Research Article

Mediterranean Marine Science
Indexed in WoS (Web of Science, ISI Thomson) and SCOPUS
The journal is available on line at http://www.medit-mar-sc.net
www.hcmr.gr
DOI: http://doi.org/10.12681/mms.27733

Potential crypticity within two decapod (Crustacea) genera: *Galathea* Fabricius, 1793 and *Eualus* Thallwitz, 1891 suggested by integrative taxonomic approach

Sheena CONFORTI[1,2] and Federica COSTANTINI[2]

[1] Eawag, Department Environmental Microbiology, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland
[2] Department of Biological, Geological and Environmental Science (BiGeA), University of Bologna, Ravenna Campus, Via Sant'Alberto 163, 48123 Ravenna, Italy

Corresponding author: Sheena CONFORTI; sheena.conforti@eawag.ch

Contributing Editor: Costas TSIGENOPOULOS

Received: 21 August 2021; Accepted: 11 December 2021; Published online: 20 June 2022

Abstract

Correct species identification and description are fundamental to understand the health status of marine ecosystems. The use of a single identification tool for species distinction may lead to species misidentifications, having major consequences for ecological studies. In the present research, we used an integrative taxonomic approach to identify benthic decapods belonging to the genera of *Galathea* Fabricius, 1793 and *Eualus* Thallwitz, 1891, collected in the Mediterranean Sea. 23 *Galathea* and 22 *Eualus* individuals were morphologically analyzed and sequenced at the mitochondrial COI gene to confirm their identity using the BOLD Identification Engine. Morphological identification revealed the presence of two *Galathea* and three *Eualus* species, while species delimitation based on DNA barcoding of COI sequences strongly suggested the presence of three *Galathea* and four *Eualus* species. Molecular analyses suggested the potential presence of two still undescribed species: one cryptic to *Galathea squamifera* and one cryptic to *Galathea intermedia*. Contrasting results obtained via morphological identification and the BOLD Identification Engine impeded the recognition of *Eualus* specimens and suggested misidentifications among BOLD reference records of *Eualus cranchii*, *Eualus occultus* and *Eualus pusiolus*. These results demonstrated that morphological identification overlooks cryptic species and that misidentifications may occur, highlighting the importance of using an integrative approach to increase the current taxonomic knowledge of benthic invertebrates.

Keywords: integrative taxonomy; barcoding; mitochondrial marker; Mediterranean Sea.

Introduction

In the last decades, marine biodiversity has been put under severe pressure by multiple anthropogenic drivers. For instance, over-exploitation of commercially valued species, habitat change, loss and degradation (Wilson *et al.*, 2008), eutrophication and hypoxic events (Vaquer-Sunyer & Duarte, 2008), introduction of non-indigenous species (Gollasch, 2006), rising sea surface temperature, increase in UV exposition and ocean acidification (Halpern *et al.*, 2008) are threatening marine ecosystems and diversity, triggering their decline. Recording the current state of biological diversity is thus of primary importance to better understand how species diversity loss can be induced by human impact (Bilgin *et al.*, 2015) since biodiversity has an important value as an indicator of environmental quality and ecosystem functioning (Bianchi & Morri, 2000).

However, the current number of taxa in the oceans is underestimated, and only a small fraction of marine species is known (Appeltans *et al.*, 2012). While the number of valid species in the oceans amounts to 239,610 (WoRMS consulted in October 2021), the total marine biodiversity is estimated to range between 300,000 and more than 10 million species (May, 1992; Poore & Wilson, 1993; Bouchet, 2006; Mora *et al.*, 2011; Costello *et al.*, 2012), and such a discrepancy is due to the different methods used to estimate it. Despite the incertitude of the estimate, the presence of a knowledge gap in ocean biodiversity is unquestionable and is a result of most taxonomic studies concentrating only on species-rich and charismatic taxa or limited to some regions of the world (Troudet *et al.*, 2017). The taxonomic revision of taxa using molecular techniques, the improved biodiversity prospecting due to improved sampling methods, and the discovery of novel habitats currently allow to describe approximately 1,500 new marine species each year (Goulletquer *et al.*, 2014). However, there is a need to increase the current rate of species description to avoid that many of them disappear before they are described (Appeltans *et al.*, 2012).

On one hand, species description based solely on morphological characters can lead to false species identification, which triggers multiple problems in ecological studies. For instance, species misidentification can over or underestimate species richness and consequently, species diversity measures (Lefébure *et al.*, 2006). Additionally, it can have implications for management and conservation plans (Bishop *et al.*, 2013; Bishop *et al.*, 2015), as well as for ecological niche modelling (Ensing *et al.*, 2013) and connectivity studies (Pante *et al.*, 2015). Errors in species identification stemming from morphological approaches are often the result of a "taxonomic impediment", which refers to the lack of dissemination of taxonomic information and of expertise in many groups of living organisms (Tautz *et al.*, 2003). In fact, the use of morphological traits for species identification has several limitations. Firstly, the morphological variability associated with phenotypic plasticity (Windig *et al.*, 2004) and dimorphism (Pérez-Barros *et al.*, 2008) in the characters employed for species recognition may give rise to uncertainty. Secondly, this partial approach overlooks morphologically similar taxa, i.e., cryptic species (Hebert *et al.*, 2003; Caputi *et al.*, 2007; Chenuil *et al.*, 2019). Thirdly, morphological keys are sometimes only effective for a particular life stage or gender, and their use is labor-intensive and requires a high level of expertise (Lefébure *et al.*, 2006; Valentini *et al.*, 2009).

On the other hand, in the last decades DNA barcoding has been used as a global standard for the identification of biological species (Hebert *et al.*, 2003). Species assignment using COI DNA barcoding is facilitated by using public libraries such as BOLD (Barcode of Life Data System) (Ratnasingham & Hebert, 2007) and GenBank (Clark *et al.*, 2015), in which sequenced DNA barcodes from unknown specimens can be compared against reference records to identify the matching species. Nevertheless, the concept of DNA barcoding as the only taxonomic tool to identify species has been subjected to criticism. Firstly, DNA barcodes are more likely to provide potentially useful information for groups that are already well studied (Rubinoff, 2006), and for which a polished species-level taxonomy exists (Petinsaari *et al.,* 2020). Secondly, the sequence length of DNA barcodes is very short compared to the entire genome and consequently it may offer only a fraction of the information needed to characterize species (Rubinoff, 2006). Finally, the success of DNA barcoding depends on the fact that interspecific variation exceeds intraspecific variation by at least one order of magnitude, establishing a "barcoding gap" (Hebert *et al.*, 2003). However, the presence of this gap is often the consequence of insufficient sampling across taxa (Wiemers & Fiedler, 2007).

To overcome the problem of using only one taxonomic approach to characterize species and to improve the accuracy of species identification, multiple studies have combined molecular and traditional, morphology-based taxonomy (Scorrano *et al.*, 2016; Bayha *et al.*, 2017).

In this study, we adopted an "integrative taxonomy" approach (Dayrat, 2005) on benthic decapods belonging to the genera *Galathea* Fabricius, 1793 and *Eualus* Thallwitz, 1891. Besides morphological identification, a region of the mitochondrial COI gene was sequenced for molecular identification. The two genera were chosen as study groups due to their ecological and economic importance, but also as examples of how data contained in public databases should be reviewed, controlled, and updated.

Decapod crustaceans are a highly diverse group of benthic invertebrates of the Mediterranean continental shelf and slope, and many of them have a high commercial value (Silva *et al.*, 2011; Colloca *et al.*, 2003). Moreover, decapod larvae represent almost 90% of the total zooplankton biomass, playing a fundamental role as prey for carnivorous species and as consumers of smaller phytoplankton, thus determining the stability of benthic and fisheries biology (Anger, 2001). Although studies on the decapod population's biology and ecology have increased during the last decades (Company *et al.*, 2008; Guijarro *et al.*, 2009), only 20% (6,269 of the 32,282 species described in WoRMS) are represented by COI barcode region sequences in BOLD (WoRMS and BOLD consulted in October 2021 (Horton *et al.*, 2020; Ratnasingham & Hebert, 2007)). These data highlight the importance of increasing the application of COI DNA barcoding to improve the analysis of decapod diversity patterns in the Mediterranean Sea and throughout the world. Furthermore, crypticity among decapod crustaceans has long been studied (Knowlton, 1986; Machordom & Macpherson, 2004), suggesting the need to adopt an integrated approach to identify species in a more accurate and conclusive manner.

Galathea squat lobsters are one of the most speciose and unwieldy groups in the Galatheidae family (Macpherson & Robainas-Barcia, 2015). In a previous study, they described the existence of 92 new *Galathea* species using morphological and molecular characters, highlighting that most of the new species were distinguished by subtle but constant morphological differences, which was consistent with molecular divergences of the mitochondrial marker COI (Macpherson & Robainas-Barcia, 2015). This suggests the utility of molecular studies, including specimens from different localities, to help clarify relationships among *Galathea* species. The same applies to the *Eualus* species, whose specimens are morphologically very similar, which makes their taxonomic distinction extremely challenging (D'Udekem-d'Acoz & Wirtz, 2002). DNA barcoding of COI has proved to be an informative molecular marker to describe new species of *Eualus* (Nye *et al.*, 2013), as well as to reveal the presence of two *Eualus cranchii* phylogeographic clades/haplogroups, suggesting the presence of cryptic species and/or genetic isolation of populations within species (Bilgin *et al.*, 2015).

Here, we aimed to *i)* correctly identify individuals of *Galathea* and *Eualus* genera up to the species level by means of morphological traits and DNA barcoding of the COI gene, *ii)* compare results obtained from morphological identification and molecular identification using different databases of BOLD IDEngine, *iii)* reveal the potential presence of crypticity within both genera using the Automatic Barcode Gap Discovery (ABGD) software

as a procedure for species delimitation (Puillandre *et al.*, 2012), and *iv)* solve phylogenetic relationships within the two genera using newly sequenced data and records of closely-related species retrieved from GenBank and BOLD public databases.

Overall, the results obtained by this study will increase the current taxonomic knowledge on marine benthic decapods inhabiting the Mediterranean Sea, which is the foundation stone for further ecological studies and biodiversity monitoring.

Materials and Methods

Field work

Specimens were collected within the framework of the co-founded ERA-NET MarTERA SEAMoBB (Solutions for Semi-Automated Monitoring of Benthic Biodiversity) project. Sampling was conducted between June and July 2019 in three localities in the Mediterranean Sea: Livorno (Italy), Palinuro (Italy) and Rovinj (Croatia) (Table 1, Fig. S.1). In each locality, 9 Autonomous Reef Monitoring Structures (ARMS) units (Leray & Knowlton, 2015) were deployed in March 2018 to allow a standardized sampling of the colonizing organisms after one year. Samples were collected by scuba divers at depths between 13 and 20 m. ARMS were removed from the benthic substrate by encapsulating the unit in a tin-PVC (polyvinyl chloride) box to prevent motile organisms from escaping. Once at the surface, ARMS were transported to shore and placed into a container filled with seawater where the tin-PVC box was removed, and the unit was systematically disassembled. Seawater from the disassembly container was filtered using 2 mm, 500 µm, and 100 µm sieves (Leray & Knowlton, 2015). Collected organisms from the 2 mm fractions were preserved in 95% ethanol and stored at 4°C for future studies. In total, 23 *Galathea* individuals and 22 *Eualus* individuals were collected and analyzed.

Morphological identification

Specimens were sorted under the stereomicroscope (Leica WILD M38®) for identification. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the software GIMP v2.10.6 (The GIMP Development Team, 2019) and Inkscape 0.92.3 (Harrington, 2005) (Fig. S.2, S.3, S.5).

To identify specimens of *Galathea*, the morphological keys provided by Poggiani (2018), and Zariquiey Álvarez (1968) were used. The main criteria of distinction for *Galathea* specimens were the following: total carapace length, shape of the rostrum, and presence/absence of spines on the epigastric region. The carapace length was measured from photographs using the software ImageJ (Rueden *et al.*, 2017). The difference between species was tested using a non-parametric Wilcoxon sum of rank test in R (R Core Team, 2018).

To identify specimens of the genus *Eualus*, the dichotomous key provided by D'Udekem-d'Acoz and Wirtz (2002) was used. The main criteria of distinction for *Eualus* specimens were the tip of the rostrum – either bifurcate or trifurcate – the presence/absence of the mandibular palp, the presence/absence of the pterygostomial tooth, and the presence/absence of spines on the merus of the 5th pereiopod.

Molecular analysis

The genomic DNA of *Galathea* individuals was isolated from their legs, while that of *Eualus* specimens from their posterior pereiopods. Genomic DNA extraction was performed using a cetyltrimethyl ammonium bromide (CTAB) protocol (Hillis & Moritz, 1990).

PCR amplification of the COI gene was performed using degenerate primers (Geller *et al.*, 2013). Each 19 µl PCR amplification contained 10 µl of Taq polymerase Master Mix (Phusion U Hot Start, ThermoFisher®), 0.6 µl of each primer (10 mM), 0.2 µl of 20 mg/ml BSA (Bovine Serum Albumin), 6.6 µl of nuclease-free H_2O (Ambion®), and 1 µl of DNA. Amplification was performed on SimpliAmp Thermal Cycler (Applied Biosystems®). Cycling parameters consisted of an initial denaturation step at 95°C for 15 min, followed by 4 cycles at 94°C for 30 s, at 50°C for 45 s, at 72°C for 1 min, further followed by 34 cycles at 94°C for 30 s, at 45°C for 45 s, at 72°C for 1 min, with a final extension step at 72°C for 8 min. Amplified fragments were purified using the ExoSAP-IT Express kit (Applied Biosystem®) following

Table 1. Summary of the sampling locations with the corresponding longitude and latitude, sampling depth (m), number of collected specimens in each location (N) and corresponding sample names.

Location	Depth [m]	Longitude [°W]	Latitude [°N]	Collected specimens of *Galathea* spp. [N]	Collected specimens of *Eualus* spp. [N]
Livorno (Italy)	13	10.34163	43.46233	4 (LIV1-4)	10 (LIV1-10)
Palinuro (Italy)	20	15.26857	40.03281	12 (PAL1-12)	4 (PAL1-4)
Rovinj (Croatia)	15	13.61834	45.09265	7 (CRO1-7)	8 (CRO1-8)

the manufacturer's recommendations, and commercially sequenced by Macrogen Europe (Amsterdam-NL).

