. edulis* squid (Supplementary Material Fig. S2). Further specimens sequenced for both COI and 16S rRNA are needed to determine whether *U. edulis* variant 3 and clade 4 represent separate species.

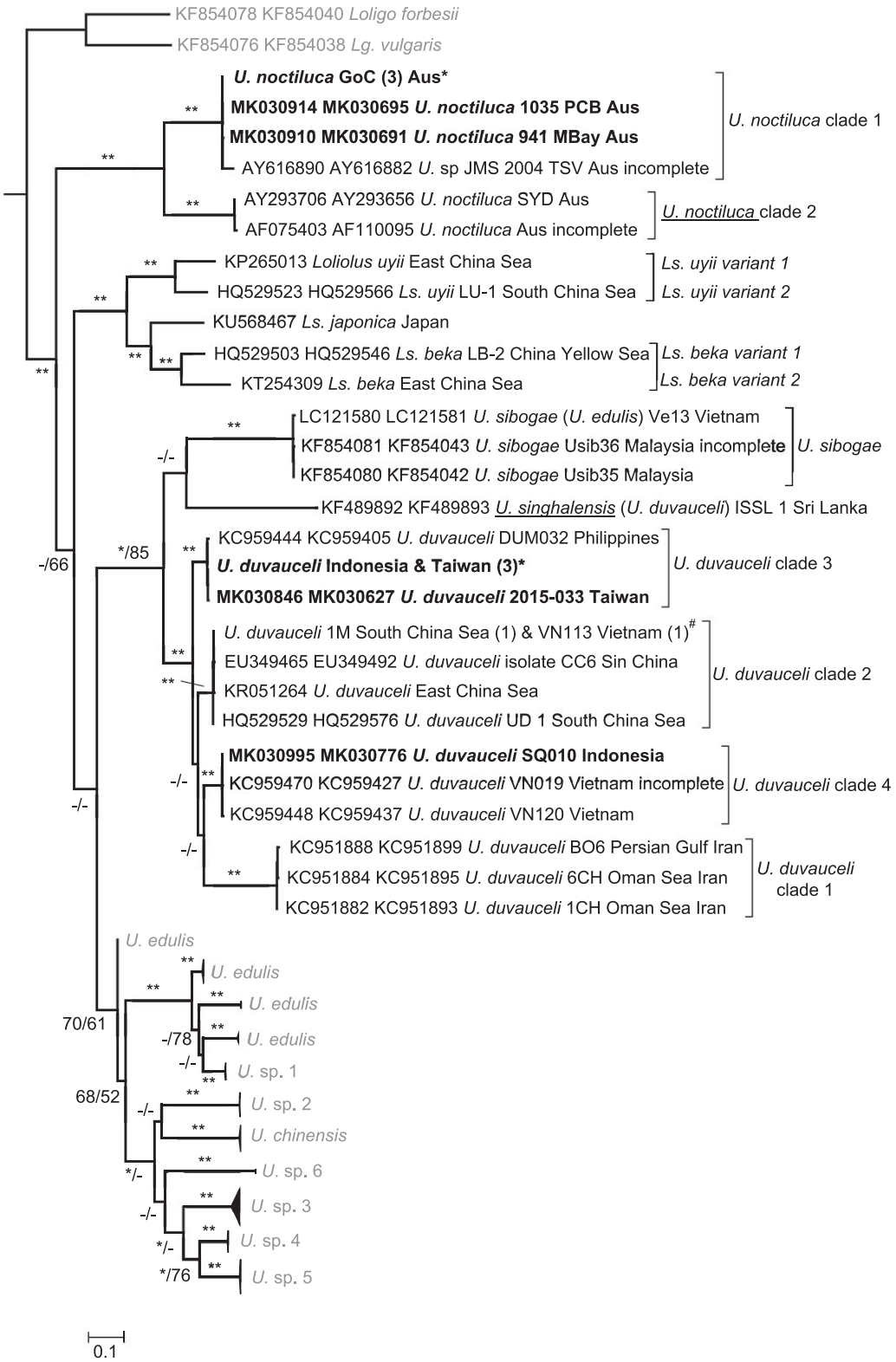

Figure 2. ML COI + 16S rRNA phylogeny of *Uroteuthis* expanding *U. sibogae*, *U. duvaucelii*, *U. noctiluca*, *Ls. beka* and *Ls. uyii* clades and variants from Figure 1. See Figure 1 caption for more detail.

