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 and Federica COSTANTINI

Mediterranean Marine Science, 23 (3) 2022

Specimens collection

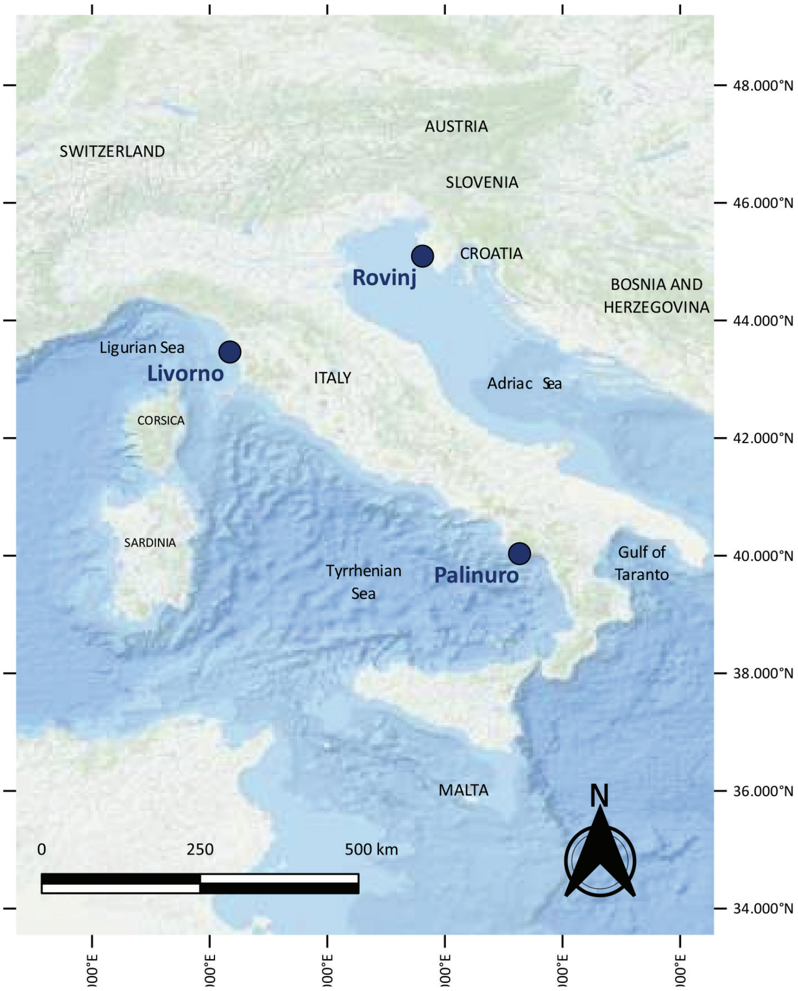

Fig. S. 1: Map illustrating the complete sampling design, comprised of three locations: Livorno (Italy), Palinuro (Italy) and Rovinj (Croatia). Nine ARMS (Autonomous Reef Monitoring Structures) were deployed in each site. The map was produced in QGIS v3.4.1 (QGIS Development Team, 2009).

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).

Sample ID	Carapace length (cm)	Shape of rostrum	Epigastric spines	Species assignation
LIV1	0.434	Long/tight	No	*G. intermedia*
LIV2	0.597	Long/tight	No	*G. intermedia*
LIV3	0.753	Long/tight	No	*G. intermedia*
LIV4	0.627	Long/tight	No	*G. intermedia*
PAL4	0.563	Long/tight	No	*G. intermedia*
PAL8	0584	Long/tight	No	*G. intermedia*
PAL10	0.753	Long/tight	No	*G. intermedia*
PAL12	2.0929	Short/large	Yes	*G. squamifera*
PAL11	1.127	Short/large	Yes	*G. squamifera*
PAL9	1.167	Short/large	Yes	*G. squamifera*
PAL7	1.452	Short/large	Yes	*G. squamifera*
PAL6	2.04	Short/large	Yes	*G. squamifera*
PAL5	1.632	Short/large	Yes	*G. squamifera*
PAL3	2.122	Short/large	Yes	*G. squamifera*
PAL2	1.429	Short/large	Yes	*G. squamifera*
PAL1	1.898	Short/large	Yes	*G. squamifera*
CRO5	2.333	Short/large	Yes	*G. squamifera*
CRO1	2.101	Short/large	Yes	*G. squamifera*
CRO3	2.465	Short/large	Yes	*G. squamifera*
CRO2	2.194	Short/large	Yes	*G. squamifera*
CRO4	1.864	Short/large	Yes	*G. squamifera*
CRO6	1.57	Short/large	Yes	*G. squamifera*
CRO7	2.445	Short/large	Yes	*G. squamifera*