Data analysis

The obtained sequences were trimmed and manually cleaned using the software BioEdit® (Hall, 1999), and later aligned in MEGA v.7 (Kumar *et al.*, 2016). All sequences were checked for double peaks in the electropherograms and for stop-codons in the amino-acid translations before running further analysis (COI sequences can be accessed under request). A dataset with sequences of *Eualus* specimens and one with sequences of *Galathea* specimens were built. For each genus, the number of haplotypes and haplotype frequencies were computed using the *pegas* package (Paradis, 2010) in R 3.5.0 (R Core Team, 2018).

Molecular identification tools

To assign specimens an identity, COI sequences were singularly blasted on the BOLD Identification System Engine (BOLD IDEngine). For each query sequence, two search databases were used: 'Species-level Barcode Records' ('spDB') and 'All Barcode Records on BOLD' ('allDB'). The 'spDB' includes every COI barcode record with a species-level identification and a minimum sequence length of 500 bp. It delivers a species identification if the query sequence shows less than 1% divergence to a reference sequence (Ratnasingham & Hebert, 2007). The 'allDB', on the other hand, contains every COI barcode record on BOLD with a minimum sequence length of 500 bp. This search only returns a list of the nearest matches and does not provide a probability of placement to a taxon. In this study, query sequences matching reference records up to species level when using the 'spDB' were classified under the resulting species. However, sequences that did not match any reference record up to species level when using the 'spDB' were classified under the species of the most similar match obtained using the 'allDB' and the similarity was reported (Table 2, Table 3). For the identification of the *Eualus* specimens, only public data were considered when using the 'allDB' database, since a preliminary analysis suggested misidentifications among some *Eualus* reference records on BOLD. Therefore, early release and private records were excluded because it was not possible to download them for further inspections. Conversely, for *Galathea* specimens early release and private records were considered as well.

Species delimitation procedure

For species delimitation, the distance-based Automatic Barcode Gap Discovery (ABGD) software (Puillandre *et al.*, 2012) was used. Aligned sequences of both datasets were uploaded to the ABGD webserver (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) to sort them into hypothetical species based on pairwise distances by detecting differences between intraspecific and interspecific variation, i.e., the barcoding gap. Before running the program, prior limits to intraspecific diversity *Pmin* and *Pmax* were set to default values of 0.001 and 0.1 respectively, while the proxy for the minimum gap width (X) was set to 1.5 (Puillandre et al., 2012). P gives approximate indications on the area where the barcoding gap should be detected and X relates to the sensitivity of the method to gap width (Puillandre et al., 2012). Pairwise distances between sequences were calculated using the two available substitution models: Jukes-Cantor (Jukes & Cantor, 1969) and Kimura K80 (Kimura, 1980). The transition/transversion rate under the K80 model was set to $k = 9$.

Phylogenetic analysis

To solve phylogenetic relationships within both genera, all publicly available BOLD and GenBank sequences corresponding to species matches or showing the highest similarity with query sequences, as well as closely related taxa, were downloaded. Moreover, intraspecific and interspecific pairwise distances of newly sequenced and retrieved records were calculated for both genera using MEGA v.7 (Kumar *et al.*, 2016).

The dataset for phylogenetic analyses of the genus *Galathea* included 72 aligned sequences, 23 of which newly sequenced and 49 retrieved from BOLD and GenBank public databases (Table S.2, S.5). A total of 556 bp of the COI mitochondrial gene were used for phylogenetic analyses. Retrieved sequences included 6 species of the genus *Galathea*: *G. dispersa* (Bate, 1859), *G. halia* (Macpherson & Robainas-Barcia, 2015), *G. intermedia* (Lilljeborg, 1851), *G. nexa* (Embleton, 1836), *G. squamifera* (Leach, 1814), and *G. strigosa* (Linnaeus, 1761), as well as one outgroup (*Pisidia longicornis*, Linnaeus 1767). Sequences of *G. squamifera* were not publicly available neither on BOLD nor on GenBank and have been provided courtesy of the University Museum of Bergen (Norway) and Norwegian Barcode of Life (https://www.norbol.org/).

The dataset for phylogenetic analyses of the genus *Eualus* included 53 aligned sequences, 22 of which newly sequenced and 31 retrieved from BOLD and GenBank public databases (Table S. 5). A total of 581 bp of the COI mitochondrial gene were used for phylogenetic analyses. Retrieved sequences included 6 species of the genus *Eualus*: *E. fabricii* (Krøyer, 1841), *E. gaimardii* (H. Milne Edwards, 1837), *E. macilentus* (Krøyer, 1841), *E. pusiolus* (Krøyer, 1841), *E. occultus* (Lebour, 1936) and *E. cranchii* (syn. *Thoralus cranchii* – Leach, 1817), as well as one outgroup (*Hippolyte commensalis* Kemp, 1925).

Phylogenetic relationships were reconstructed using both the maximum likelihood (ML) method (Fukami & Tateno, 1989) and Bayesian inference (Ronquist & Huelsenbeck, 2003). ML trees reconstruction and outputs are available in the supplementary material (Fig. S.

Table 2. Summary of the species identification of *Galathea* specimens obtained from several identification methods. The first column indicates the specimen's name and the second the haplotype of each specimen obtained from the sequencing of the COI gene. Results obtained using morphological characters are listed in the column 'Morphological identification'. The 'ABGD' column indicates the group in which each specimen was clustered when sequences were analyzed using the Automatic Barcode Gap Discovery Software (ABGD). Results of molecular identification of *Galathea* specimens using the BOLD ID Engine are reported in the table as well. Molecular identification was obtained using the databases 'Species-level Barcode Records' (BOLD 'spDB') and 'All Barcode Records on BOLD' (BOLD 'allDB'). For each database, the match of the specimen sequence with the closest species is listed ('Match'). Using the 'spDB', results were: 'none' (impossible to match the specimen to any reference record on BOLD), 'confirmed' (specimen confirmed to belong to that species), 'suggested' (specimen only suggested to belong to that species). 'Sequence availability' indicates the availability of the matched sequence in the database to external users, which is either 'public' (the sequence can be accessed and downloaded) or 'early release' (the sequence is not available for download and geographical and sequencing information are not always available, indicated as na). 'Location' indicates where the BOLD reference records were collected: DE = German Bight, NW = Oslo – Norway, SW = Skagerrak – Sweden, VA = Vanuatu.

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'				
				Match	Similarity [%]	Sequence availability	Record ID	Match	Similarity [%]	Sequence availability	Record ID	Location
CRO1	I	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.34	early release	*na*	*G. squamifera*	99.34	early release	*na*	NW
								G. strigosa	88.06	public	SWEMA471_15	SW
CRO2	II	*G. squamifera*	Grp. 2	*none*				*G. squamifera*	91.75	early release	*na*	NW
CRO3	III	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99	early release	*na*	*G. squamifera*	99	early release	*na*	NW
								G. strigosa	88.56	public	SWEMA471_15	SW
CRO4	II	*G. squamifera*	Grp. 2	*none*				*G. squamifera*	91.61	early release	*na*	NW
CRO5	IV	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.31	early release	*na*	*G. squamifera*	99.31	early release	*na*	NW
								G. strigosa	87.92	public	SWEMA471_15	SW
CRO6	V	*G. squamifera*	Grp. 2	*none*				*G. squamifera*	91.77	early release	*na*	NW
								G. halia	87.37	public	GBCM6663_17	VA
CRO7	II	*G. squamifera*	Grp. 2	*none*				*G. squamifera*	91.77	early release	*na*	NW

Continued

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'				
LIV1	VI	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.28	early release	*na*	*na*
								G. intermedia	88.94	public	BNSDE145_12	DE
LIV2	VII	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89	public	BNSDE145_12	DE
LIV3	VIII	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.55	early release	*na*	*na*
								G. intermedia	89.22	public	BNSDE145_12	DE
LIV4	IX	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.54	early release	*na*	*na*
								G. intermedia	89.22	public	BNSDE145_12	DE
PAL1	X	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.31	early release	*na*	*G. squamifera*	99.31	early release	*na*	NW
								G. strigosa	87.92	public	SWEMA471_15	SW
PAL2	XI	*G. squamifera*	Grp. 1	*G. squamifera (suggested)*	98.62	early release	*na*	*G. squamifera*	98.62	early release	*na*	NW
								G. strigosa	88.24	public	SWEMA471_15	SW
PAL3	XII	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.19	early release	*na*	*G. squamifera*	99.19	early release	*na*	NW
								G. strigosa	88.57	public	SWEMA471_15	SW
PAL4	XIII	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.22	early release	*na*	*na*
								G. strigosa	88.72	early release	*na*	*na*
								G. intermedia	88.66	public	BNSDE143_12	DE
PAL5	XIV	*G. squamifera*	Grp. 1	*G. squamifera (suggested)*	98.93	early release	*na*	*G. squamifera*	98.93	early release	*na*	NW

Continued

Table 2 continued

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'				
								G. strigosa	87.56	public	SWEMA471_15	SW
PAL6	XV	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.21	early release	*na*	*G. squamifera*	99.21	early release	*na*	NW
								G. strigosa	87.92	public	SWEMA471_15	SW
PAL7	IV	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.3	early release	*na*	*G. squamifera*	99.3	early release	*na*	NW
								G. strigosa	87.91	public	SWEMA471_15	SW
PAL8	XVI	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.55	early release	*na*	*na*
								G. strigosa	88.89	early release	*na*	*na*
								G. intermedia	88.83	public	BNSDE143_12	DE
PAL9	XVII	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.31	early release	*na*	*G. squamifera*	99.31	early release	*na*	NW
								G. strigosa	87.86	public	SWEMA471_15	SW
PAL10	XVIII	*G. intermedia*	Grp. 3	*none*				*G. intermedia*	89.43	early release	*na*	*na*
								G. intermedia	89.11	public	BNSDE145_12	DE
PAL11	XIX	*G. squamifera*	Grp. 1	*G. squamifera (confirmed)*	99.02	early release	*na*	*G. squamifera*	99.02	early release	*na*	NW
								G. strigosa	88.29	public	SWEMA471_15	SW
PAL12	XX	*G. squamifera*	Grp. 1	*G. squamifera (suggested)*	97.5	early release	*na*	*G. squamifera*	97.5	early release	*na*	NW
								G. strigosa	86.07	public	SWEMA471_15	SW

Table 3. Summary of the species identification of *Eualus* specimens obtained from several identification methods. The first column indicates the specimen's name and the second the haplotype of each specimen, obtained from the sequencing of the COI gene. Results obtained using morphological characters are listed in the column 'Morphological identification'. The 'ABGD' column indicates the group in which each specimen was clustered when sequences were analyzed using the Automatic Barcode Gap Discovery Software (ABGD). Results of molecular identification of *Galathea* specimens using the BOLD ID Engine are reported in the table as well. Molecular identification was obtained using the database 'Species-level Barcode Records' (BOLD 'spDB') and the database 'All Barcode Records on BOLD' (BOLD 'allDB'). For each database, the match of the specimen sequence with the closest species is listed ('Match'). Using the 'spDB', results were: 'none' (impossible to match the specimen to any reference record on BOLD), 'confirmed' (specimen confirmed to belong to that species), 'suggested' (specimen only suggested to belong to that species). 'Sequence availability' indicates the availability of the matched sequence in the database to external users, which is either 'public' (the sequence can be accessed and downloaded) or 'early release' (the sequence is not available for download and geographical and sequencing information are not always available, indicated as na). 'Location' indicates where the BOLD reference records were collected: DE = German Bight, PT = Alentejo – Portugal, SP = northern Spain, SW = Skagerrak – Sweden, UK = Anglesey – United Kingdom.

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'			
				Match	Similarity [%]	Sequence availability	Record ID	Match	Similarity [%]	Record ID	Location
CRO1	I	*E. cranchii*	Grp. 1			*none*		*E. cranchii*	84.53	JSDUK184_08	UK
CRO2	II	*E. cf cranchii*	Grp. 1			*none*		*E. cranchii*	84.35	JSDUK184_08	UK
CRO3	II	*E. cf cranchii*	Grp. 1			*none*		*E. cranchii*	84.35	JSDUK184_08	UK
CRO4	III	*E. cf cranchii*	Grp. 2	*E. cranchii (confirmed)*	99.12	private	*na*	*E. cranchii*	85.14	JSDUK184_08	UK
CRO5	III	*E. cranchii*	Grp. 2	*E. cranchii (confirmed)*	99.12	private	*na*	*E. cranchii*	85.14	JSDUK184_08	UK
CRO6	IV	*E. cf occultus*	Grp. 1			*none*		*E. cranchii*	84.54	JSDUK184_08	UK
CRO7	V	*E. cf cranchii*	Grp. 3	*E. occultus (suggested)*	98.3	early release	*na*	*E. cranchii*	98.47	JSDUK184_08	UK
CRO8	VI	*E. cranchii*	Grp. 2	*E. cranchii (confirmed)*	99.65	private	*na*	*E. cranchii*	84.95	JSDUK184_08	UK
LIV1	VII	*E. occultus*	Grp. 3	*E. occultus (suggested)*	98.3	early release	*na*	*E. cranchii*	98.28	JSDUK184_08	UK
LIV 2	VIII	*E. cf cranchii*	Grp. 4	*E. cranchii, E. occultus, E. pusiolus (suggested)*	99.82	public	BNSC322_11	*E. cranchii*	99.82	BNSC322_11	DE

Continued

Table 3 continued

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'			
								E. occultus	98.76	BCASB018_16	SP
								E. pusiolus	98.59	SWEMA676_15	SW
								E. occultus	98.24	SWEMA526_15	SW
LIV3	IX	*E. cranchii*	Grp. 3			*none*		*E. cranchii*	93.44	JSDUK185_08	UK
LIV4	V	*E. cf occultus*	Grp. 3	*E. occultus (suggested)*	98.48	early release	*na*	*E. cranchii*	98.47	JSDUK184_08	UK
LIV5	X	*E. cranchii*	Grp. 3	*E. cranchii (suggested)*	97.71	public	JSDUK184_08	*E. cranchii*	97.71	JSDUK184_08	UK
LIV6	XI	*E. cranchii*	Grp. 3	*E. occultus (suggested)*	98.86	early release	*na*	*E. cranchii*	98.85	JSDUK184_08	UK
LIV7	XII	*E. cf occultus*	Grp. 2	*E. cranchii (suggested)*	98.77	private	*na*	*E. cranchii*	85.06	JSDUK184_08	UK
LIV8	II	*E. pusiolus*	Grp. 1			*none*		*E. cranchii*	84.35	JSDUK184_08	UK
LIV9	XIII	*E. cranchii*	Grp. 1			*none*		*E. cranchii*	84.35	JSDUK184_08	UK
LIV10	II	*E. cranchii*	Grp. 1			*none*		*E. cranchii*	84.35	JSDUK184_08	UK
PAL1	XIV	*E. cf occultus*	Grp. 4	*E. cranchii (suggested)*	98.07	public	BNSC322_11	*E. cranchii*	98.07	BNSC322_11	DE
								E. cranchii	97.72	MLALE062_12	PT
								E. occultus	97.36	BCASB018_16	SP
								E. pusiolus	97.19	SWEMA676_15	SW