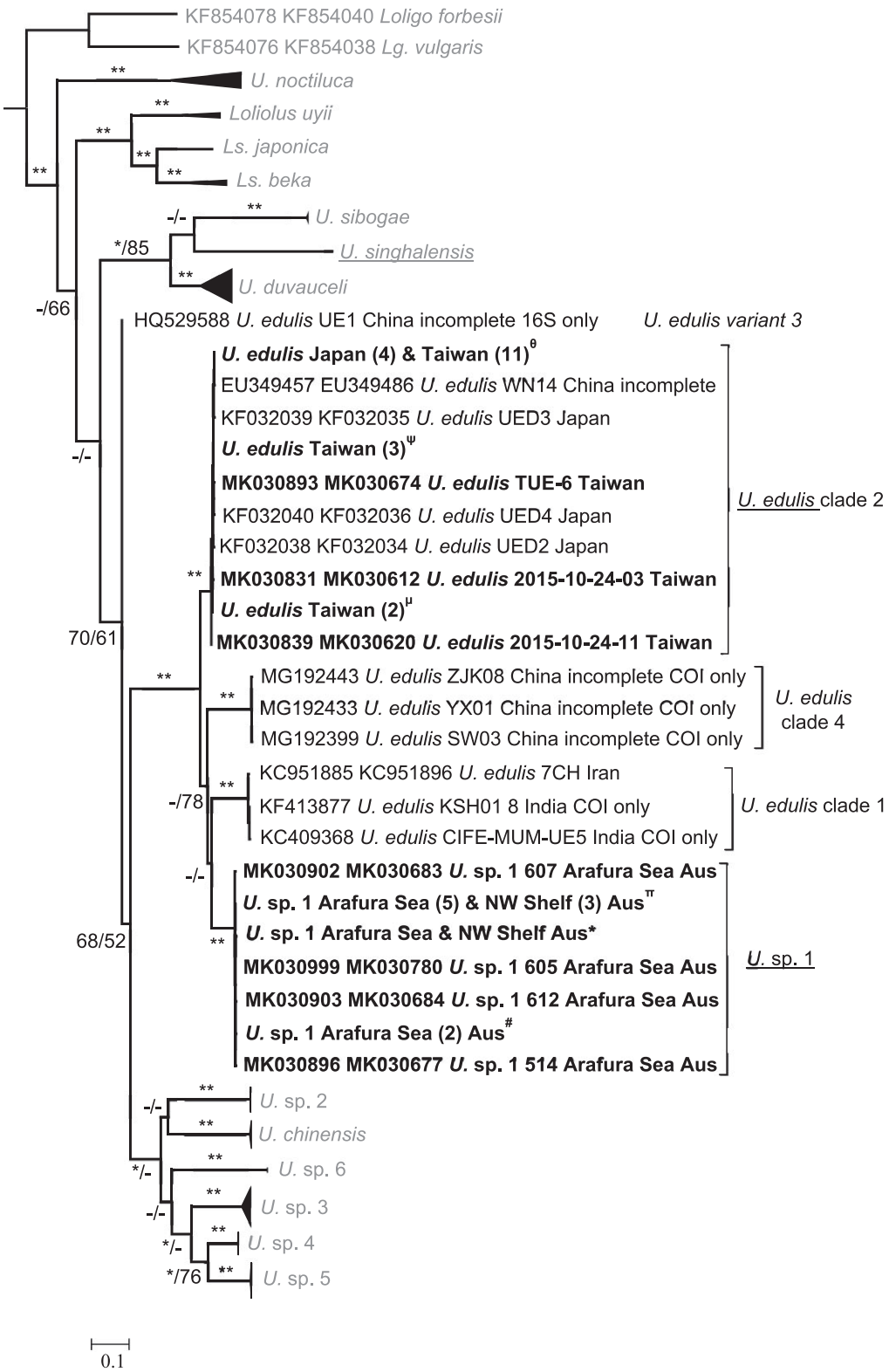

Figure 3. ML COI + 16S rRNA phylogeny of *Uroteuthis* expanding *U. edulis* and *U.* sp. 1 clades and variants from Figure 1. See Figure 1 caption for more detail.

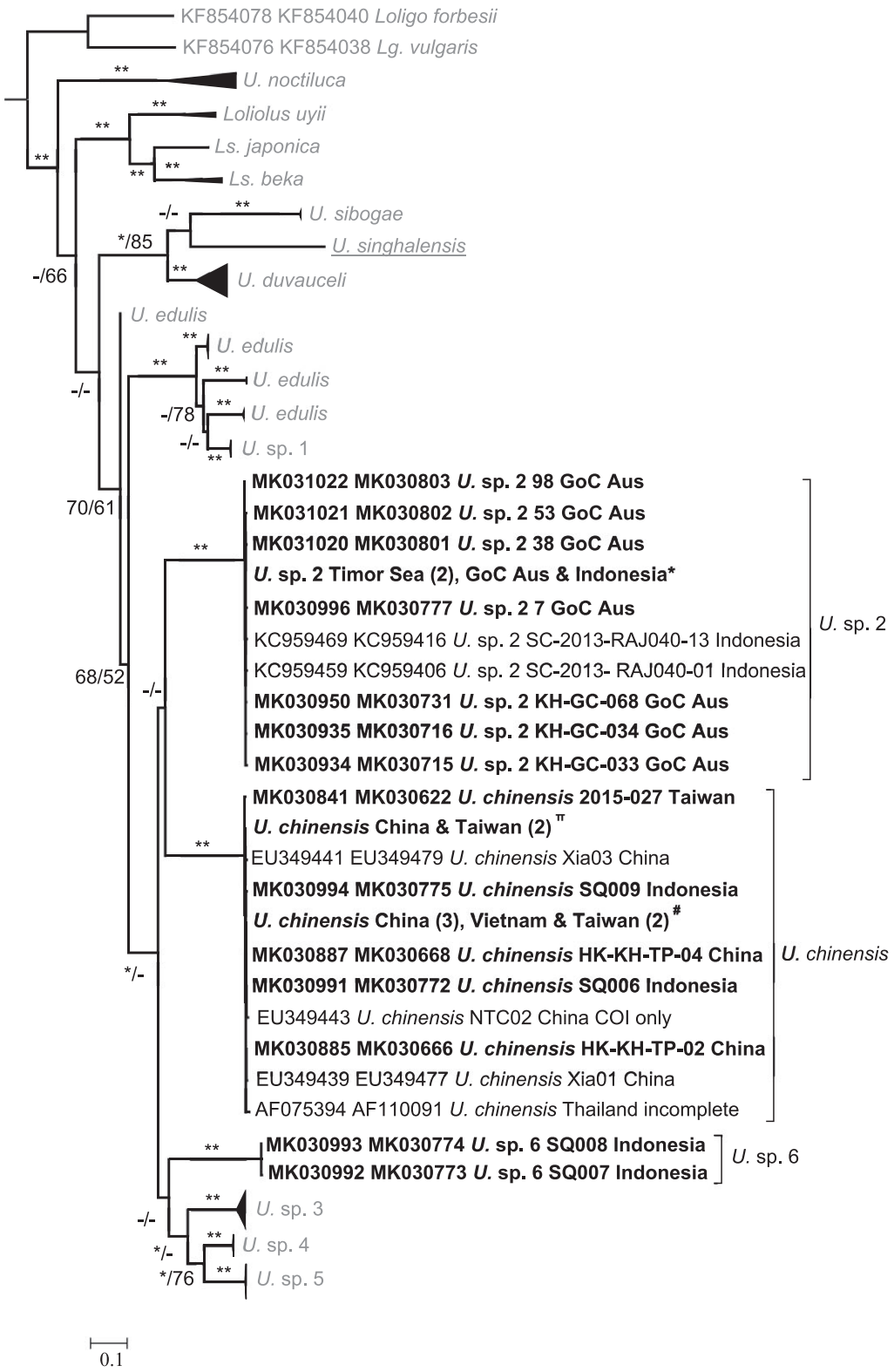

Figure 4. ML COI + 16S rRNA phylogeny of *Uroteuthis* expanding *U.* sp. 6, *U.* sp. 2 and *U. chinensis* clades from Figure 1. See Figure 1 caption for more detail.