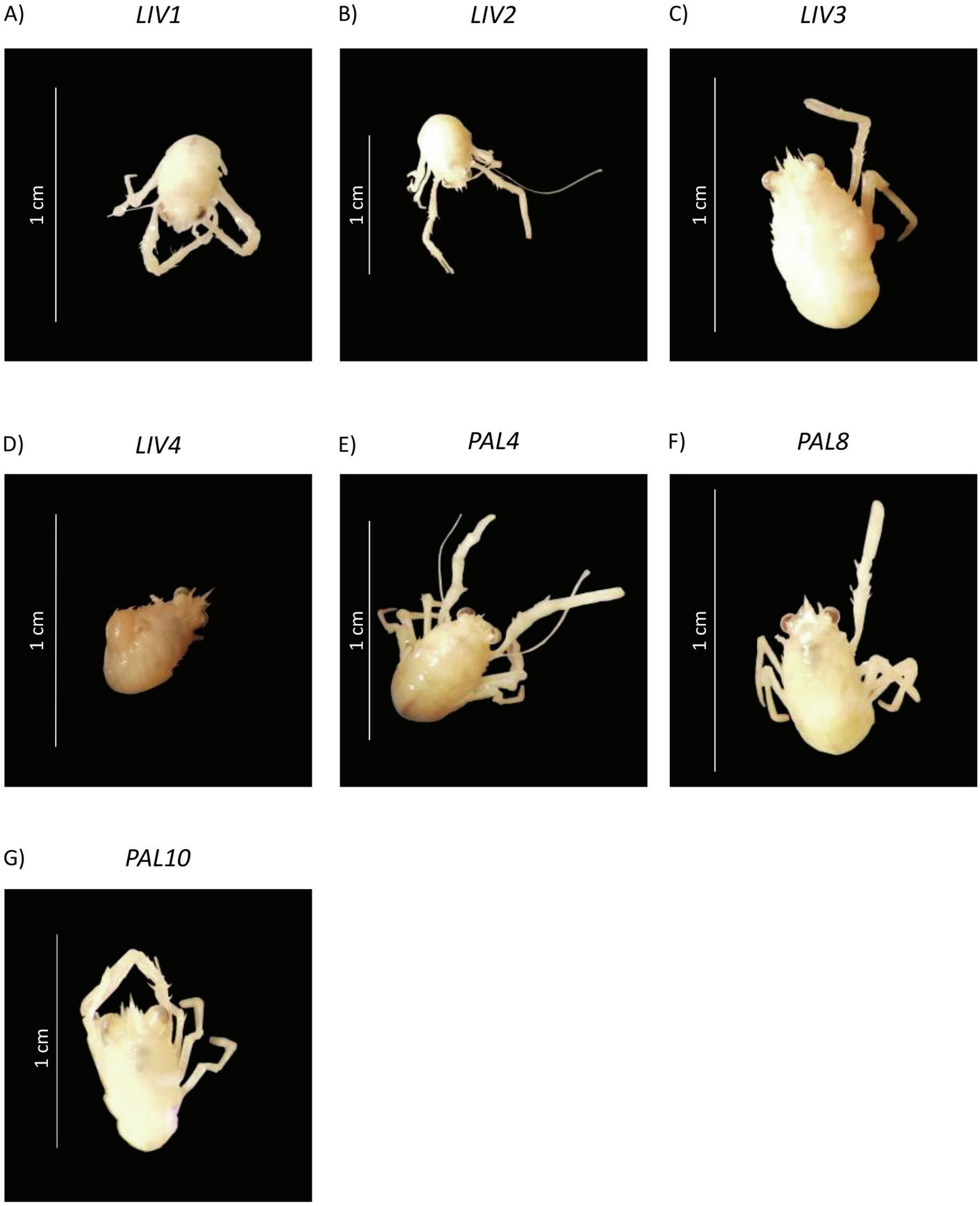

Fig. S. 2: Pictures of specimens morphologically identified as G. intermedia with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the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