Continued

Table 3 continued

Specimen	Haplotype	Morphological identification	ABGD	BOLD 'spDB'				BOLD 'allDB'			
								E. occultus	96.84	SWEMA526_15	SW
PAL2	XV	*E. cranchii*	Grp. 3	*E. occultus (suggested)*	97.35	early release	*na*	*E. cranchii*	97.32	JSDUK184_08	UK
PAL3	VIII	*E. cf occultus*	Grp. 4	*E. cranchii, E. occultus, E. pusiolus (suggested)*	99.82	public	BNSC322_11	*E. cranchii*	99.82	BNSC322_11	DE
								E. cranchii	99.12	MLALE062_12	PT
								E. occultus	98.76	BCASB018_16	SP
								E. pusiolus	98.59	SWEMA676_15	SW
								E. occultus	98.24	SWEMA526_15	SW
PAL4	XVI	*E. cf cranchii*	Grp. 4	*E. cranchii, E. occultus, E. pusiolus (suggested)*	99.47	public	BNSC322_11	*E. cranchii*	99.47	BNSC322_11	DE
								E. cranchii	99.12	MLALE062_12	PT
								E. cranchii	99.11	BNSC325_11	DE
								E. occultus	98.94	BCASB018_16	SP
								E. pusiolus	98.59	SWEMA676_15	SW
								E. occultus	98.24	SWEMA526_15	SW

6, Fig. S. 7). The function ModelTest in the *phangorn* 2.5.5 package (Schliep, 2011; Schliep *et al.*, 2017) in R (R Core Team, 2018) was used to find the best model of evolution, and the best-fitting model for the data was chosen based on the Akaike Information Criterion (Akaike, 1973). For both genera, the best model of evolution was found to be the Hasegawa-Kishino-Yano+Gamma distribution+Inversion parameter (HKY+G+I) (Hasegawa *et al.*, 1985). To further confirm the choice of model, the difference between likelihoods resulting from the use of the best-fitting model and the one having one parameter less was tested by computing an Analysis of Variance (ANOVA). Bayesian inference was conducted with partition-specific settings in MrBayes v3.2.7 (Ronquist & Huelsenbeck, 2003). For the genus *Galathea*, a Bayesian Markov Chain Monte Carlo (MCMC) tree search with two independent runs with four chains each was run for 1,440,000 generations and trees were sampled every 100 generations. For the genus *Eualus*, on the other hand, a Bayesian MCMC tree search with two independent runs with four chains each was run for 2,000,000 generations and trees were sampled every 1,000 generations. Convergence diagnostics of the two independent runs were analyzed using Tracer v1.7.1 (Rambaut *et al.*, 2018) and after reaching the stationary phase, the first 25% of trees were discarded in both genera. Trees were edited using the software FigTree v1.4.4 (Rambaut & Drummond, 2012).

Results

Galathea genus

Morphological identification

Morphological identification allowed to separate the *Galathea* specimens into two distinct species: 16 individuals were identified as *G. squamifera* and seven as *G. intermedia* (Table 2). The main difference between the two species was found to be the carapace length. Indeed, the mean carapace length of *G. intermedia* (0.616 cm ± 0.112 cm) was significantly lower (p = 0.00021) than that of *G. squamifera* (1.87 cm ± 0.427 cm) (Fig. S. 4). Another useful discrimination criterion was the shape of the rostrum: the one of *G. squamifera* was short and large, with four evident lateral teeth (LT) on each side and an apical tooth (AT), clearly longer than those on the sides (Fig. 1A), whereas *G. intermedia* had a long and tight rostrum, with an AT and four rather inconspicuous LT on each side (Fig. 1B). Finally, in the epigastric region of *G. squamifera* there were two spines that were absent in *G. intermedia* (Fig. 1A). The morphological characteristics for each individual are summarized in the supplementary material (Table S.1).

COI divergence assessment

The COI gene fragments (556 bp) of the 23 specimens of *Galathea* analyzed produced 20 unique haplotypes. Almost every haplotype matched a different specimen, except haplotype II (shared by three specimens) and haplotype IV (shared by two) (Fig. 2). Specimens sharing haplotype II were all collected in Croatia, while individuals sharing haplotype IV were collected in Croatia and Palinuro (Fig. 1B). Fig. 2A clearly illustrates that the number of mutations between morphologically identified species *G. intermedia* and *G. squamifera* (n = 81) is higher than within species (*max* n = 33 within G. *squamifera*), supporting species subdivision obtained from morphological identification. However, the number of mutations between haplotypes V and XI of *G. squamifera* is higher than other numbers of mutations observed within species from the main haplogroups of *G. squamifera*. Indeed, due to the large genetic distance observed between these two

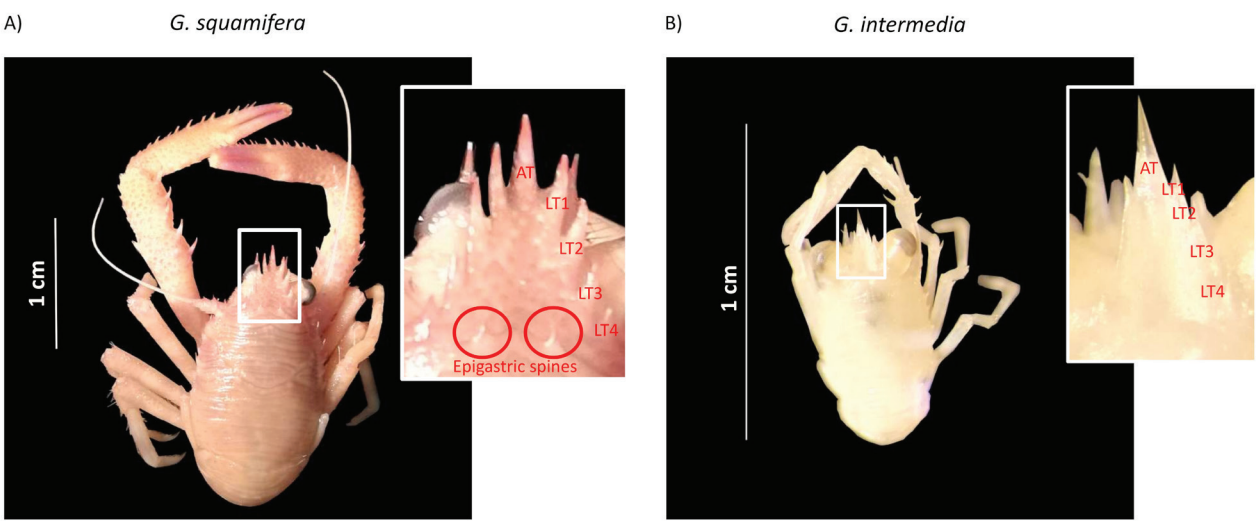

Fig. 1: Morphological differences between *Galathea squamifera* and *G. intermedia*. Pictures of samples of A) *Galathea squamifera* (PAL6) and B) *Galathea intermedia* (PAL10). Enlargements show the rostrum of each species, comprising one apical tooth (AT) and four lateral teeth (LT) on each side. The red circles in A) highlight the two spines on the epigastric region of *G. squamifera*. The photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and an f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the software GIMP v2.10.6 and Inkscape 0.92.3.

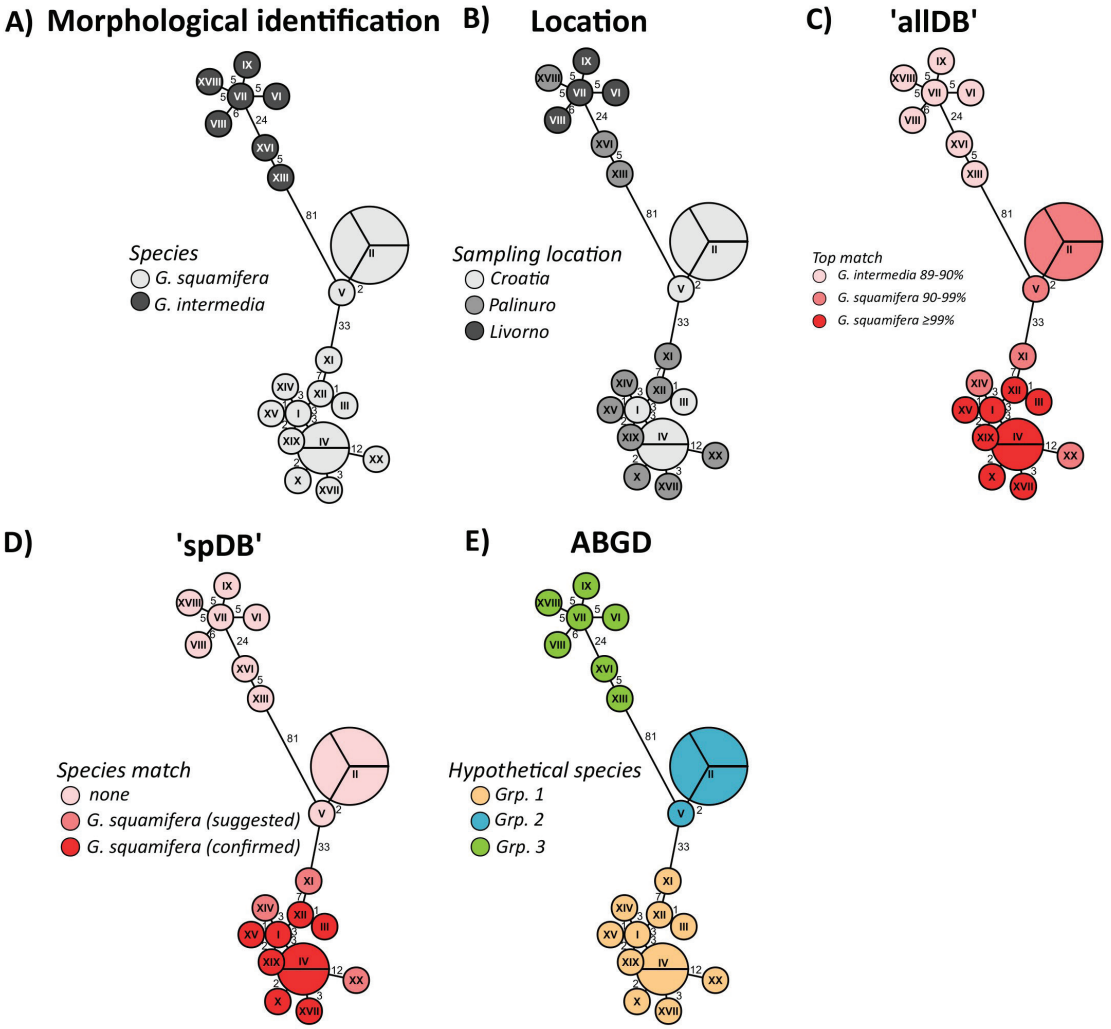

Fig. 2: Haplotype network of *Galathea* genus based on 23 COI sequences. The network is replicated five times to illustrate: A) results obtained from morphological identification, B) geographical distribution, C) results obtained from molecular identification using all barcode records on BOLD (allDB) and D) species-level barcode records on BOLD (spDB), E) results obtained from species delimitation using Automatic Barcode Gap Discovery (ABGD). Each circle represents one haplotype (named using Roman numerals). The size of the circles indicates the haplotype frequencies, and the subdivisions of the pies indicate the number of specimens sharing that haplotype. Numbers between circles indicate the number of mutations between haplotypes.

haplotypes, the overall mean distance within morphologically identified *G. squamifera* was found to be larger (4% ± 0.45%) than the overall mean distance observed within *G. intermedia* (2.75% ± 0.38 %).

Molecular identification using BOLD

Molecular identification of *Galathea* specimens using BOLD IDEngine delivered different outcomes when using the two databases 'allDB' and 'spDB' (Table 2, Fig. 2C and D). On one hand, nine specimens were matched to the species *G. squamifera* with a similarity higher than 99% (Table 2) using 'spDB', while three others were only suggested to be *G. squamifera* since the similarity with reference records was lower, namely ranging between 97.5 and 98.93%. As for the remaining 11 specimens, no match or suggestions were delivered using 'spDB' (Fig. 2D). All specimens confirmed or suggested as *G. squamifera* were matched to early release records sampled in Oslo (Norway). Because the latter are not publicly accessible, no sequencing data were available. Specimens associated with *G. squamifera* were sampled in Croatia and in Palinuro but not in Livorno (Table 2, Fig. 2B). On the other hand, results obtained using 'allDB' were coherent with those obtained from morphological identification. Indeed, the number of individuals showing a significant similarity (between 91.62 and 99.34%) with early release records of *G. squamifera* sampled in Oslo (Norway) were 16 (4 more than when using 'spDB'). Moreover, the remaining seven individuals showed the greatest similarity (between 88.66 and 89.55%) with reference records of *G. intermedia* (Fig. 2C). Some of them were sampled in the German Bight of the North Sea (Germany), while for some others the sampling location is unknown because they were early release records (Table 2).

Species delimitation using ABGD software

Species delimitation based on COI sequences using ABGD suggested the presence of three distinct groups of taxa among collected *Galathea* specimens (Fig. 2E, Table 2), as opposed to the two indicated by both morphological and molecular identification using BOLD IDEngine. These results were obtained with the two substitutions models used – either Jukes-Cantor (Jukes & Cantor, 1969) or Kimura K80 (Kimura, 1980). The subdivision into three hypothetical species perfectly matched the structure of the haplotype network (Fig. 2C). Group 1 included 12 specimens, all of which corresponded to those suggested or confirmed as *G. squamifera* using both 'spDB' and 'allDB' (Fig. 2C and D). Group 2 comprised four analyzed specimens, for which no match/suggestions to reference records was provided using 'spDB'. However, when 'allDB' was employed, they showed the greatest similarity with reference records of *G. squamifera* (Fig. 2C and D). Group 3 contained the remaining seven specimens, for which no match/suggestion was provided using 'spDB' neither, while they showed the highest similarity with *G. intermedia* when using 'allDB' (Fig. 2C and D).