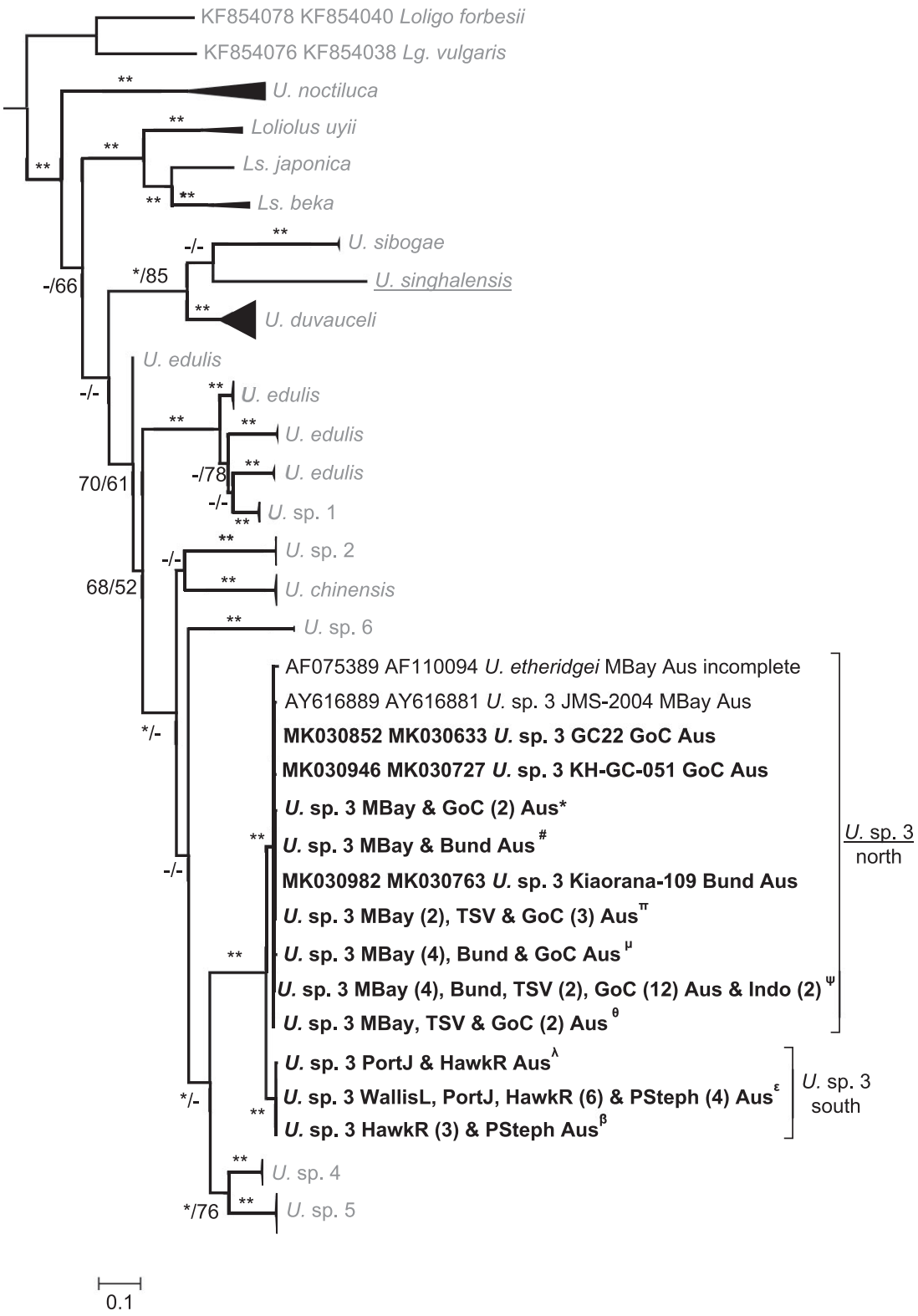

Figure 5. ML COI + 16S rRNA phylogeny of *Uroteuthis* expanding *U.* sp. 3 clade from Figure 1. See Figure 1 caption for more detail.

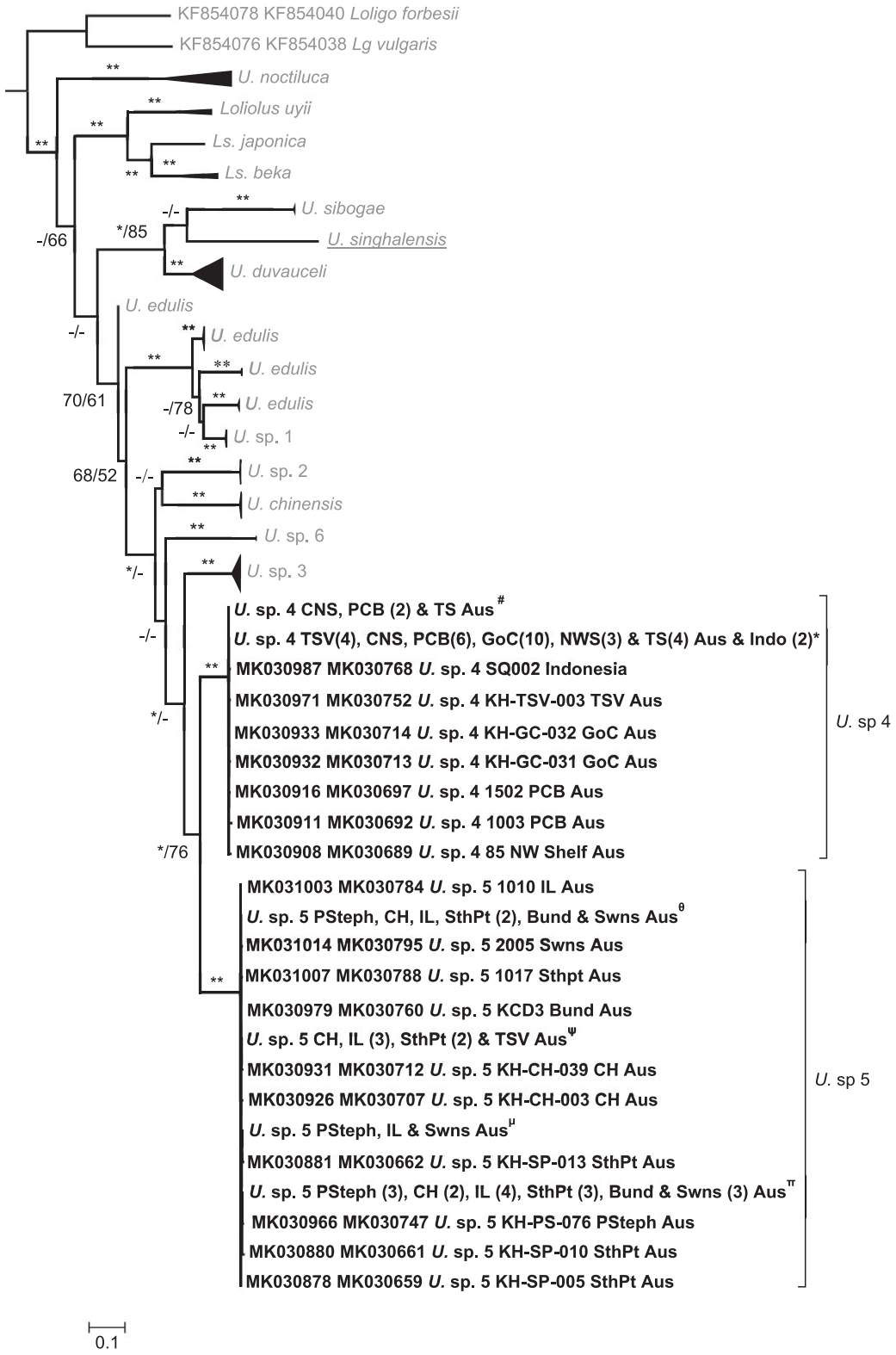

Figure 6. ML COI + 16S rRNA phylogeny of *Uroteuthis* expanding *U.* sp. 4 and *U.* sp. 5 clades from Figure 1. See Figure 1 caption for more detail.