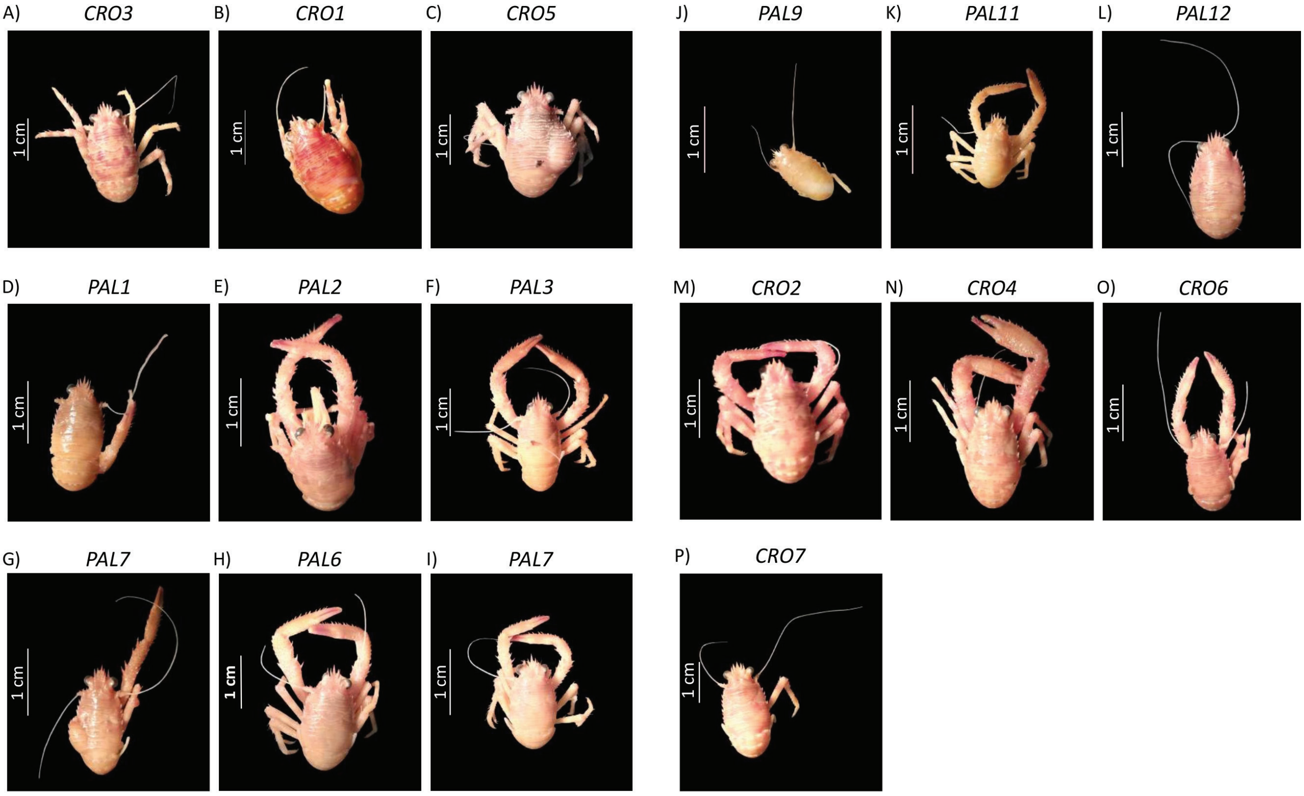

Fig. S. 3: Pictures of specimens morphologically identified as G. squamifera with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the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