Phylogenetic inference using mitochondrial COI gene

To calculate the interspecific and intraspecific mean pairwise distances, results obtained from species delimitation using ABGD were considered, and the three suggested groups of taxa were treated as additional species to the six *Galathea* species downloaded (Table S. 2), for a total of nine *Galathea* species and one outgroup (*P. longicornis*). Range values of mean pairwise distances observed on the whole dataset were within species 0% ± 0% (*G. squamifera* reference records) to 2.54% ± 0.46% (ABGD Group 3) (Table S. 3), and within species 1.1% ± 0.4% to 19% ± 1.7% (Table S. 4). Nevertheless, considering Group 1 of ABGD and *G. squamifera* as the same species, the range of mean pairwise distances between species increased, spanning from 8.22% ± 1.1% to 19% ± 1.7 %. The intraspecific mean pairwise difference of *G. squamifera* and ABGD Group 1, together, was very similar to the one of only *G. squamifera* reference records (0.95% ± 0.17 %). Phylogenetic trees clearly show that newly sequenced specimens are divided into four clades (Fig. 3 and Fig. S. 6). One clade corresponds to Group 1 of ABGD (orange colored), one to Group 2 of ABGD (blue colored) and two to Group 3 of ABGD (green colored). The 12 specimens constituting Group 1 of ABGD (Table 2) fall within the same clade of *G. squamifera* reference records, reinforcing the results obtained from molecular identification using BOLD IDEngine, which indicated a match/suggestion of these specimens up to species level with *G. squamifera* (Fig. 2D). Conversely, specimens representing the two other groups of ABGD compose clades separated from any other *Galathea* species existing in Europe for which a COI sequence was available for download, and their species identity could not be deduced by phylogenetic reconstructions. Indeed, Group 2 builds the sister clade of G. *squamifera*, while specimens belonging to Group 3 of ABGD show a mean pairwise distance between them of 2.54% ± 0.46% (Table S. 3) and compose two sister clades of *G. intermedia* (Fig. 3 and Fig. S. 6). Bayesian reconstruction is supported by values of posterior probability (PP) almost always higher than 96%, except for the clade formed by *G. nexa* and *G. dispersa* (56%). This low PP value can be explained by the fact that only two sequences corresponding to *G. nexa* were used to reconstruct the phylogenetic tree. Unfortunately, no additional records were available for downloading. Additionally, low PP values can be observed also within clades formed by a single species or ABGD group, but these poorly supported splits are not considered in the results here reported. Evolutionary analyses clearly separated specimens of Group 2 and those of Group 1 into two sister clades that are clustered together into a well-supported (PP = 100%) monophyletic group, which could indicate the presence of cryptic species. Group 3 forms a well-supported (PP = 100%) monophyletic group, together with *G. intermedia* as well, suggesting the presence of a second cryptic species in the dataset. The mean pairwise distance calculated within specimens of Group 1 and Group 2 amounts to a total of 4.29% ± 0.48 %, which is higher than the highest intraspecific mean pairwise distance observed within *Galathea* species of the studied dataset (Group 3 = 2.54% ± 0.46 %) (Table S. 3). Nevertheless, the genetic distance between Group 1 and Group 2 samples is lower (8.2% ± 1.1 %) (Table S. 4) than the lowest mean interspecific pairwise distance between the *Galathea* species of the studied dataset (12.3% ± 1.4 %). The genetic distance found between the *G. intermedia* reference records collected in the North Sea and the newly sequenced organisms constituting Group 3 (11% ± 1.4 %) remains lower than the interspecific levels observed for the genus *Galathea* (Table S. 4), suggesting a lack of gene flow between North Sea and Mediterranean Sea. All 19 reference records of *G. intermedia* originate from the North Sea, and no records sampled in the Mediterranean were available for download. Additionally, Group 3 is characterized by an intraspecific mean pairwise distance (2.54% ± 0.46 %) that is at least twice as large as those calculated within the remaining eight *Galathea* species (Table S. 3). Indeed, specimens of Group 3 form two sister clades in the phylogenetic reconstructions rather than only one, as the other ABGD detected groups.

Eualus genus

Morphological identification

Morphological identification resulted in three *Eualus* species among the 22 specimens analyzed: *E. occultus, E. cranchii* and *E. pusiolus*. One individual was assigned to the species *E. pusiolus*, nine samples were identified with certainty and six with uncertainty as *E. cranchii*, and one sample was identified with certainty and five with uncertainty as *E. occultus* (Fig. 4A, Table 3). The morphological analysis was mainly based on the rostrum, which

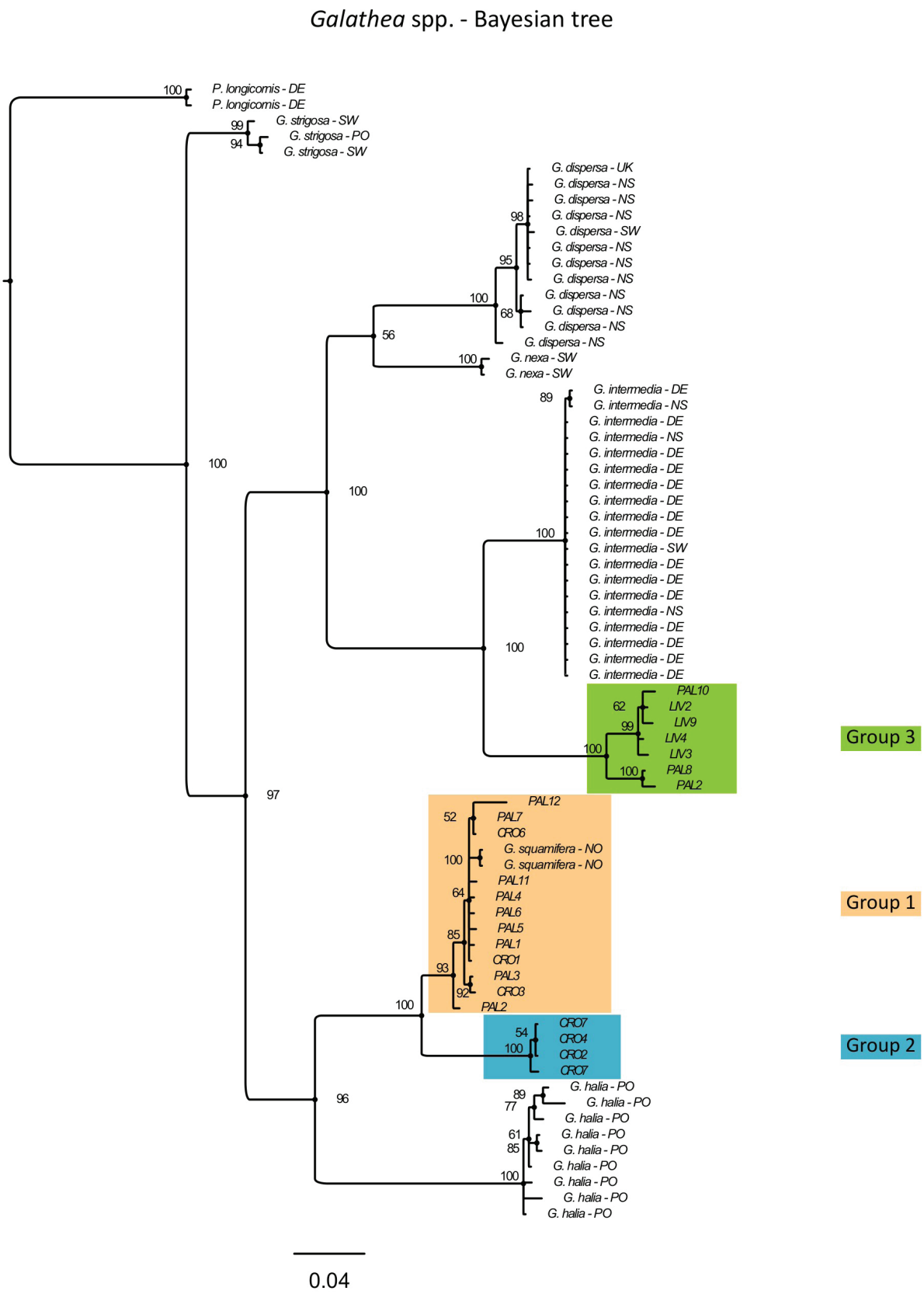

Fig. 3: Bayesian phylogenetic reconstruction of *Galathea* specimens based on COI sequences. MCMC searches were performed based on the HKY+G+I substitution model (Hasegawa et al., 1985). The tree was rooted using the outgroup *Pisidia longicornis* (Linnaeus, 1767). Posterior probabilities (PP) are represented in percentages next to the nodes. Colors represent species delimitation obtained from the ABGD software. The abbreviations next to reference records indicate the sampling locations: DE = German Bight (North Sea), NO = Norway, NS = North Sea, PO = Pacific Ocean, SW = Sweden (Skagerrak), and UK = United Kingdom – North Sea. The analysis involved 72 nucleotide sequences. Evolutionary analysis was conducted in MrBayes v3.2.7 (Ronquist and Huelsenbeck, 2003) and in FigTree v1.4.4 (Rambaut and Drummond, 2012).

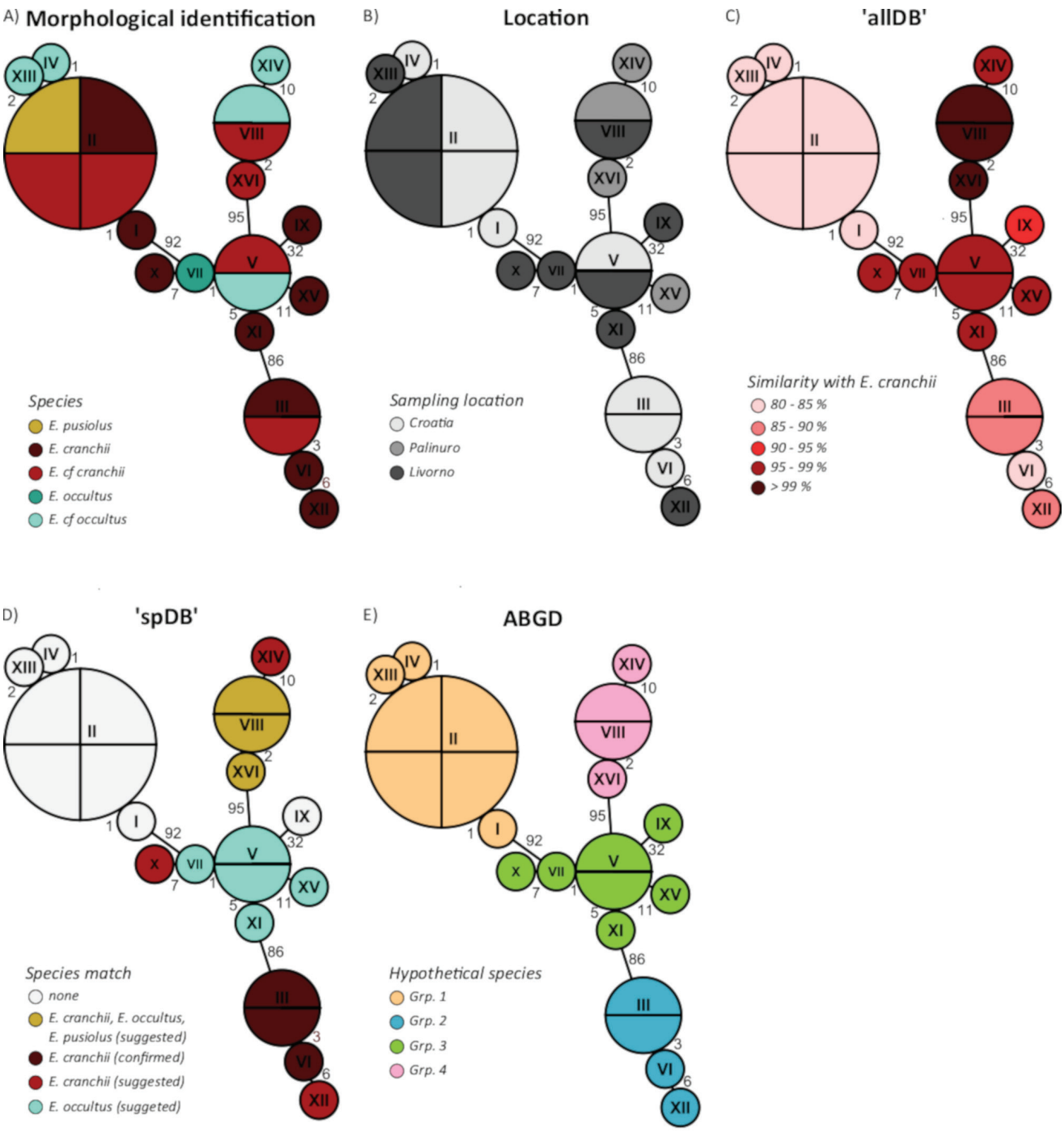

Fig. 4: Haplotype network of *Eualus* genus based on N=22 COI sequences. The network is repeated five times to illustrate: A) results obtained from morphological identification, B) geographical distribution, C) results obtained from molecular identification using all barcode records on BOLD (allDB) and D) species-level barcode records on BOLD (spDB), E) results obtained from species delimitation using Automatic Barcode Gap Discovery. Colors of A), C) and D) are coordinated: red shades refer to *E. cranchii,* green shades to *E. occultus* and yellow shades to *E. pusiolus*, *E. occultus* and *E. cranchii*. Each circle represents one haplotype (named using Roman numerals). The size of the circles indicates haplotype frequencies and the subdivisions of the pies indicate the number of specimens sharing that haplotype. Numbers between circles indicate the number of mutations between haplotypes.

was found to be at least bifurcated and shorter than the scaphocerite in all individuals. *E. pusiolus* was identified thanks to the presence of a pterygostomial tooth, which is absent in *E. cranchii* and *E. occultus* (Fig. 5A). Specimens identified with certainty as *E. cranchii* were distinguished from *E. occultus* because they had a trifurcated rostrum (Fig. 5B), which never occurs in *E. occultus*, and the merus of the 4th and 5th pereiopods with spines arranged in comb-like fashion on the flexor border (Fig. 5C). On the other hand, specimens identified with dubiety as *E. cranchii* had a bifurcated rostrum, such as *E. occultus* (Fig. 5D), and their identification was based on the