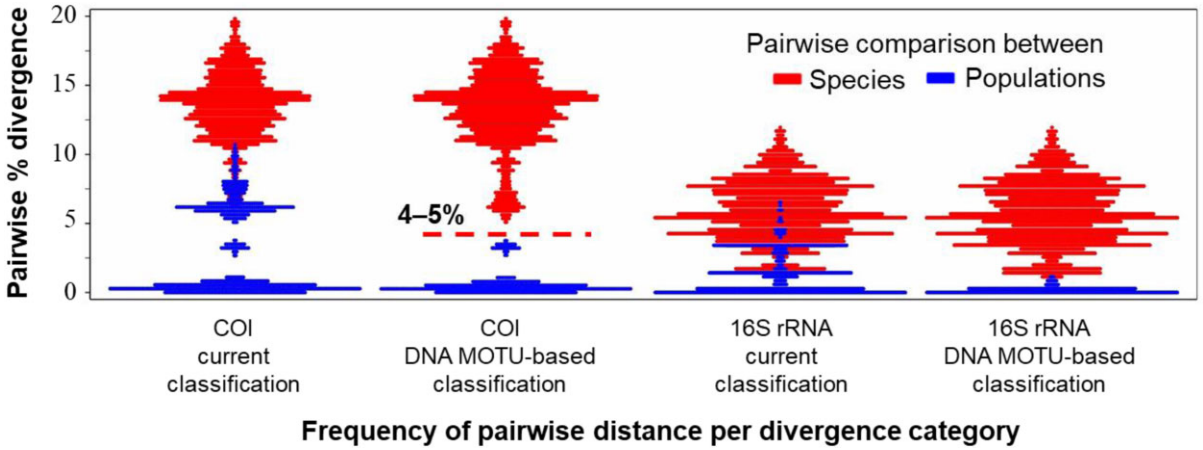

Figure 7. Summary of pairwise genetic distances between populations and species for mitochondrial COI and 16S rRNA highlighting the overlap between population and species comparisons caused by species complexes *vs* a DNA-based species barcode gap to define molecular operational taxonomic units.

Uroteuthis chinensis *and* U. *sp. 2 through* U. *sp. 6*

Clade *U.* sp. 6 has been differentiated based on two Indonesian samples (Fig. 4). This clade groups weakly with *U.* sp. 2 from northern Australia and more strongly in a polytomy that also contains *U. chinensis* and *U.* sp. 3 to *U.* sp. 5 (Fig. 4). Geographic structure was detected in the inshore species, *U.* sp. 3 from the COI gene but 16S rRNA did not resolve any structure (Fig. 5). The southern *U.* sp. 3 specimens were only found in waters off New South Wales from Wallis Lake to Port Jackson, and northern *U.* sp. 3 specimens were more widespread extending from Indonesia through the Gulf of Carpentaria and south to Moreton Bay in southern Queensland (Fig. 8). Sister to inshore *U.* sp. 3 were two deeper shelf clades, *U.* sp. 4 and *U.* sp. 5, the latter differentiated for the first time in the current study (Fig. 6). These two clades were also geographically separated from each other (and in both COI and 16S rRNA) with a small region of overlap identified off Townsville in northern Queensland (Fig. 8). The extent of the ranges of *U.* sp. 1, *U.* sp. 3 and *U.* sp. 4 in Western Australia is likely underestimated because of limited geographic sampling in that region (Fig. 8).

Species diagnostic mutations

Diagnostic COI single-nucleotide polymorphisms (SNPs) that differentiate among 23 of the 24 putative *Uroteuthis* and *Loliolus* species (*U. edulis* clade 3 was excluded as no COI sequence is available for comparison) are provided in Supplementary Material Table S2. Between 1 and 6 diagnostic SNPs were identified in 21 of the putative species with only *U. duvaucelii* clade 2 and *U.* sp. 5 specimens requiring multilocus sequence typing for identification.

DISCUSSION

This DNA-based analysis has provided a framework for improving the taxonomy and systematics of the genus *Uroteuthis*. The study supports previous research suggesting that many of the currently accepted species, including *U. duvaucelii*, *U. noctiluca*, *Ls. beka*, *Ls. uyii* and *U. edulis*, are species complexes (Yeatman & Benzie, 1994; Anderson, 2000a, b; Sin *et al.*, 2009; Dai *et al.*, 2012; Sales *et al.*, 2013; Anderson *et al.*, 2014; Jin *et al.*, 2022; Krishnan *et al.*, 2022). The study also confirmed that five undescribed Australian *Uroteuthis* species group with, but are genetically distinct from, *U. edulis* and *U. chinensis*. By clarifying the identity of the *Uroteuthis* squids caught in Australia's northern and east coast fisheries, this study provides valuable information to inform stock assessments and help improve resource management.

Nuclear DNA markers 18S rDNArDNA and rhodopsin were found to be too conserved for resolving closely related *Uroteuthis* species. Although the mitochondrial DNA COI and 16S rRNA sequences were unique for every species, COI provided the greatest resolution as a species-level marker for *Uroteuthis*, with interspecific divergences of up to 12%, compared with 7% for 16S rRNA. We propose a strict 4–5% COI barcode gap (maximum population divergence of 4% and minimum species divergence of 5%) and an average COI barcode gap of 0.65–12% (mean population difference to mean species difference) for the genus *Uroteuthis*. The 0.65% average level of population divergence for *Uroteuthis* is higher than that reported for other coleoid genera—Mastigoteuthidae 0.2% (Braid, McBride & Bolstad, 2014), Gonatidae 0.5% (Katugin *et al.*, 2017) and Onychoteuthidae 0.4% (Lischka, Braid & Bolstad, 2018)—but likely reflects the widespread geographic sampling in this study. Similar COI divergence levels to those reported here were observed by Sales *et al.* (2013) among cryptic species of *Doryteuthis pleii* (COI divergence 7.7%) and *D. pealeii* (COI divergence 4.9%). Our results provide further support that *Uroteuthis* is a paraphyletic genus, with *Loliolus* nested within, and taxonomic revision of both genera is warranted.

Using a COI barcode gap of 4–5%, the species complex of *U. duvaucelii* likely contains four species. The phylogenetic position of *U. noctiluca* and the node grouping *U.* sp. 2 with *U.* sp. 6 have poor branch support and should be treated as trichotomies until genetic markers with higher resolving power are sequenced. The species complexes of *U. noctiluca*, *Ls. beka* and *Ls. uyii* likely comprise at least two species each, and *U. edulis* likely comprises at least three, possibly four species, in addition to *U.* sp. 1. Although these species were not the primary focus of this study, their inclusion was necessary to provide the baseline intra- and interspecific divergence levels from which to clarify Australian species.

In northern and eastern Australian waters, this study differentiated five genetically distinct *Uroteuthis* species. A sixth new species revealed in this study, *U.* sp. 6, was only collected from Indonesia. At the time of writing this manuscript, there were a further 10 species of *Uroteuthis* with no DNA sequences available to compare. Since then, two have been sequenced, *U. singhalensis and U. bengalensis* (Jothinayagam, 1987) (Krishnan *et al.*, 2022). Neither of these species match *U.* sp. 6. The remaining eight species, *U. bartschi* Rehder, 1945, *U. vossi* (Nesis, 1982), *U. pickfordi* (Adam, 1954), *U. reesi* (Voss, 1962), *U. abulati* (Adam, 1955), *U. arabica* (Ehrenberg, 1831), *U. machelae* (Roeleveld & Augustyn, 2005) and