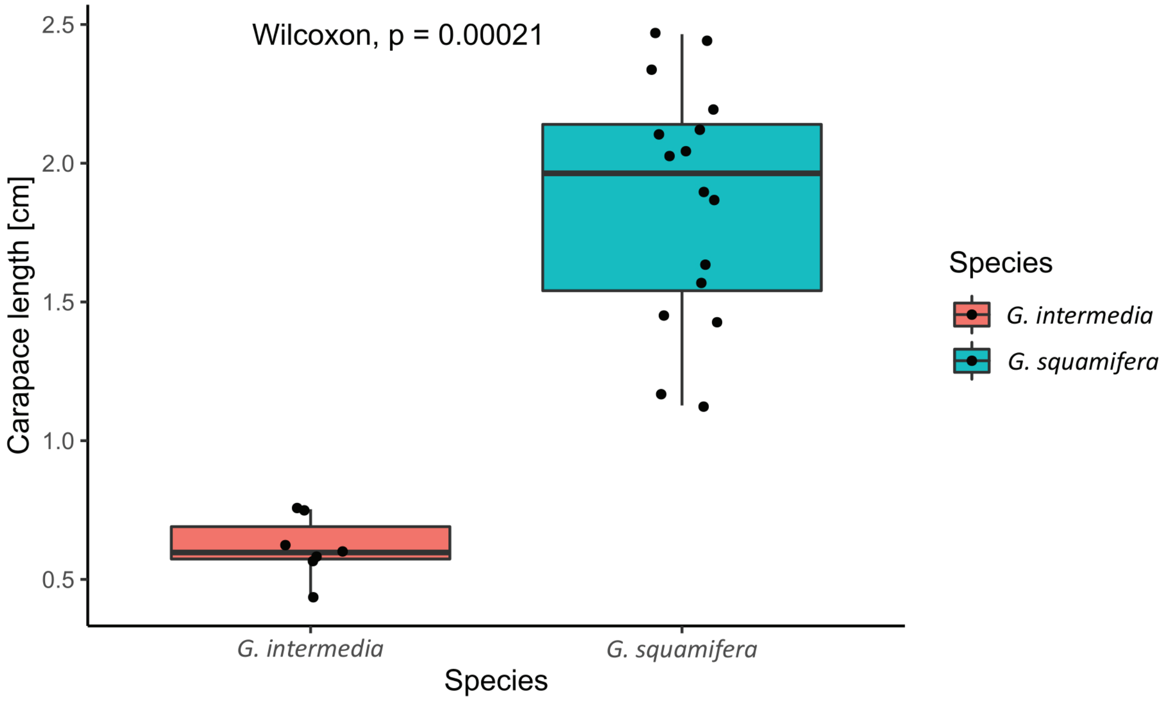

Fig. S. 4: Boxplot displaying the carapace length of *G. intermedia* (red) and *G. squamifera* (blue). A non-parametric Wilcoxon rank sum test was performed to test for significant differences between the mean carapace length of *G. intermedia* and that of *G. squamifera.* The detected difference appeared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