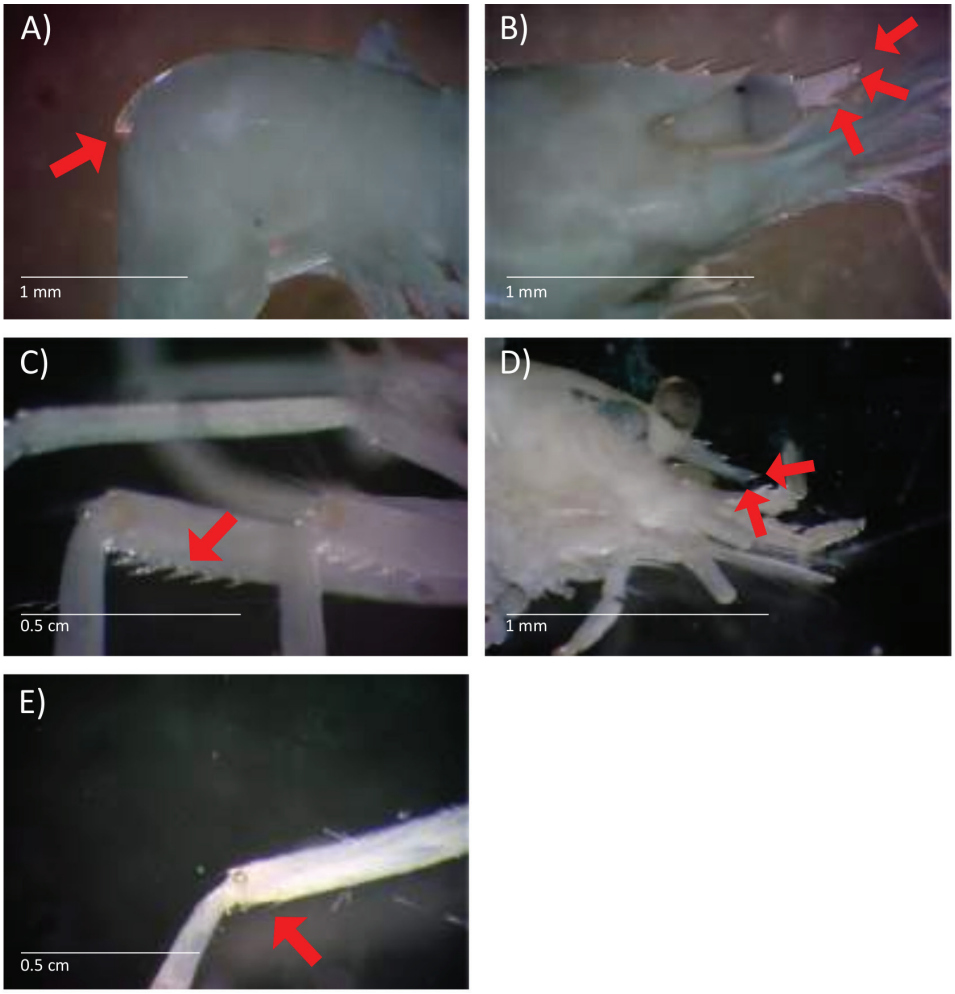

Fig. 5: Morphological differences between *E. cranchii, E. occultus* and *E. pusiolus*. The photographs show: A) pterygostomial tooth of *E. pusiolus*; B) trifurcated rostrum of *E. cranchii*; C) merus of the 4th and 5th pereiopods with spines arranged in comb-like fashion on the flexor border of *E. cranchii*; D) bifurcated rostrum of *E. occultus*; and E) merus of the 5th pereiopod without spines arranged in comb-like fashion on the flexor border of *E. occultus.* The photographs of each organism were taken using a MoticamBTU® camera connected to a Leica WILD M8® stereomicroscope.

fact that the merus of the 5th pereiopod of *E. cranchii* had spines arranged in comb-like fashion on the flexor border, while the one of *E. occultus* did not (Fig. 5E). However, damages to the 5th pereiopod in several specimens impeded to identify them with certitude. In specimens where the 5th pereiopod was completely absent, identification as either *E. cranchii* or *E. occultus* was based on *i)* the similarity with confirmed specimens, *ii)* the presence of the mandibular palp in *E. occultus*, which is absent in *E. cranchii*, and *iii)* the shape of the dorso-distal process of the 3rd segment of the antennular process, which is narrowly triangular in *E. occultus* and broadly triangular in *E. cranchii*. However, since mandibular palps can be easily lost and the difference in the shape of the 3rd segment of the antennular process is minimal, the identification of these samples remained unconfirmed (*E. cf cranchii* and *E. cf occultus*) (Table 3).

COI divergence assessment

The COI gene fragments (573 bp) of the 22 specimens of *Eualus* analyzed produced 16 unique haplotypes, four of which were shared by more than one individual (haplotype III, V and VIII), while all the others were represented by only one specimen (Fig. 4). Haplotype II was shared by four individuals, two sampled in Croatia and two in Livorno (Fig. 4B). Notably, despite their sharing the same haplotype, morphological traits allowed to identify one of them as *E. pusiolus*, another as *E. cranchii* and the other two as *E. cf cranchii* (Fig. 4A).

Molecular identification using BOLD

Results obtained from molecular identification of *Eualus* specimens based on COI sequences using BOLD IDEngine are summarized in Table 3 and Fig. 4C and D. Outputs obtained from the use of the two databases 'spDB' and 'allDB' were not compared since with the first one, results related to private and early release records were considered as well, while with 'allDB' only public results were taken into account to avoid the lack of sequencing and geographical information. Unfortunately, no sequencing or geographical information for private records nor for early release records is publicly

accessible. Using 'spDB', three specimens were identified with certainty as *E. cranchii* with a probability of placement between 99.12 and 99.65%, 11 individuals were only suggested and not confirmed as *E. cranchii* and/or *E. occultus* and/or *E. pusiolus*, and the remaining eight did not return any match against reference records available on BOLD (Fig. 4D). On the other hand, by removing private and early release records when using 'allDB', all 22 specimens showed the highest similarity – between 84.35 and 99.82% – only with *E. cranchii* reference records (Table 3, Fig. 4C). Both molecular and morphological identification pointed out the presence of three species within the studied dataset: *E. cranchii*, *E. occultus* and *E. pusiolus*. However, molecular identification up to species level (Fig. 4D) showed very different results than morphological identification (Fig. 4A). Firstly, out of the 15 potential specimens of *E. cranchii* identified using morphological traits (Table 3, Fig. 4A), only three specimens were identified with certainty as *E. cranchii* using COI sequences (Table 3, Fig. 4D). Secondly, out of the six potential specimens of *E. occultus* identified using morphological characters (Table 3, Fig. 4A), only two were at least suggested to be *E. occultus* by molecular identification (Table 3, Fig. 4D). Thirdly, the only specimen morphologically identified as *E. pusiolus* (Fig. 4A) did not match any species on BOLD using molecular identification, even though *E. pusiolus* reference records were available on the database (Fig. 4D). Finally, using 'spDB' and 'allDB', three and four specimens respectively were matched to three species simultaneously (*E. cranchii*, *E. occultus* and *E. pusiolus*) with almost the same probability of placement (>99%) (Table 3). Results obtained from molecular identification using BOLD IDEngine further show that when private and early release records on BOLD were not considered (Table 3, Fig. 4C), the percentage of similarity between query sequences and *E. cranchii* reference records decreased consistently compared to when all available sequences were taken into account (Table 3, Fig. 4D). For example, the three specimens that matched a private record of *E. cranchii* with a probability of placement between 99.12 and 99.65% using 'spDB' (CRO4, CRO5 and CRO8) showed a similarity with *E. cranchii* of just 85.14% when only public records were taken into account (Table 3, Fig. 4C). Specimen LIV7 showed the highest similarity (98.77%) with the same private record of *E. cranchii*, but when only public records were considered, its similarity with an *E. cranchii* record decreased to 85.06% (Table 3). Additionally, the four samples together constitute a group that is separate from the others in the haplotype network (haplotypes III, VI and XII) (Fig. 4). However, since the specimens were collected in Croatia and Livorno (Fig. 4B), a genetic divergence from other *E. cranchii* reference records due to geographical isolation is unlikely.

Species delimitation using ABGD software

Species delimitation based on COI sequences using the ABGD indicated the presence of four hypothetical species within the analyzed specimens (Table 3, Fig. 4E), as opposed to the three suggested by morphological and molecular identification using BOLD IDEngine. Results were the same with both substitution models used, either with Jukes-Cantor (Jukes & Cantor, 1969) or Kimura K80 (Kimura, 1980). Species delimitation obtained from ABGD was the only one perfectly matching the structure of the haplotype network (Fig. 4E). Group 1 included seven specimens corresponding to haplotypes I, II, IV and XIII, which showed the lowest similarity with *E. cranchii* records using 'allDB' and for which no match/suggestion was provided using 'spDB' (Fig. 4C and D). Group 2 comprised four specimens corresponding to haplotypes III, VI, and XII, which were matched to a private *E. cranchii* record using 'spDB' (Fig. 4D), but they showed a low similarity with *E. cranchii* records (< 90%) using 'allDB' (Fig. 4C). Group 3 comprised seven other samples, corresponding to haplotypes V, VII, XI, XV, IX and X, which showed the highest similarity with early release records of *E. occultus* and with *E. cranchii* sampled in Anglesey, UK (Fig. 4D). Group 4 includes the remaining four specimens and is represented by haplotypes VIII, XIV and XVI, which were found to be *E. cranchii*, *E. occultus* and *E. pusiolus* with the same percentage of similarity with both 'spDB' and 'allDB'.

Phylogenetic inference using mitochondrial COI gene

To calculate the interspecific and intraspecific mean pairwise distances (p-distance), results obtained from species delimitation using ABGD were considered, and the four suggested groups of taxa were treated as additional species to the six *Eualus* sequences retrieved (Table S. 5), for a total of ten *Eualus* species and one outgroup (*Hippolyte commensalis*). Range values of mean p-distance observed on the whole dataset were within species 0.07% ± 0.07% (*E. fabricii*) to 13.66% ± 0.91% (*E. cranchii*) (Table S. 6) and within species 1.8% ± 0.4% to 23% ± 1.4% (Table S. 7). A relatively high mean intraspecific p-distance, compared to other values observed within the dataset, was observed for *E. pusiolus* as well (9.52% ± 0.80 %) (Table S. 6). The lowest interspecific mean p-distance was observed between Group 4 of ABGD and *E. occultus* (1.8% ± 0.4 %), while the highest between Group 1 of ABGD and *E. gaimardii* (23% ± 1.4 %). The mean p-distance between the outgroup *H. commensalis* and *Eualus* species ranged between 22.6% ± 1.6% with *E. pusiolus* and 24.9% ± 1.7% with Group 2 of ABGD (Table S. 7). The mean intraspecific p-distance within groups of ABGD was found to be between 0.2% ± 0.08% for Group 1 and 2.53% ± 0.37% for Group 3. Phylogenetic reconstruction clearly divided newly sequenced *Eualus* specimens into four distinct clades (Fig. 6, Fig. S. 7). The tree was supported by values of posterior probability (PP) almost always higher than 93%, except for the separation between ABGD Group 4 and ABGD Groups 1, 2 and 3 (72%) and for the separation between ABGD Groups 2 and 3 (61%). The subdivision into four clades perfectly matched species delimitation using ABGD software (Table 3), with the orange colored clade

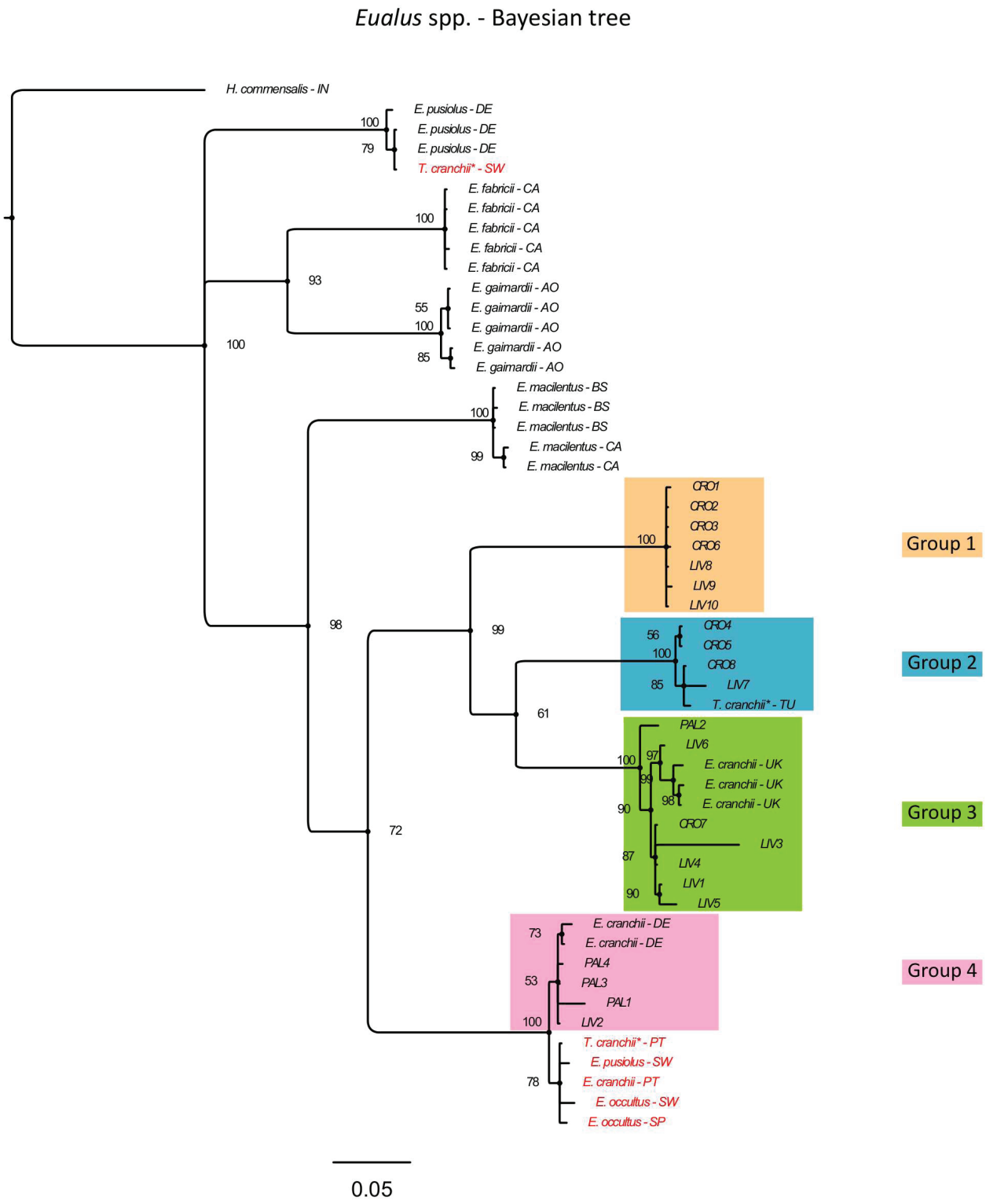

Fig. 6: Bayesian phylogenetic reconstruction of *Eualus* specimens based on COI sequences. MCMC searches were performed based on the HKY+G+I substitution model (Hasegawa et al., 1985). The tree was rooted using the outgroup *Hippolyte commensalis* (Kemp, 1925). Posterior probabilities (PP) are represented in percentages next to the nodes. Colors represent species delimitation obtained from ABGD software. The abbreviations next to the reference records indicate the sampling location: AO = Arctic Ocean, BS = Bering Sea, CA = Canada, DE = German Bight (North Sea), IN = Indonesia, PT = Portugal, SP = northern Spain, SW = Sweden (Skagerrak), TU = Turkey, and UK = Anglesey – UK (North Sea). Sequences in red indicate doubtful species assignment on public databases. The analysis involved 53 nucleotide sequences. Evolutionary analyses were conducted in MrBayes v3.2.7 (Ronquist and Huelsenbeck, 2003) and in FigTree v1.4.4 (Rambaut and Drummond, 2012). **Thoralus cranchii* is currently accepted as *Eualus cranchii*, but these sequences were available on the databases under the name *Thoralus cranchii*.