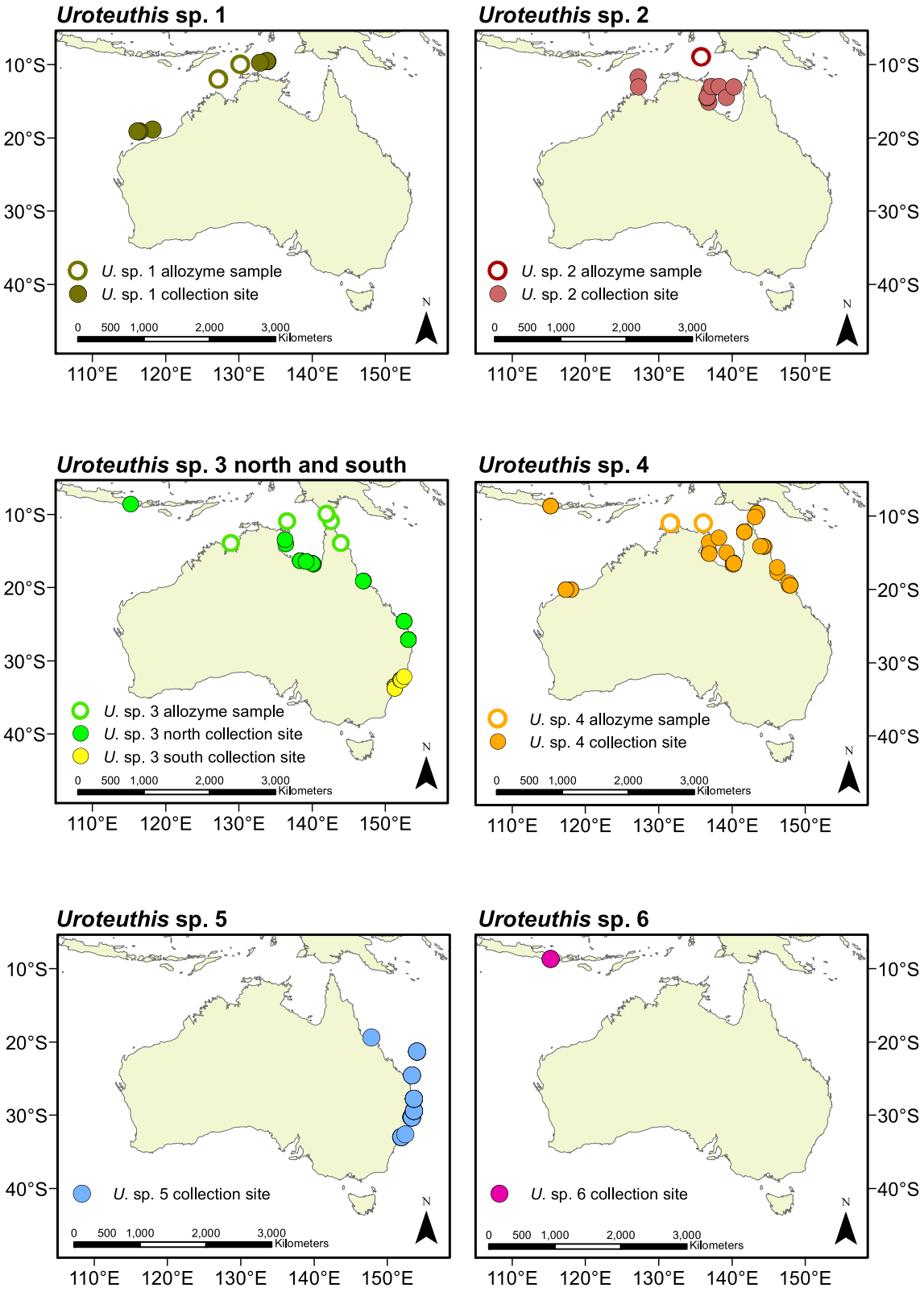

Figure 8. Schematic maps showing the sample collection sites and known species distributions of the newly differentiated *Uroteuthis* species. Sample collection sites used in this study are shown as filled circles; locations of allozyme samples analysed in earlier studies are shown as unfilled circles. Note the species are found on continental shelves and are unlikely to occur in deep oceanic waters (i.e. across the Timor Sea and Indonesian archipelago, and out into the Coral Sea) and may also occur outside these boundaries at sites as yet unsampled. Thus, the distribution zones are approximate, particularly in the west and southwest of Australia where sampling effort was low.

U. robsoni (Alexeyev, 1992) do not have morphologies akin to the medium to large body size of *U.* sp. 6, they have not been mistaken for *U. edulis* or *U. chinensis* in the literature, and they have not been reported in or close to Indonesian or Australian waters.

Of the Australian *Uroteuthis* species, *U.* sp. 1 falls within the *U. edulis* species complex and the remaining four species (*U.* sp. 2 to *U.* sp. 5) group with *U. chinensis*. These results support the allozyme results of Yeatman and coauthors (Yeatman, 1993; Yeatman & Benzie, 1993, 1994); however, an additional new species, *U.* sp. 5, was identified in this study. Like *U.* sp. 4, it is a deeper shelf species, but has a more southern distribution along the east coast of Australia. Sampling for the earlier studies focused on northern Australian waters, which may explain why they failed to detect this fifth species.

Distinct, non-overlapping geographic ranges may explain the genetic diversity of the *U.* sp. 1 and *U. edulis* species complex. Samples of *U.* sp. 1 have only been taken from northern Australia (the North West Shelf off Western Australia and the Arafura Sea between Australia and Western New Guinea) while samples in *U. edulis* clade 1 originate from Iran and India and those in *U. edulis* clade 2 (likely the type species) are predominantly from East Asia. Samples of both *U. edulis* clade 3 and clade 4 were collected from China, but more research is needed to obtain complete sequence data (COI and 16S rRNA) along with morphological data before the validity of these two species can be confirmed. A recent study of COI sequences from 131 *U. edulis* sourced from Chinese waters by Jin *et al.* (2022) supports the presence of *U. edulis* clades 2 and 4, suggesting that variant 3 may be an artefact.

The geographic range of *U.* sp. 2 includes northern Australia but also extends north to Indonesia. Yeatman & Benzie (1994) observed differences in depth preference between *U.* sp. 1, restricted to deeper water (>100 m), and *U.* sp. 2 collected from more intermediate depths (20–120 m).

Fisheries harvesting squids along the east coast of Australia are primarily catching three *Uroteuthis* species: a widely distributed inshore species, *U.* sp. 3, which is subdivided into northern (Queensland, north of Moreton Bay) and southern (New South Wales, south of Wallis Lakes) stocks; and two geographically restricted deeper shelf species, *U.* sp. 4, occurring from Townsville and further north in Queensland and *U.* sp. 5, occurring from Townsville and further south into New South Wales (Fig. 8). The three species, along with *U.* sp. 6, are closely related to *U. chinensis*. The northern and southern stocks of *U.* sp. 3 are approaching the borderline of speciation (4% COI divergence) while *U.* sp. 4 and *U.* sp. 5 have fully transitioned (6% COI divergence). Two factors influenced the decision not to divide *U.* sp. 3 into two species: one, the 16S rRNA marker was unable to differentiate the two clades, and two, the inshore estuary habitat preference of *U.* sp. 3 suggests that this species may display an isolation-by-distance genetic signal. In contrast, the geographic overlap of the deeper shelf species *U.* sp. 4 and *U.* sp. 5 in waters off Townsville suggests that they have been reproductively isolated, although using mitochondrial DNA markers alone we cannot exclude the possibility of hybrids where the two species overlap. Therefore, fine-scale sampling and further genetic analysis with more variable nuclear markers would be beneficial for all three species.