Eualus genus

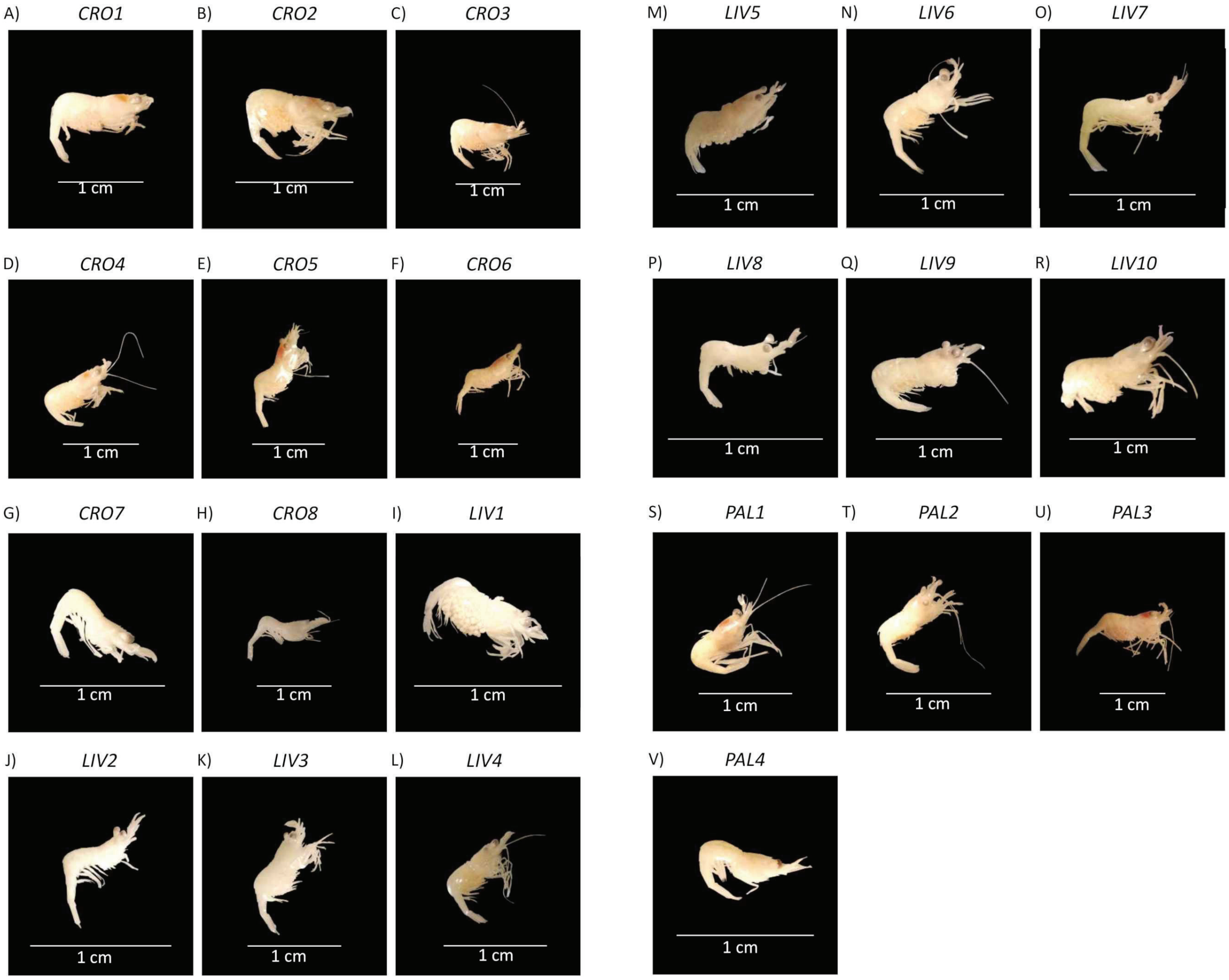

Fig. S. 5: Pictures of specimens morphologically identified as *Eualus* spp. with scale bars corresponding to 1 cm. Photographs of each organism were taken using a Huawei P20 Lite with a 16 MP camera and f/2.2 lens directly from the binocular of the stereomicroscope and subsequently processed using the 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 esti-*

mated 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).

Species	Number of sequences	BOLD IDs	GenBank IDs	Location
G. intermedia	19	BNSC183_10	KT209399	North Sea
		BNSC184_10	KT208435	North Sea
		BNSC509_14	KT208820	North Sea (English Channel)
		BNSDE022_11	KT209259	North Sea (Doggerbank)
		BNSDE023_11	KT209371	North Sea (Doggerbank)
		BNSDE024_11	KT208989	North Sea (Doggerbank)
		BNSDE025_11	KT209451	North Sea (Doggerbank)
		BNSDE143_12	KT208531	North Sea (German Bight)
		BNSDE144_12	KT208699	North Sea (German Bight)
		BNSDE145_12	KT209169	North Sea (German Bight)
		BNSDE146_12	KT208807	North Sea (German Bight)
		BNSDE147_12	KT209218	North Sea (German Bight)
		BNSDE148_12	KT209077	North Sea (German Bight)
		BNSDE235_12	KT208833	North Sea (German Bight)
		BNSDE236_12	KT209422	North Sea (German Bight)
		BNSDE237_12	KT209465	North Sea (German Bight)
		BNSDE238_12	KT208794	North Sea (German Bight)
		BNSDE279_12	KT209491	North Sea
		SWEMA530_15	MG935178	Kattegatt (Sweden)
G. strigosa	3	SWEMA470_15	MG935275	Skagerrak (Sweden)