corresponding to Group 1, the blue colored one to Group 2, the green colored one to Group 3 and the pink colored one to Group 4 (Fig. 6). Specimens constituting Group 1 of ABGD fell within a clade without any reference record in it, thus their species identity could not be inferred from the tree. Specimens of Group 2 clustered together with a reference record of *E. cranchii* (syn. *T. cranchii*) sampled in Turkey, while individuals belonging to Group 3 fell within a clade with a reference record of *E. cranchii* sampled in Anglesey, UK. The interspecific mean p-distance between Group 2 and Group 3 was found to amount to 15.8% ± 1.5 %. On the other hand, specimens belonging to Group 4 fell within a clade together with three distinct species: *E. pusiolus* collected in Sweden, *E. occultus* collected in Sweden and northern Spain, *E. cranchii* collected in the German Bight of the North Sea and in Portugal. These samples correspond to those that matched three different species using 'allDB' and 'spDB' in BOLD IDEngine (Table 3, Fig. 4C and 4D). Strangely, on one hand, a reference record of *E. pusiolus* sampled in Sweden clustered with *E. cranchii* and *E. occultus* reference records as well as with specimens constituting Group 4 of ABGD instead of falling in the clade comprising three *E. pusiolus* individuals collected in the German Bight of the North Sea (Fig. 6, Fig. S. 1). However, on the other hand, a reference record of *E. cranchii* (syn. *T. cranchii*) sampled in Sweden clustered with three different reference records of *E. pusiolus* (Fig. 6, Fig. S. 7).

Discussion

The use of the appropriate taxonomy and the accurate identification of species are fundamental to correctly conduct ecological studies and monitor biodiversity (Bortolus, 2008; Molinari-Jobin *et al.*, 2012; Pante *et al.*, 2015). In the last decade, integrative taxonomy has proven to be a very efficient and reliable approach to reveal the existence of new species (De Queiroz, 2007; Schlick-Steiner *et al.*, 2010; Yeates *et al.*, 2011). Here, the combined use of morphological and molecular approaches to assign an identity to species revealed the potential presence of crypticity within the genera *Galathea* and *Eualus*. As regards the *Galathea* genus, morphological identification led to an underestimation of species richness compared to results obtained from species delimitation based on COI sequences using ABGD software, confirming that the use of morphological characters to assign an identity to specimens could result in biodiversity underestimation (Lefébure *et al.*, 2006), and that the approach overlooks morphologically similar taxa (Hebert *et al.*, 2003; Caputi *et al.*, 2007). The observed divergences in the results – obtained by applying different identification tools – highlight the importance of using an integrative approach to assign an identity to species, as already mentioned in multiple studies (Rubinoff, 2006; Wiemers & Fiedler, 2007; Valentini *et al.*, 2009).

Galathea genus

Morphology alone allowed to identify all *Galathea* specimens as either *G. squamifera* or *G. intermedia*, but DNA barcoding of the COI gene confirmed the presence of *G. squamifera* and suggested the existence of two still undescribed species: one cryptic to *G. squamifera* and one cryptic to *G. intermedia*. However, no morphological differences could be observed between described and presumed cryptic species; for this reason, based on the present data only, no formal decision on the taxonomic status of these species will be taken until a higher number of specimens and/or additional independent molecular markers are analyzed.

Not only did morphological identification underestimate the number of *Galathea* species within the studied dataset, but it also assigned a doubtful identity to specimens. For example, it allowed to recognize some specimens as *G. intermedia*, although phylogenetic reconstructions based on COI sequences clustered those specimens in a clade sister to *G. intermedia* reference records, and molecular identification using 'spDB' in BOLD IDEngine did not assign them to *G. intermedia* reference records available in the library. Similarly, some specimens morphologically identified as *G. squamifera* were clustered in a clade sister to *G. squamifera* in the phylogenetic reconstructions, and since molecular identification using 'spDB' in BOLD IDEngine did not associate these specimens with *G. squamifera* reference records neither, results obtained from morphological identification remain doubtful. Unlike the 'spDB', the use of 'allDB' delivered the same results as the morphological approach. Therefore, the choice of the database has a significant effect on the results displayed: thus, the BOLD IDEngine should be used with caution, and results compared against other databases and interpreted judiciously. Nevertheless, phylogenetic reconstructions partially supported results obtained from morphological identification since specimens suggested as *G. squamifera* by BOLD IDEngine clustered within the same clade as *G. squamifera* reference records. This underlines again the importance of relying on multiple identification tools to assign an identity to specimens. Indeed, molecular identification through DNA barcoding can also have limitations, such as the fact that useful information is mainly available for groups that are already well studied (Rubinoff, 2006) and, unfortunately, no previous molecular studies on *G. squamifera* or *G. intermedia* had been conducted. Nonetheless, here DNA barcoding of the COI gene was confirmed to be a suitable genetic marker to distinguish between *Galathea* species, as already highlighted in previous studies (Silva *et al.*, 2011; Macpherson & Robainas-Barcia, 2015; Raupach *et al.*, 2015).

While the haplotype network already suggested a mistake in morphological identification, since a significant number of mutations (n=33) was observed between two specimens – which were both identified as *G. squamifera* – ABGD questioned even more the morphological results by separating the two specimens into two distinct groups of taxa. The advantage of using ABGD software is that

group subdivisions do not have to be necessarily considered distinct species; quite the contrary, since it is based on the barcoding gap and since there are still conflicting opinions regarding its existence and suitability for separating species (Meier *et al.*, 2006; Wiemers & Fiedler, 2007), they can be considered as distinct haplogroups belonging to the same species, which could be genetically more distant compared to within groups distance because of the lack of gene flow between them. Nevertheless, the results indicated that specimens of Group 1 and Group 2 cannot be simply considered as belonging to two separate phylogeographic groups, since both include samples collected from the same geographic location (Croatia). Additionally, samples collected in the North Sea (Oslo-Norway) fell within the same clade as samples collected in the Mediterranean Sea, indicating the absence of genetic isolation due to physical separation between *G. squamifera* populations, and reinforcing the hypothesis that *G. squamifera* (Group 1) and Group 2 correspond to two cryptic species.

Conversely, phylogenetic trees indicated that *G. intermedia* reference records, collected in the German Bight of the North Sea (Germany), formed a well-supported separate clade from specimens of Group 3 collected in the Mediterranean Sea. This observation suggests that the two clades might correspond to two distinct phylogeographic groups of *G. intermedia*, between which gene flow is absent. In support of this argument, there is no evidence of migration of *G. intermedia* neither within nor outside the German Bight (Kronenberger & Türkay, 2003). Described levels of intra-/interspecific variation showed interesting results, although they must be considered with caution since some taxa were characterized by only two specimens and this number is not sufficient to fully grasp the genetic diversity of the species (Zhang *et al.*, 2010). Firstly, the mean p-distance between *G. intermedia* and specimens of Group 3 matched those found in a previous study between *Galathea* species, which ranged from 7.2% to 24.6% (Macpherson & Robainas-Barcia, 2015). This result is inconsistent with the assumption that there are two phylogeographic groups within *G. intermedia* and suggests that the two clades correspond to two cryptic species resulting from an allopatric speciation. The mean p-distance between *G. squamifera* Group 1 and Group 2 was comprised within the range mentioned above as well, reinforcing the theory that the specimens of Group 2 belong to a cryptic species of *G. squamifera*. Secondly, the interspecific mean p-distance between specimens of Group 1 and *G. squamifera* is much lower than the range described by Macpherson *et al.* (2015), which suggests that they are the same species. In fact, the intraspecific mean p-distance of the merged groups remained similar to values reported within other *Galathea* species investigated, further supporting results obtained from morphological and molecular identification, which revealed that the specimens of Group 1 were *G. squamifera*. Finally, the intraspecific mean p-distance of Group 3 is the highest of all analyzed taxa: consequently, phylogenetic reconstructions divided the specimens of this group into two sister clades. However, the present data do not allow to conclude whether specimens of Group 3 belong to two distinct species or to a single species characterized by a higher genetic divergence than other *Galathea* species. Even though the mean intraspecific p-distance of Group 3 is already higher than within other taxa, the low number of specimens forming Group 3 is not sufficient to describe most of the genetic diversity of a species (Zhang *et al.*, 2010), suggesting that a larger sample size could expand the barcoding gap between these specimens and reveal the presence of two distinct species.

Eualus genus

Within the genus *Eualus*, morphological and molecular identification showed contrasting results and the identity of the 22 individuals collected in the Mediterranean Sea has remained unclear. Morphological traits allowed to identify all *Eualus* specimens as *E. cranchii* or *E. occultus* or *E. pusiolus*, but species delimitation using ABGD and phylogenetic reconstruction based on COI sequences suggested the existence of four groups of taxa, thus uncovering the potential presence of an unknown species.

Morphological identification of *Eualus* specimens revealed to be very difficult due to the presence of subtle morphological differences; hence, only half of the specimens could be identified with certainty, demonstrating that the use of morphological traits for species identification has several limitations (Windig *et al.*, 2004; Caputi *et al.*, 2007; Pérez-Barros *et al.*, 2008). Morphological keys are sometimes only effective for a particular life stage or gender (Hebert *et al.*, 2003; Lefébure *et al.*, 2006; Valentini *et al.*, 2009): since in this study individuals were not selected based on sex, this morphological trait was not properly used, leading to possible species misidentifications. Nevertheless, two of the specimens sharing haplotype II were unambiguously morphologically identified as two distinct species (*E. cranchii* and *E. pusiolus*), raising the possibility that the COI is not a good genetic marker to distinguish between these two species, even if several previous studies had confirmed the appropriateness of the COI gene to distinguish between cryptic and non-cryptic *Eualus* species (Silva *et al.*, 2011; Nye *et al.*, 2013; Bilgin *et al.*, 2015; Vassily *et al.*, 2017). Indeed, COI barcodes may offer only a fraction of the information needed to characterize species and may not be representative of the whole genome (Rubinoff, 2006). Hence, independent nuclear markers such as SNPs or RAD tags may be useful for a better species delineation of marine crustaceans (Raupach *et al.*, 2015).

Just like morphological identification, the use of the BOLD IDEngine suggested the existence of three species: *E. cranchii*, *E. occultus* and *E. pusiolus*. However, several samples did not match any reference record on BOLD, indicating that the species in the dataset could be more than just three. Furthermore, some specimens were assigned to multiple species simultaneously, and it is very likely that their presence is due to species misidentifications among BOLD reference records corresponding to *E. cranchii*, *E. occultus* and *E. pusiolus*: for example,

some specimens showed the same similarity with a reference record of *E. occultus* and one of *E. pusiolus*, both collected in Skagerrak (Sweden), but the similarity between these two records is too significant to believe that they belong to distinct species. In fact, phylogenetic reconstructions based on the COI gene suggested that the distance between *E. occultus* sampled in Sweden and *E. pusiolus* sampled in the German Bight of the North Sea is much higher by clustering them in separated clades. False species identifications are very common in taxonomy (Tautz *et al.*, 2003) and previous studies had already showed that BOLD is a potential source of species assignment misplacement (Bilgin *et al.*, 2015; Meiklejohn *et al.,* 2019). Moreover, the presence of numerous early release and private reference records questions the effectiveness of using BOLD IDEngine to distinguish *Eualus* specimens. For example, only three *Eualus* specimens were assigned up to species level with a probability of placement higher than 99%. However, they were matched to a single private record of *E. cranchii*, and when only public reference records were considered, the same specimens showed a probability of placement within *E. cranchii* of only 85.06%. Hence, the fact that the private record in question is indeed an *E. cranchii* is doubtful since it appears to be very genetically different than the several public ones.