A morphological study of the three east coast species (*U.* sp. 3, 4 and 5) is ongoing, but the three species are proving difficult to differentiate due to overlapping morphometry and considerable intraspecific and gender-based variation (K. Hall *et al.*, unpubl.). While combined depth and location information can help separate many specimens, those collected from intermediate depths (i.e. 10–30 m) or near Townsville currently require genetic analysis. It is also challenging to discern which of the three species might be *U. etheridgei*, given the imprecise type locality, which was provided as 'Australian Seas (? S.E.)' in Berry (1918) to indicate an unknown capture location off southeastern Australia, and the small (possibly immature) holotype and paratype specimens.

Future research on the genus *Uroteuthis* should focus on sequencing type specimens, or where this is not possible, at least high-quality voucher specimens originating from the type locality with reliable source information, so that DNA sequences can be accurately assigned to described species. More variable nuclear DNA markers are also needed for the genus and will likely be acquired with high-throughput sequencing. Once species delimitation is resolved, baseline information relating to species morphology and ranges can be verified, noting that the origin of capture of market-sourced specimens may be uncertain.

CONCLUSION

Applying a 4–5% mitochondrial DNA COI barcode gap has partitioned *U. duvaucelii*, *U. noctiluca* and *U. edulis* into 10 species. This study also provides the first mitochondrial DNA sequences for four *Uroteuthis* species (*U.* sp. 1, *U.* sp. 4, *U.* sp. 5 and *U.* sp. 6—the latter two we propose are new species in the genus). Additional sequencing from specimens of *U.* sp. 2 and *U.* sp. 3 aligns results from previous allozyme studies with published sequences, consolidating sample identities and expanding the species' known distributions. One of the new species, *U.* sp. 6, has only been collected from Indonesia. Five species, *U.* sp. 1 through *U.* sp. 5, were studied in greater detail. They occur in Australian waters with two, *U.* sp. 1 and *U.* sp. 2, predominantly found off northern Australia. Three species dominate catches off the east coast; *U.* sp. 3 is a widely distributed inshore species and *U.* sp. 4 and *U.* sp. 5 are deeper shelf species with narrower distributions that overlap off Townsville in Queensland. The new information provided by this study will be used to better inform fisheries managers for the sustainable ongoing harvest of *Uroteuthis* species from Australian waters.

ACKNOWLEDGEMENTS

Sincere thanks are extended to Professor C. C. Lu, Chuan-Wen Ho, Gary Fry, Matthew McHugh, Matthew Harrison, Mitch Burns, Daniel Johnson, Lounds Seafoods and the Coffs Harbour, Iluka, and Clarence River Fishermen's Co-operatives for assistance in sourcing squid samples. We also thank Adnan Maussalli for specimens and early contributions to genetic screening. Thanks also to Kerri Chandra and Katherine Delaney for assistance with creating the graphs and maps. Valuable feedback was also received during journal peer review.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of Molluscan Studies* online.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

FUNDING

This work was funded by the Hermon Slade Foundation (Project ID HSF 15-09 Hall), the New South Wales Department of Primary Industries and the Queensland Department of Agriculture and Fisheries.

DATA AVAILABILITY

Sequences generated within this project are publicly available online through GenBank (National Centre for Biotechnology

Information) (accession numbers are listed in Supplementary Material Table S1). Sequence alignments are publicly available on Dryad (doi.org/10.5061/dryad.gf1vhhmvr).

AUTHORS' CONTRIBUTIONS

Samples sourced by M.C.D., J.D. (née Yeatman) and K.CH. Genetic screening and marker selection by D.B., R.S. and J.A.T.M. Funding sourced by K.C.H. Genetic analysis and manuscript prepared by J.A.T.M. with input from M.C.D., K.C.H., J.D. and J.R.O.

REFERENCES

ADAM, W. 1954. Cephalopoda: IIIe partie. IV. Céphalopodes à l'exclusion des genres *Sepia. Sepiella et Sepioteuthis Siboga Expeditie*, **55c**: 123–193, 140 figs, 124 pls.

ADAM, W. 1955. Cephalopodes. Résultats scientifiques des campagnes de la Calypso, I. Campagnes 1951–1952 en Mer Rouge. *Annales de l'Institut Océanographique, Paris*, **30**: 185–194.

ALEXEYEV, D.O. 1992. The systematic position of bioluminescent squids of family Loliiginidae (Cephalopoda, Myopsida). *Zoologicheskiy Zhurnal*, **71**: 12–23. [in Russian]

ALLCOCK, A.L., LINDGREN, A. & STRUGNELL, J. 2015. The contribution of molecular data to our understanding of cephalopod evolution and systematics: a review. *Journal of Natural History*, **49**: 1373–1421.

ANDERSON, F.E. 2000a. Phylogeny and historical biogeography of the loliginid squids (Mollusca: Cephalopoda) based on mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution*, **15**: 191–214.

ANDERSON, F.E. 2000b. Phylogenetic relationships among loliginid squids (Cephalopoda: Myopsida) based on analyses of multiple data sets. *Zoological Journal of the Linnean Society*, **130**: 603–633.

ANDERSON, F.E., BERGMAN, A., CHENG, S.H., PANKEY, M.S. & VALINASSAB, T. 2014. Lights out: the evolution of bacterial bioluminescence in Loliginidae. *Hydrobiologia*, **725**: 189–203.

ARKHIPKIN, A.I., RODHOUSE, P.G.K., PIERCE, G.J., SAUER, W., SAKAI, M., ALLCOCK, L., ARGUELLES, J., BOWER, J.R., CASTILLO, G., CERIOLA, L., CHEN, C.-S., CHEN, X., DIAZ-SANTANA, M., DOWNEY, N., GONZÁLEZ, A., AMORES, J.G., GREEN, C.P., GUERRA, A., HENDRICKSON, L.C., IBÁÑEZ, C., ITO, K., JEREB, P., KATO, Y., KATUGIN, O.N., KAWANO, M., KIDOKORO, H., KULIK, V., LAPTIKHOVSKY, V.V., LIPINSKI, M., LIU, B., MARIÁTEGUI, L., MARIN, W., MEDINA, A., MIKI, K., MIYAHARA, K., MOLTSCHANIWSKYJ, N., MOUSTAHFID, H., NABHITABHATA, J., NANJO, N., NIGMATULLIN, C.M., OHTANI, T., PECL, G., PEREZ, J., PIATKOWSKI, U., SAIKLIANG, P., SALINAS-ZAVALA, C.A., STEER, M., TIAN, Y., UETA, Y., VIJAI, D., WAKABAYASHI, T., YAMAGUCHI, T., YAMASHIRO, C., YAMASHITA, N. & ZEIDBERG, L.D. 2015. World squid fisheries. *Reviews in Fisheries Science & Aquaculture*, **23**: 92–252.

BARTON, D.P. & MORGAN, J.A.T. 2016. A morphological and genetic description of pentastomid infective nymphs belonging to the family Sebekidae Sambon, 1922 in fish in Australian waters. *Folia Parasitologica*, **63**: 1.