Continued

Table S3 continued

Species	Number of sequences	BOLD IDs	GenBank IDs	Location
		SWEMA471_15	MG935004	Skagerrak (Sweden)
		na	MK049302.1	Pacific Ocean
G. halia	9	GBCM10958_17	KP203595	Pacific Ocean (Vanuatu)
		GBCM11341_17	KP203603	Pacific Ocean (Vanuatu)
		GBCM5930_17	KP203592	Pacific Ocean (New Caledonia)
		GBCM5931_17	KP203732	Pacific Ocean (Philippines)
		GBCM6293_17	KP203719	Pacific Ocean (New Caledonia)
		GBCM6412_17	KP203596	Pacific Ocean (Solomon Islands)
		GBCM6663_17	KP203604	Pacific Ocean (Vanuatu)
		GBCM6775_17	KP203591	Pacific Ocean (New Caledonia)
		GBCM7511_17	KP203594	Pacific Ocean (Vanuatu)
G. dispersa	12	BNSC142_10	KT209581	North Sea
		BNSC143_10	KT208726	North Sea
		BNSC144_10	KT208594	North Sea
		BNSC145_10	KT209403	North Sea
		BNSDE088_11	KT208548	North Sea
		BNSDE089_11	KT209174	North Sea
		BNSDE091_11	KT209320	North Sea
		BNSDE134_12	KT208618	North Sea
		BNSDE135_12	KT209070	North Sea
		BNSDE181_12	KT208465	North Sea
		JSDUK030_08	JQ306007	North Sea (UK)
		SWEMA574_15	MG935384	Skagerrak (Sweden)
G. nexa	2	SWEMA532_15	MG935130	Skagerrak (Sweden)
		SWEMA579_15	MG935299	Skagerrak (Sweden)
G. squamifera	2	Provided with courtesy by University Museum of Bergen (Norway) and Norwegian Barcode of Life	*na*	Norway

Species	Number of sequences	BOLD IDs	GenBank IDs	Location
Pisidia longicornis (outgroup)	2	BNSDE165_12	KT208496	North Sea (German Bight)
		BNSC177_10	KT209482	North Sea

Table S. 2: Summary of *Galathea* species and outgroup downloaded from BOLD and GenBank.

Species	Mean p-distance [%]	Std. Error [%]
G. squamifera	**0.00**	**0.00**
G. nexa	**0.36**	**0.21**
G. dispersa	**0.85**	**0.23**
G. halia	**1.31**	**0.26**
G. strigosa	**0.84**	**0.32**
G. intermedia	**0.04**	**0.03**
Group 1	**0.93**	**0.18**
Group 2	**0.22**	**0.16**
Group 3	**2.54**	**0.46**
Group 1 and *2*	4.29	0.48
G. squamifera and *Group 1*	0.95	0.17

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.

	P. longicornis	*G. squamifera*	*G. nexa*	*G. dispersa*	*G. halia*	*G. strigosa*	*G. intermedia*	*Group 1*	*Group 2*	*Group 3*
P. longicornis		0.017	0.019	0.018	0.019	0.016	0.017	0.017	0.017	0.018
G. squamifera	0.191		0.015	0.015	0.018	0.014	0.015	0.004	0.012	0.017
G. nexa	0.196	0.138		0.013	0.014	0.014	0.014	0.015	0.015	0.015
G. dispersa	0.210	0.145	0.122		0.017	0.016	0.015	0.015	0.014	0.016
G. halia	0.211	0.156	0.157	0.183		0.016	0.017	0.017	0.016	0.017
G. strigosa	0.168	0.124	0.123	0.140	0.157		0.015	0.013	0.014	0.015
G. intermedia	0.223	0.171	0.153	0.147	0.190	0.146		0.015	0.015	0.014
Group 1	0.195	0.010	0.142	0.149	0.154	0.127	0.173		0.011	0.016
Group 2	0.194	0.085	0.155	0.149	0.152	0.156	0.187	0.082		0.015
Group 3	0.216	0.175	0.154	0.168	0.179	0.166	0.110	0.178	0.186	