In the phylogenetic reconstruction, specimens constituting ABGD Group 1 fell within a clade without any reference record in it, suggesting the existence of a still unknown and undescribed species among the analyzed organisms. This assumption was suggested by the use of BOLD IDEngine as well, either by using 'spDB', which did not match those specimens to any reference record, or by using 'allDB', which classified those specimens as those with the lowest similarity to *E. cranchii*. Nevertheless, the phylogenetic reconstruction did not allow to clarify the conflicting species assignment obtained from morphological identification and the BOLD IDEngine. For example, while specimens of ABGD Group 2 clustered in a clade containing a reference record of *E. cranchii* collected in Turkey, specimens of ABGD Group 3 clustered in a separate clade containing reference records of *E. cranchii* collected in Anglesey-UK, indicating that Group 2 and Group 3 represent two phylogeographic groups of *E. cranchii*. However, only public reference records available for downloading were used to build phylogenetic reconstructions, and the lack from the phylogenetic tree of the early release and private records that were considered when using BOLD IDEngine could lead to misinterpretations of results. In fact, while phylogenetic analyses suggested that specimens of Group 3 were *E. cranchii*, 'spDB' identified almost all of them as *E. occultus* because of their similarity with some early release records not available for download. The assignment of specimens constituting ABGD Group 4 to a species was problematic as well. On one hand, the low interspecific p-distance observed between specimens of ABGD Group 4 and *E. occultus* suggested that these specimens could be *E. occultus*. Moreover, morphological identification categorized two specimens as *E. cf occultus* and the other two as *E. cf cranchii*, and since their identification has remained dubious, it is probable that the two identified *E. cf cranchii* were in reality *E. occultus*. Molecular identification using 'spDB' partially confirmed this inference, since it suggested that three out of four specimens may potentially be *E. occultus*. On the other hand, however, once that only public reference records were considered, specimens of ABGD Group 4 matched *E. cranchii* reference records with a probability of placement > 95%. Finally, Bayesian reconstruction clustered specimens of ABGD Group 4 with three distinct species of *Eualus*, making it impossible to conclude whether they were *E. cranchii*, *E. occultus* or *E. pusiolus*. The aggregation of these reference records within the same clade, although corresponding to distinct species, supported the already introduced hypothesis that there are false species identifications among reference records in BOLD, an assumption that can be supported by the high values of intraspecific mean p-distance measured within *E. cranchii* and *E. pusiolus* compared to the one observed within other *Eualus* species, such as *E. fabricii* and *E. macilentus*.

In conclusion, the use of additional genetic markers, as well as of a higher number of samples and reference records, would allow to ascertain the identity of specimens, and to obtain a more reliable species delimitation. On the whole, the use of 'integrative taxonomy' allowed to reveal the potential presence of crypticity within both genera and demonstrated that species misidentifications due to morphologically similar taxa can often occur and this issue has to be considered with particular attention when performing ecological or biodiversity studies.

Acknowledgements

This work was supported by the European project SEAMoBB: "Solutions for Semi-Automated Monitoring of Benthic Biodiversity" funded by ERA-NET MarTERA. Sheena Conforti was supported by the IMBRSea master program (International Master of Science in Marine Biological Resources), as an Erasmus Mundus Joint Master's Degree (EMJMD) scholarship Awardee. We would like to express our thanks to all the diving centers that supported us during the field activities (Rovinj Sub, Rovinji, Croatia; Palinuro Sub, Palinuro, Italy; Accademia Blu Diving Center, Livorno, Italy. We would also like to thank the CIBM, Centro Interuniversitario di Biologia Marina ed Ecologia Applicata "G.Bacci", Livorno, Italy and the Ruđer Bošković Institute, Rovinj, Croatia, for their support during the sample processing on the field, as well as all the scientists that supported the molecular and morphological analyses in the laboratory.

References

Akaike, H., 1973. Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrika,* 60 (2), 255-265.

Anger, K., 2001. The biology of decapod crustacean larvae

(Crustacean Issues 14). *Aquaculture*, 216 (1-4), 383-387.

Appeltans, W., Ahyong, S.T., Anderson, G., Angel, M.V., Artois, T. *et al.*, 2012. The magnitude of global marine species diversity. *Current Biology*, 22 (23), 2189-2202.

Bate, C.S., 1859. On the importance of an examination of the structure of the integument of Crustacea in the determination of doubtful species - Application to the genus *Galathea*, with the description of a new species of that genus. *Journal of the Proceedings of the Linnean Society. Zoology,* 3, 1-4.

Bayha, K.M., Collins, A.G, Gaffney, P.M., 2017. Multigene phylogeny of the scyphozoan jellyfish family Pelagiidae reveals that the common U.S. Atlantic Sea nettle comprises two distinct species (*Chrysaora quinquecirrha* and *C. chesapeakei*). *PeerJ*, 5, e3863.

Bianchi, C.N., Morri, C., 2000. Marine biodiversity of the Mediterranean Sea: situation, problems and prospects for future research. *Marine Pollution Bulletin*, 40 (5), 367-376.

Bilgin, R., Utkan, M.A., Kalkan, E., Karhan, S.Ü., Bekbolet, M., 2015. DNA barcoding of twelve shrimp species (Crustacea: Decapoda) from Turkish seas reveals cryptic diversity. *Mediterranean Marine Science*, 16 (1), 36-45.

Bishop, J.D.D., Roby, C., Yunnie, A.L.E., Wood, C.A., Lévêque, L. *et al.*, 2013. The Southern Hemisphere ascidian Asterocarpa humilis is unrecognised but widely established in NW France and Great Britain. *Biological Invasions*, 15 (2), 253-260.

Bishop, J.D.D., Wood, C.A., Lévèque, L., Yunnie, A. L.E., Viard, F., 2015. Repeated rapid assessment surveys reveal contrasting trends in occupancy of marinas by non-indigenous species on opposite sides of the western English Channel. *Marine Pollution Bulletin*, 95 (2), 699-706.

Bortolus, A., 2008. Error Cascades in the Biological Sciences: The unwanted consequences of using bad taxonomy in ecology. *AMBIO: A Journal of the Human Environment*, 37 (2), 114-118.

Bouchet, P., 2006. The magnitude of marine biodiversity. p. 31-62. In: *The exploration of marine biodiversity: scientific and technological challenges*. Duarte, C. M. (Ed). Fundación BBVA, Bilbao.

Caputi, L., Andreakis, N., Mastrototaro, F., Cirino, P., Vassillo, M. *et al.,* 2007. Cryptic speciation in a model invertebrate chordate. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (22), 9364-9369.

Chenuil, A., Cahill, A. E., Délémontey, N., Du Salliant du Luc, E., Fanton, H., 2019. Problems and Questions Posed by Cryptic Species. A Framework to Guide Future Studies. p. 77-106. In: *From Assessing to Conserving Biodiversity. History, Philosophy and Theory of the Life Sciences*, vol 24. Casetta, E., Marques da Silva, J., Vecchi, D. (Eds). Springer, Cham.

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2015. GenBank. *Nucleic Acids Research*, 44 (D1), D67-D72.

Colloca, F., Cardinale, M., Belluscio, A., Ardizzone, G., 2003. Pattern of distribution and diversity of demersal assemblages in the central Mediterranean Sea. *Estuarine, Coastal and Shelf Science,* 56 (3-4), 469-480.

Company, J.B., Puig, P., Sardà, F., Palanques, A., Latasa, M. *et al.,* 2008. Climate influence on deep sea populations. *PLoS ONE*, 3 (1), e1431.

Costello, M.J., Wilson, S., Houlding, B., 2012. Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Systematic Biology,* 61 (5), 871-883.

Dayrat, B., 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society,* 85 (3), 407-415.

De Queiroz, K., 2007. Species concepts and species delimitation. *Systematic biology,* 56 (6), 879-886.

Dewalt, R.E., 2011. DNA barcoding: A taxonomic point of view. *Journal of the North American Benthological Society,* 30 (1), 174-181.

D'Udekem-d'Acoz, C., Wirtz, P., 2002. Observations on some interesting coastal Crustacea Decapoda from the Azores, with a key to the genus Eualus Thallwitz, 1892 in the Northeastern Atlantic and the Mediterranean. *Arquipelago. Life and Marine Science*, 19A, 67-84.

Embleton, R., 1836. List of Malacostraca Podophthalmata, found on the coasts of Berwickshire and North Durham, years 1832-1841. *History of the Berwickshire Naturalists' Club,* 1, 69-72.

Ensing, D.J., Moffat, C.E., Pither, J., 2013. Taxonomic identification errors generate misleading ecological niche model predictions of an invasive hawkweed. *Botany,* 91 (3), 137-147.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution,* 39 (4), 783-791.

Fukami, K., Tateno, Y., 1989. On the maximum likelihood method for estimating molecular trees: uniqueness of the likelihood point. *Journal of molecular evolution,* 28 (5), 460-464.

Geller, J., Meyer, C., Parker, M., Hawk, H. *et al.,* 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 (5), 851-861.

Gollasch, S., 2006. Overview on introduced aquatic species in European navigational and adjacent waters. *Helgoland Marine Research,* 60, 84-89.

Goulletquer, P., Gros, P., Bœuf, G., Weber, J., 2014. *Biodiversity in the marine environment.* Springer, Cham, 198 pp.

Guijarro, B., Massutí, E., Moranta, J., Cartes, J.E., 2009. Short spatio-temporal variations in the population dynamics and biology of the deep-water rose shrimp *Parapenaeus longirostris* (Decapoda: Crustacea) in the western Mediterranean. *Scientia Marina,* 73 (1), 183-197.

Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series,*, 41, 95-98.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F. *et al.,* 2008. A global map of human impact on marine ecosystems. *Science,* 319 (5865), 948-952.

Harrington, B., Engelen, J., 2004. Inkscape. *Software available at http://www.Inkscape.org.*

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 (2), 160-174.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences,* 270 (1512), 313-321.

Hillis, D.M., Moritz, C., 1990. Molecular systematics. *Science,* 250 (4983), 1022-1023.

Jukes, T.H., Cantor, C.R., 1969. Evolution of Protein Molecules. p. 21-132. In: *Mammalian Protein Metabolism.* Munro, H. N., (Ed). Academic Press, New York.

Kemp, S., 1925. Notes on Crustacea Decapoda in the Indian Museum. XVII On various Caridea. *Records of the Indian Museum,* 27, 249-342.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* 16 (2), 111-120.

Knowlton, N., 1986. Cryptic and sibling species among the decapod crustacea. *Journal of Crustacean Biology,* 6 (3), 356-363.

Kronenberger, K., Türkay, M., 2003. A population study of *Galathea intermedia* (Crustacea: Decapoda: Anomura) in the German Bight. *Journal of the Marine Biological Association of the United Kingdom,* 83 (1), 133-141.

Krøyer, H., 1841. Fire nye arter af slægten Cuma Edw. *Naturhistorisk Tidsskrift Ser II,* 3(6), 503-535.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, 33 (7), 1870-1874.

Leach, W.E., 1813-1815. Crustaceology. 7(1): 383–384 [1813], 7(2), 385–437, 765–766 [1814], 9(1), Pl. CCXXI [1815]. In: *The Edinburgh Encyclopaedia.* Brewster, D. (Ed). Balfour, Edinburgh.

Lebour, M.V., 1936. Notes on the Plymouth species of Spirontocaris (Crustacea). *Proceedings of the Zoological Society of London,* 106 (1), 89-104, plates 1-7.

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

Leray, M., Knowlton, N., 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences,* 112 (7), 2076-2081.

Liljeborg, O., 1851. Norges Crustacéer. *Ofversigt af Konglige Vetenskaps-Akademiens Förhandlingar,* 8, 19-25.

Linnaeus, C., 1761. Fauna Suecica sistens Animalia Sueciae Regni: Distributa per Classes, Ordines, Genera, Species, cum Differentiis Specierum, Synonymis Auctorum, Nominibus Incolarum, Locis Natalium, Descriptionibus insectorum. *Editio altera, auctior. Stockholmiae, Stockhom, Sweden,* 48, 1-578.

Machordom, A., Macpherson, E., 2004. Rapid radiation and cryptic speciation in squat lobsters of the genus Munida (Crustacea, Decapoda) and related genera in the South West Pacific: molecular and morphological evidence. *Molecular Phylogenetics and Evolution,* 33 (2), 259-279.

Macpherson, E., Robainas-Barcia, A., 2015. Species of the genus *Galathea Fabricius*, 1793 (Crustacea, Decapoda, Galatheidae) from the Indian and Pacific Oceans, with descriptions of 92 new species. *Zootaxa*, 3913 (1), 1-335.

May, R.M., 1992. Bottoms up for the oceans. *Nature* 357 (6376), 278-279.

Meier, R., Shiyang, K., Vaidya, G., Ng, P. K., 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology,* 55 (5), 715-728.

Meiklejohn, K.A., Damaso, N., Robertson, J.M., 2019. Assessment of BOLD and GenBank - Their accuracy and reliability for the identification of biological materials. *PloS ONE,* 14 (6), e0217084.

Milne Edwards, H., 1834-1840. Histoire Naturelle des Crustacés, Comprenant l´Anatomie, la Physiologie et la Classification de ces Animaux. *Encyclopédique Roret, Paris,* (I), i-xxxv + 1-468, (II), 1-532, (III), 1-638, Plates 1–42.

Molinari-Jobin, A., Kéry, M., Marboutin, E., Molinari, P., Koren, I. *et al.*, 2012. Monitoring in the presence of species misidentification: the case of the Eurasian lynx in the Alps. *Animal Conservation,* 15 (3), 266-273.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How many species are there on earth and in the ocean? *PLoS Biology,* 9 (8), 1-8.

Nye, V., Copley, J.T., Linse, K., 2013. A new species of *Eualus* Thallwitz, 1892 and new record of *Lebbeus antarcticus* (Hale, 1941) (Crustacea: Decapoda: Caridea: Hippolytidae) from the Scotia Sea. *Deep Sea Research Part II: Topical Studies in Oceanography,* 92, 145-156.

Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D. *et al.,* 2015. Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology,* 24 (3), 525-544.

Paradis, E., 2010. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics,* 26 (3), 419-420.

Pérez-Barros, P., D'Amato, M.E., Guzmán, N.V., Lovrich, G.A., 2008. Taxonomic status of two South American sympatric squat lobsters, *Munida gregaria* and *Munida subrugosa* (Crustacea: Decapoda: Galatheidae), challenged by DNA sequence information. *Biological Journal of the Linnean Society,* 94 (2), 421-434.

Petinsaari, M., Ratnasingham, S., Miller, S.E., Hebert, P.D.N., 2020. BOLD and GenBank revisited – Do identification errors arise in the lab or in the sequence libraries? *PLOS ONE,* 15 (4), e0231814.

Poggiani, L., 2018. *I crostacei del mare di Gano e del bacino del Metauro*. Fondazione Cassa di Risparmio di Fano, Fano, 287 pp.

Poore, G.C.B., Wilson, G.D.F., 1993. Marine species richness. *Nature,* 361 (6413), 597-598.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology,* 21 (8), 1864-1877.

QGIS Development Team, 2009. *QGIS Geographic Information System*. Open Source Geospatial Foundation.