BERRY, S.S. 1918. Report on the Cephalopoda obtained by the F.I.S. "Endeavour" in the Great Australian Bight and other Southern Australian localities. In: *Biological results of the fishing experiments carried on by the F.I.S. "Endeavour"*. Vol. 4 (H.C. Dannevig, ed.), pp. 203–298. Department of Trade and Customs, Sydney.

BRAID, H.E., MCBRIDE, P.D. & BOLSTAD, K.S.R. 2014. Molecular phylogenetic analysis of the squid family Mastigoteuthidae (Mollusca, Cephalopoda) based on three mitochondrial genes. *Hydrobiologia*, **725**: 145–164.

CRACRAFT, J. & DONOGHUE, M.J. 2004 *Assembling the tree of life*. Oxford University Press, New York.

DAI, L., ZHENG, X., KONG, L. & LI, Q. 2012. DNA barcoding analysis of Coleoidea (Mollusca: Cephalopoda) from Chinese waters. *Molecular Ecology Resources*, **12**: 437–447.

DARRIBA, D., POSADA, D., KOZLOV, A.M., STAMATAKIS, A., MOREL, B. & FLOURI, T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution*, **37**: 291–294.

D'ORBIGNY, A. 1835–1848. *Histoire Naturelle Générale et Particuliere des Céphalopodes Acétabulifères Vivants et Fossiles* (Férussac, A. and d'Orbigny, A., eds). pp. 361. J.B. Bailliere, Paris.

DUNNING, M. 1998. Loliginidae—inshore squids, pencil squids. In: *FAO species identification guide for fishery purposes. the living marine resources of the Western Central Pacific*. Vol. 2 (K.E. Carpenter & V.H. Niem, eds), pp. 764–780. FAO, Rome.

DUNNING, M., MCKINNON, S. & YEOMANS, K. 2000. *Development of a northern Australian squid fishery. FRDC Final Report Project No. 94/017*. Queensland Department of Primary Industries Queensland, Australia.

DUNNING, M., NORMAN, M. & REID, A. 1998. Cephalopods: introduction and general remarks. In: *FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific*. Vol. 2 (K.E. Carpenter & V.H. Niem, eds), pp. 687–826. FAO, Rome.

EHRENBERG, C.G. 1831. *Symbolae physicae, seu, icones et descriptiones corporum naturalium novorum aut minus cognitorum: quae ex itineribus per Libyam Aegyptum Nubiam Dongalam Syriam Arabiam et habessiniam publico institutis sumptu friderici Guilelmi Hemprich et christiani Godofredi Ehrenberg: studio annis 1820-25 redierunt*. Berolini?, Berlin, Pars Zoologica. v.4. Animalis Evertebrata exclusis Insectis.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution: International Journal of Organic Evolution*, **39**: 783–791.

FLOURI, T., IZQUIERDO-CARRASCO, F., DARRIBA, D., ABERER, A.J., NGUYEN, L.-T., MINH, B.Q., VON HAESELER, A. & STAMATAKIS, A. 2015. The phylogenetic likelihood library. *Systematic Biology*, **64**: 356–362.

GELLER, J., MEYER, C., PARKER, M. & HAWK, H. 2013. Redesign of PCR primers for mitochondrial cytochrome *c* oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, **13**: 851–861.

GRAY, J.E. 1849. *Catalogue of the mollusca in the collection of the British Museum: I: Cephalopoda Antepedia*. British Museum Natural History, London.

HASEGAWA, M., KISHINO, H. & YANO, T.-A. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**: 160–174.

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

HOYLE, W.E. 1885. XIX. Diagnoses of new species of Cephalopoda collected during the cruise of HMS 'Challenger.' Part I. The Octopoda. *Annals and Magazine of Natural History*, **15**: 222–236.

HOYLE, W.E. 1886. Report on the Cephalopoda collected by H.M.S. "Challenger" during the years 1873–76. In: *Report on the scientific results of the voyage of H.M.S. "Challenger" during the Years 1873–76, Zoology*. Vol. 16 (C.W. Thompson & J. Murray, eds). Eyre & Spottiswoode, London.

HUELSENBECK, J.P. & RONQUIST, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754–755.

JEREB, P. & ROPER, C. 2010. *Cephalopods of the world. An annotated and illustrated catalogue of species known to date. Volume 2. Myopsid and Oegopsid Squids. FAO species catalogue for fishery purposes*. Vol. 4. FAO, Rome.

JIANG, L., KANG, L., WU, C., CHEN, M. & LÜ, Z. 2018. A comprehensive description and evolutionary analysis of 9 Loliginidae mitochondrial genomes. *Hydrobiologia*, **808**: 115–124.

JIN, Y., WANG, C., LI, N. & FANG, Z. 2022. Species identification of two Loliginidae cryptic species in China Seas with morphological and molecular methods. *Regional Studies in Marine Science*, **55**: 102549.

JOTHINAYAGAM, J. 1987. Cephalopoda of the Madras Coast. Zoological Survey of India. *Technical Monograph*, **15**: 1–85.

KATUGIN, O.N., CHICHVARKHINA, O.V., ZOLOTOVA, A.O. & CHICHVARKHIN, A.Y. 2017. DNA barcoding for squids of the family Gonatidae (Cephalopoda: Teuthida) from the boreal North Pacific. *Mitochondrial DNA Part A*, **28**: 41–49.

KIMURA, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences*, **78**: 454–458.

KRISHNAN, N., SUKUMARAN, S., SEBASTIAN, W. & GOPALAKRISHNAN, A. 2022. Morphological and molecular investigations on squids of the genera *Uroteuthis* and *Loliolus* from the Indian coast to resolve taxonomic ambiguities. *Malacologia*, **64**: 215–230.

KUMAR, S., STECHER, G. & TAMURA, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**: 1870–1874.

LAMARCK, J.-B. 1798. Extrait d'un mémoire sur le genre de la séche, du calmar et poulpe, vulgairement nommés, polypes de mer. *Bulletin des Sciences, par la Société Philomatique de Paris*, **2**: 129–131.

LARKIN, M.A., BLACKSHIELDS, G., BROWN, N.P., CHENNA, R., MCGETTIGAN, P.A., MCWILLIAM, H., VALENTIN, F., WALLACE, I.M., WILM, A. & LOPEZ, R. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**: 2947–2948.

LISCHKA, A., BRAID, H.E. & BOLSTAD, K.S.R. 2018. DNA barcoding from the hooked squids (Cephalopods: Onychoteuthidae) of the Sargasso Sea. *Hydrobiologia*, **808**: 107–113.

LU, C. 2001. Cephalopoda. In: *Mollusca: Aplacophora, polyplacophora, scaphopoda, cephalopoda, Vol. 17.2: Zoological catalogue of Australia* (A. Wells & W.K. Houston, eds), pp. 129–308. CSIRO Publishing, Melbourne.

LU, C., ROPER, C.F. & TAIT, R. 1985. A revision of *Loliolus* (Cephalopoda; Loliginidae), including *L. noctiluca*, a new species of squid from Australian waters. *Proceedings of the Royal Society of Victoria*, **97**: 59–85.