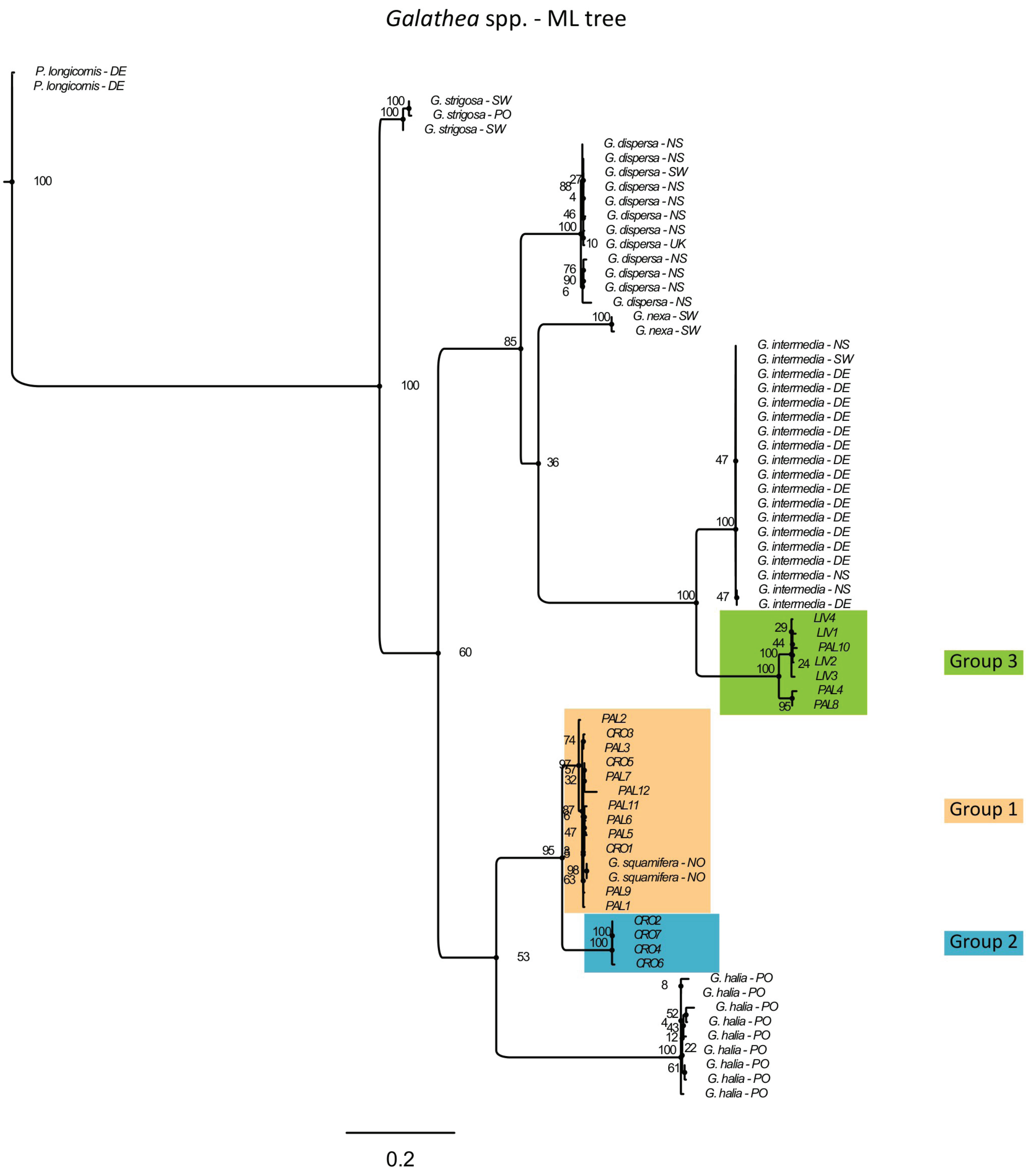

Fig. S. 6: Maximum likelihood (ML) phylogenetic reconstruction of *Galathea* specimens based on COI sequences. The ML method was based on the HKY+G+I substitution model (Hasegawa et al., 1985). The tree was rooted using the outgroup Pisidia longicornis (Linnaeus, 1767). The initial tree for the heuristic 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.