R Core Team, 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A., Drummond, A.J., 2012. *FigTree version 1.4.0*.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in bayesian phylogenetics using Tracer 1.7. *Systematic Biology,* 67 (5), 901-904.

Ratnasingham, S., Hebert, P.D.N., 2007. The Barcode of Life Data System. *Molecular Ecology Notes*, 77 (3), 355-364.

Raupach, M.J., Barco, A., Steinke, D., Beermann, J., Laakmann, S., 2015. The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. *PLoS ONE,* 10 (9), 1-23.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* 19 (12), 1572-1574.

Rubinoff, D., 2006. Utility of mitochondrial DNA barcodes in species conservation. *Conservation Biology,* 20 (4), 1026-1033.

Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A. E. *et al.,* 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC bioinformatics,* 18 (1), 529.

Schlick-Steiner, B.C., Steiner, F.A., Seifert, B., Stauffer, C., Christian, E. *et al.,* 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual review of entomology,* 55, 421-438.

Schliep, K.P., 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics,* 27 (4), 592-593.

Schliep, K., Potts, A.J., Morrison, D.A., Grimm, G.W., 2017. Intertwining phylogenetic trees and networks. *Methods in Ecology and Evolution,* 8 (10), 1212-1220.

Scorrano, S., Aglieri, G., Boero, F., Dawson, M.N., Piraino, S., 2016. Unmasking *Aurelia* species in the Mediterranean Sea: an integrative morphometric and molecular approach. *Zoological Journal of the Linnean Society,* 180 (2), 243-267.

Silva, J.M., Creer, S., dos Santos, A., Costa, A.C., Cunha, M.R. *et al.,* 2011. Systematic and evolutionary insights derived from mtDNA COI Barcode diversity in the Decapoda (crustacea: Malacostraca). *PLoS ONE,* 6 (5), e19449.

Tarrío, R., Rodrıguez-Trelles, F., Ayala, F. J., 2000. Tree rooting with outgroups when they differ in their nucleotide composition from the ingroup: the drosophila saltans and willistoni groups, a case study. *Molecular Phylogenetics and Evolution,* 16 (3), 344-349.

Tautz, D., Arctander, P., Minelli, A., Thomas, R.H., Vogler, A. P., 2003. A plea for DNA taxonomy. *Trends in Ecology and Evolution,* 18 (2), 70-74.

Thallwitz, J., 1891. Decapoden-Studien, insbesondere basirt auf A.B. Meyer's Sammlungen im Ostindischen Archipel, nebst einer Aufzählung der Decapoden und Stomatopoden des Dresdener Museums. *Abhandlungen und Berichte des Königlichen Zoologischen und Anthropologisch-Etnographischen Museums zu Dresden*, 3 (3), 1-56.

The GIMP Development Team, June 2019. *GIMP*. Version 2.10.12. https: //www.gimp.org (Accessed February 2021).

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., Legendre, F., 2017. Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports,* 7 (1), 1-14.

Valentini, A., Pompanon, F., Taberlet, P., 2009. DNA barcoding for ecologists. *Trends in Ecology and Evolution,* 24 (2), 110-117.

Vaquer-Sunyer, R., Duarte, C.M., 2008. Thresholds of hypoxia for marine biodiversity. *Proceedings of the National Academy of Sciences of the United States of America,* 105 (40), 15452-15457.

Vassily, S., Simakova, U., Griffiths, H., 2017. When rare things are relatively common: caridean shrimps of the South Orkney area and distinct characteristics of the Antarctic species of Eualus (Crustacea: Decapoda: Caridea). p. 91. In: *Book of abstracts: XIIth SCAR Biology Symposium. Leuven, Belgium, 10-14 July 2017*. Anton Van de Putte (Ed.). Scientific Committee on Antarctic Research (SCAR), Cambridge, United Kingdom.

Wiemers, M., Fiedler, K., 2007. Does the DNA barcoding gap exist? - A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology,* 4 (8), 1-16.

Wilson, S.K., Fisher, R., Pratchett, M.S., Graham, N.A.J., Dulvy, N.K. *et al.*, 2008. Exploitation and habitat degradation as agents of change within coral reef fish communities. *Global Change Biology,* 14 (12), 2796-2809.

Windig, J.J., de Kovel, C.G.F., Jong, G., 2004. Genetics and mechanics of plasticity. p. 31-49. In: *Phenotypic Plasticity - Functional and Conceptual Approaches*. DeWitt, T. J., Scheiner, S. M. (Eds). Oxford University Press, New York.

WoRMS Editorial Board, 2021. *World Register of Marine Species (WoRMS)*. https://www.marinespecies.org (Accessed April 2021)

Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L. *et al.,* 2011. Integrative taxonomy, or iterative taxonomy? *Systematic entomology,* 36 (2), 209-217.

Zariquiey Álvarez, R., 1968. Crustáceos Decápodos Ibéricos. *Investigación Pesquera*, 32, 1-510.

Zhang, A.B., He, L.J., Crozier, R.H., Muster, C., Zhu, C-D., 2010. Estimating sample sizes for DNA barcoding. *Molecular Phylogenetics and Evolution,* 54 (3), 1035-1039.

Supplementary Data

The following supplementary information is available online for the article:

Specimens' collection

Fig. S.1: Map illustrating the complete sampling design, comprised of three locations: Livorno (Italy), Palinuro (Italy) and Rovinj (Croatia). Nine ARMS (Autonomous Reef Monitoring Structures) were deployed in each site. The map was produced in QGIS v3.4.1 (QGIS Development Team, 2009).

Morphological identification

Galathea genus

Table S. 1: Morphological characteristics of each Galathea specimen analyzed. Specimen IDs indicate the sampling locations: 'LIV' = Livorno, 'PAL' = Palinuro, 'CRO' = Croatia. Carapace length (cm), shape of the rostrum and presence/absence of epigastric spines are the key morphological traits used for identification. Based on these morphological traits, specimens were assigned to either G. intermedia or G. squamifera (Species assignment).

Fig. S. 2: Pictures of specimens morphologically identified as G. intermedia with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the software GIMP v2.10.6 (The GIMP Development Team, 2019) and Inkscape 0.92.3 (Harrington, 2005).

Fig. S. 3: Pictures of specimens morphologically identified as G. squamifera with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the software GIMP v2.10.6 (The GIMP Development Team, 2019) and Inkscape 0.92.3 (Harrington, 2005).

Fig. S. 4: Boxplot displaying the carapace length of *G. intermedia* (red) and *G. squamifera* (blue). A non-parametric Wilcoxon rank sum test was performed to test for significant differences between the mean carapace length of *G. intermedia* and that of *G. squamifera.* The detected difference appeared to be significant (p-value = 0.00021). Carapace lengths were measured in cm using ImageJ (Rueden *et al*., 2017).

Eualus genus

Fig. S. 5: Pictures of specimens morphologically identified as *Eualus* spp. with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the software GIMP v2.10.6 (The GIMP Development Team, 2019) and Inkscape 0.92.3 (Harrington, 2005).

Phylogenetic analysis

Reconstruction ML trees

ML trees were generated in R (R Core Team, 2018) using the phangorn 2.5.5 package (Schliep, 2011; Schliep *et al.*, 2017). The function ModelTest() in the phangorn 2.5.5 package (Schliep, 2011; Schliep *et al.*, 2017) in R (R Core Team, 2018) was used to find the best model of evolution, and the best-fitting model for the data was chosen based on the Akaike Information Criterion (Akaike, 1973). Distance models chosen for both datasets were optimized using the function optim.pml(), and parameters estimated by ML. In order to improve the topology search, a stochastic rearrangement was chosen when optimizing the phylogenetic trees. To additionally confirm the choice of the model, the difference between likelihoods resulting from using the best fitting model and that having one parameter less was tested computing an Analysis of Variance (ANOVA). Trees were outgroup rooted in both genera to prevent differences between ingroup and outgroup in the substitution processes from making the root unstable (Tarrío *et al.*, 2000). Finally, a bootstrap technique (bootstrap values = 100) was applied to assess the confidence levels of each clade of the observed trees (Felsenstein, 1985). For the *Galathea* genus dataset, the best model of evolution was found to be the Hasegawa-Kishino-Yano+Gamma distribution+Inversion parameter (HKY+G+I) (Hasegawa *et al.*, 1985). The estimated maximum likelihood parameters were the following: discrete gamma categories k = 4; proportion of invariant sites a = 0.52; shape parameter s = 0.69; base frequencies A = 0.34, C = 0.16, G = 0.11, T = 0.39. For the *Eualus* genus dataset, the best model of evolution was HKY+G+I (Hasegawa *et al.*, 1985) as well. The estimated maximum likelihood parameters were the following: discrete gamma categories k = 4; proportion of invariant sites a = 0.45; shape parameter s = 1.02; base frequencies A = 0.28, C = 0.20, G = 0.15, T = 0.37. Trees were edited using the software FigTree v1.4.4 (Rambaut & Drummond, 2012).

Galathea genus

Table S. 3: Table summarizing the mean intraspecific pairwise distance within the *Galathea* species. The number of base differences per site from averaging over all sequence pairs within each group is expressed as a percentage [%]. Standard error estimates too are expressed as a percentage [%] in the last column, and they were estimated using 100 bootstrap replications. The evolutionary analysis involved 72 nucleotide sequences and was conducted in MEGA v.7 (Kumar et al., 2016).

Table S. 4: Table summarizing the mean interspecific pairwise distance between *Galathea* species. The number of base differences per site from averaging over all sequence pairs between groups are shown in decimal form. Standard error estimates, above the diagonal (in grey), are shown in decimal form too, and were estimated using 100 bootstrap replications. The evolutionary analysis involved 72 nucleotide sequences and was conducted in MEGA v.7 (Kumar et al., 2016). * Pisidia longicornis is the outgroup used for phylogenetic reconstructions.

Fig. S. 6: Maximum likelihood (ML) phylogenetic reconstruction of *Galathea* specimens based on COI sequences. The ML method was based on the HKY+G+I substitution model (Hasegawa et al., 1985). The tree was rooted using the outgroup Pisidia longicornis (Linnaeus, 1767). The initial tree for the heuristic search was obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the number of nucleotide differences. A discrete Gamma distribution to model evolutionary rate differences among sites (G = 4 categories) as well as a proportion of invariant sites (I = 0.52) were used. The tree with the highest log likelihood (-3182.3) is shown. The percentage of trees in which the associated taxa clustered together (bootstrap percentage BP_{ML}) is shown next to the branches and they were estimated with 100 bootstrap replications (Felsenstein, 1985). Colors represent species delimitation obtained from ABGD software. The abbreviations next to reference records indicate the sampling locations: DE = German Bight (North Sea), NO = Norway, NS = North Sea, PO = Pacific Ocean, SW = Sweden (Skagerrak), and UK = United Kingdom – North Sea. The analysis involved 72 nucleotide sequences. Evolutionary analysis was conducted in R 3.5.0 (R Core Team, 2018) and in FigTree v1.4.4 (Rambaut and Drummond, 2012).

Eualus genus

Table S. 5: Summary of *Eualus* species and outgroup downloaded from BOLD and GenBank.

Table S. 6: Table summarizing the mean intraspecific pairwise distance within *Eualus* species. The number of base differences per

site from averaging over all sequence pairs within each group are shown as a percentage [%]. Standard error estimates are shown in the last column as a percentage [%] too and were estimated using 100 bootstrap replications. The evolutionary analysis involved 53 nucleotide sequences and was conducted in MEGA v.7 (Kumar *et al.*, 2016). **E. cranchii* and *T. cranchii* are grouped together since they are synonyms and indicate the same species.

Table S. 7: Table summarizing the mean interspecific pairwise distance between *Eualus* species. The number of base differences per site from averaging over all sequence pairs between groups are shown. Specimens records classified under the synonyms *T. cranchii* and *E. cranchii* were clustered together in the same group. Standard error estimates are shown above the diagonal (in grey) and were estimated using 100 bootstrap replications. The evolutionary analysis involved 53 nucleotide sequences and was conducted in MEGA v.7 (Kumar et al., 2016). **H. commensalis* (*Hippolythe commensalis*) is the outgroup used for the phylogenetic reconstruction. ***E. cranchii* and *T. cranchii* are grouped together since both names indicate the same species.

Fig. S. 7: Maximum likelihood (ML) phylogenetic reconstruction of *Eualus* specimens based on COI sequences. The ML method was based on the HKY+G+I model (Hasegawa et al., 1985). The tree was rooted using the outgroup *Hippolyte commensalis* (Kemp, 1925). The initial tree for the heuristic search was obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the number of nucleotide differences. A discrete Gamma distribution to model evolutionary rate differences among sites (G = 4 categories) as well as a proportion of invariant sites (I = 0.45) were used. The tree with the highest log likelihood (-3757.269) is shown. The percentage of trees in which the associated taxa clustered together (bootstrap percentage BPML) is shown next to the branches and were estimated with 100 bootstrap replications (Felsenstein, 1985). Colors represent the species delimitation obtained from the ABGD software. Abbreviation next to reference records indicate sampling locations: AO = Arctic Ocean, BS = Bering Sea, CA = Canada, DE = German Bight (North Sea), IN = Indonesia, PT = Portugal, SP = northern Spain, SW = Sweden (Skagerrak), TU = Turkey, and UK = Anglesey – UK (North Sea). Sequences in red indicate doubtful species assignment on public databases. The analysis involved 53 nucleotide sequences. Evolutionary analyses were conducted in R 3.5.0 (R Core Team, 2018) and in FigTree v1.4.4 (Rambaut and Drummond, 2012). **Thoralus cranchii* is currently accepted as *Eualus cranchii*, but these sequences were available on databases under the name *Thoralus cranchii.*

Maximum likelihood (ML) phylogenetic reconstruction both clearly divided newly sequenced *Eualus* specimens in four distinct clades (Fig. S. 2). ML reconstruction showed bootstrap percentages (BPML) of at least 50% in most clades except for separation of ABGD Group 4 from ABGD Groups 1, 2 and 3 (46%) and for the separation between *E. pusiolus* and the clade including *E. macilentus* (22%) (Fig. S. 3).

ML reconstruction showed bootstrap percentages (BPML) of at least 50% in most clades except for the separation of ABGD Group 4 from ABGD Groups 1, 2 and 3 (46%) and for the separation between *E. pusiolus* and the clade including *E. macilentus* (22%) (Fig. S. 4).