MEIER, R., ZHANG, G. & ALI, F. 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. *Systematic Biology*, **57**: 809–813.

MILLER, S., DYKES, D. & POLESKY, H. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, **16**: 1215.

NATSUKARI, Y. 1984. Taxonomical and morphological studies on the Loliginid Squids IV: two new genera of the Family Loliginidae. *Venus (Japanese Journal of Malacology)*, **43**: 229–239.

NATSUKARI, Y. & OKUTANI, T. 1975. Taxonomic and morphological studies on the loliginid squids I: identity of *Loligo chinensis* Gray, 1849, redescription of the type specimen and taxonomic review (Cephalopoda: Loliginidae). *Venus (Japanese Journal of Malacology)*, **34**: 85–91.

NESIS, K.N. 1982. *Abridged key to the cephalopod molluscs of the world's ocean*. Light and Food Industry Publishing House, Moscow. [in Russian]

OKUTANI, T. 2005. Past, present and future studies on cephalopod diversity in tropical West Pacific. *Phuket Marine Biological Center Research Bulletin*, **66**: 39–50.

ORTMANN, A. 1891. Cephalopoden von Ceylon, gesammelt von den Herren Dres. Sarasin. *Zoologische Jahrbucher*, **5**: 669–678.

PALUMBI, S.R. 1996. Macrospatial genetic structure and speciation in marine taxa with high dispersal abilities. In: *Molecular zoology: advances, strategies and protocols* (J.D. Ferraris & S.R. Palumbi, eds), pp. 101–117. Wiley-Liss, Inc, New York.

PUILLANDRE, N., LAMBERT, A., BROUILLET, S. & ACHAZ, G. 2012. ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, **21**: 1864–1877.

QUOY, J. & GAIMARD, J. 1832. Voyage de decouvertes de l'Astrolabe pendant les annees 1826–1829 sous le commendement de MJ Dumont d'Urville. *Zoologie*, **2**: 1–320.

RAUPACH, M.J., MAYER, C., MALYUTINA, M. & WÄGELE, J.-W. 2009. Multiple origins of deep-sea *asellota* (Crustacea: Isopoda) from shallow waters revealed by molecular data. *Proceedings of the Royal Society B: Biological Sciences*, **276**: 799–808.

REHDER, H. 1945. A new genus and species of squids from the Philippines. *Proceedings of the Biological Society of Washington*, **58**: 21–26.

RODHOUSE, P.G. 2005. World squid resources. In: *Review of the state of world marine fisheries*. Vol. FAO Fisheries Technical Paper 457 (J. Csirke, ed.), pp. 175–187. Marine Resources Service, Fishery Resources Division, Food and Agriculture Organisation of the United Nations, Rome.

ROELEVELD, M.A. & AUGUSTYN, C. 2005. Description of a new species of *Uroteuthis* (*Photololigo*) from the Mozambique Channel. *Phuket Marine Biological Center Research Bulletin*, **66**: 97–107.

SALES, J.B., SHAW, P.W., HAIMOVICI, M., MARKAIDA, U., CUNHA, D.B., READY, J., FIGUEIREDO-READY, W.M., SCHNEIDER, H. & SAMPAIO, I. 2013. New molecular phylogeny of the squids of the family Loliginidae with emphasis on the genus *Doryteuthis* Naef, 1912: mitochondrial and nuclear sequences indicate the presence of cryptic species in the southern Atlantic Ocean. *Molecular Phylogenetics and Evolution*, **68**: 293–299.

SALES, J.B.L., MARKAIDA, U., SHAW, P.W., HAIMOVICI, M., READY, J.S., FIGUEREDO-READY, W.M., ANGIOLETTI, F., CARNEIRO, M.A., SCHNEIDER, H. & SAMPAIO, I. 2014. Molecular phylogeny of the genus *Lolliguncula* Steenstrup, 1881 based on nuclear and mitochondrial DNA sequences indicates genetic isolation of populations from North and South Atlantic, and the possible presence of further cryptic species. *PLoS ONE*, **9**: e88693.

SASAKI, M. 1929. A monograph of dibranchiate cephalopods of the Japanese and adjacent waters. *Journal College of Agriculture Hokkaido University*, **20**: 1–357, 330pls.

SIN, Y.W., YAU, C. & CHU, K.H. 2009. Morphological and genetic differentiation of two loliginid squids, *Uroteuthis* (*Photololigo*) *chinensis* and *Uroteuthis* (*Photololigo*) *edulis* (Cephalopoda: Loliginidae), in Asia. *Journal of Experimental Marine Biology and Ecology*, **369**: 22–30.

STEENSTRUP, J. 1856. *Original description not documented*. https://www.molluscabase.org/aphia.php?p=taxdetails&id=140270 (17 May 2023, date last accessed).

SWOFFORD, D.L. 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*. Sinauer Associates, Sunderland, MA.

TAKEMOTO, K. & YAMASHITA, M. 2012. Complete nucleotide sequences of mitochondrial DNA of long-finned squid *Loligo edulis*. *Fisheries Science*, **78**: 1031–1039.

TAMURA, K. & NEI, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**: 512–526.

TAVARÉ, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. In: *Lectures on mathematics in the life sciences*, vol. **17**, pp. 57–86. American Mathmatical Society, Providence, RI.

VECCHIONE, M., BRAKONIECKI, T.F., NATSUKARI, Y. & HANLON, R.T. 1998. A provisional generic classification of the family Loliginidae. *Smithsonian Contributions to Zoology*, **586**: 215–222.

VECCHIONE, M., SHEA, E., BUSSARAWIT, S., ANDERSON, F., ALEXEYEV, D., LU, C., OKUTANI, T., ROELEVELD, M., CHOTIYAPUTTA, C. & ROPER, C. 2005. Systematics of Indo-West Pacific loliginids. *Phuket Marine Biological Center Research Bulletin*, **66**: 23–26.

VOSS, G.L. 1962. Six new species and two new subspecies of cephalopods from the Philippine Islands. *Proceedings of Biological Society of Washington*, **75**: 169–176.

WAKIYA, Y. & ISHIKAWA, M. 1921. Some Japanese myopsid cephalopods, with descriptions of four new species. *Zoological Magazine, Tokyo*, **33**: 279–292, pls. 1–2. [in Japanese]

WIENS, J.J. 2006. Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics*, **39**: 34–42.

YEATMAN, J. 1993. *Morphological aspects of Australian Photololigo spp. (Loliginidae: Cephalopoda)*. PhD thesis, James Cook University of North Queensland, Townsville.

YEATMAN, J. & BENZIE, J.A.H. 1993. Cryptic speciation in *Loligo* from northern Australia. In: *Recent advances in fishery biology* (T. Okutani, R.K. O'Dor & T. Kubodera, eds), pp. 641–652. Tokai University Press, Tokyo.

YEATMAN, J. & BENZIE, J.A.H. 1994. Genetic structure and distribution of *Photololigo* spp. in Australia. *Marine Biology*, **118**: 79–87.

ZHANG, J., KAPLI, P., PAVLIDIS, P. & STAMATAKIS, A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**: 2869–2876.