Species	Number of sequences	BOLD IDs	GenBank IDs	Location
E. occultus	2	SWEMA526_15	*na*	Skagerrak (Sweden)
		BCASB018_16	*na*	northern Spain
E. cranchii	6	BNSC322-11	KT209433	North Sea (German Bight)
		BNSC325-11	KT209005	North Sea (German Bight)
		JSDUK183-08	JQ306048	Anglesey (UK)
		JSDUK184-08	JQ306049	Anglesey (UK)
		JSDUK185-08	JQ306050	Anglesey (UK)
		MLALE062-12	KF369190	Alentejo Litoral (Portugal)
T. cranchii (syn. *E. cranchii*)	3	*na*	KJ841675.1	Turkey
		na	MG935123.1	Kattegatt (Sweden)
		na	KF369190.1	Portugal
E. pusiolus	4	BNSDE248-12	KT209346	North Sea (German Bight)
		BNSDE249-12	KT209527	North Sea (German Bight)
		BNSDE250-12	KT209086	North Sea (German Bight)
		SWEMA676-15	MG935120	Skagerrak (Sweden)
E. gaimardi	5	ARBEN110-18	*na*	Arctic Ocean
		ARBEN111-18	*na*	Arctic Ocean
		ARBEN112-18	*na*	Arctic Ocean
		BEAUF093-11	*na*	Arctic Ocean
		BEAUF094-11	*na*	Arctic Ocean
E. macilentus	5	ASGRD112-18	*na*	Bering Sea
		ASGRD113-18	*na*	Bering Sea
		ASGRD114-18	*na*	Bering Sea
		WW088-07	FJ581633	**Gulf of Saint Lawrence (Canada)**
		WW089-07	FJ581632	Newfoundland (Canada)
E. fabricii	5	ARCMI428-14	MG319594	Nunavut (Canada)
		ARCMI449-14	MG315523	Nunavut (Canada)
		ARCMI450-14	MG319600	Nunavut (Canada)
		ARCMI459-14	MG316293	Newfoundland (Canada)
		ARCMI465-14	MG311791	Newfoundland (Canada)
Hippolyte commensalis (outgroup)	1	GBCMD21537-19	KX589001	Indonesia

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.

Species	Mean p-distance [%]	Std. Error [%]
E. fabricii	0.07	0.07
E. macilentus	0.55	0.24
E. gaimardii	0.48	0.19
E. occultus	1.21	0.42
E. pusiolus	9.52	0.80
E. cranchii*	13.66	0.91
Group 1	0.20	0.08
Group 2	1.03	0.29
Group 3	2.53	0.37
Group 4	1.03	0.27

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.

	*H. commensalis**	*E. fabricii*	*E. macilentus*	*E. gaimardii*	*E. occultus*	*E. cranchii***	*E. pusiolus*	*Group 1*	*Group 2*	*Group 3*	*Group 4*
*H. commensalis**		**0.230**	**0.242**	**0.233**	**0.018**	**0.015**	**0.016**	**0.019**	**0.017**	**0.016**	**0.018**
E. fabricii	**0.019**		**0.014**	**0.013**	**0.015**	**0.014**	**0.014**	**0.014**	**0.015**	**0.017**	**0.015**
E. macilentus	**0.018**	**0.186**		**0.015**	**0.014**	**0.012**	**0.013**	**0.015**	**0.013**	**0.016**	**0.014**
E. gaimardii	**0.019**	**0.156**	**0.207**		**0.015**	**0.013**	**0.015**	**0.014**	**0.015**	**0.015**	**0.015**
E. occultus	**0.231**	**0.192**	**0.173**	**0.191**		**0.008**	**0.011**	**0.197**	**0.186**	**0.178**	**0.018**
*E. cranchii***	**0.227**	**0.203**	**0.186**	**0.198**	**0.107**		**0.010**	**0.179**	**0.156**	**0.126**	**0.106**
E. pusiolus	**0.226**	**0.194**	**0.185**	**0.185**	**0.148**	**0.159**		**0.203**	**0.208**	**0.199**	**0.148**
Group 1	**0.242**	**0.193**	**0.180**	**0.230**	**0.017**	**0.013**	**0.014**		**0.014**	**0.015**	**0.017**
Group 2	**0.249**	**0.211**	**0.197**	**0.207**	**0.016**	**0.012**	**0.014**	**0.170**		**0.015**	**0.017**
Group 3	**0.231**	**0.224**	**0.206**	**0.221**	**0.016**	**0.009**	**0.015**	**0.163**	**0.158**		**0.016**
Group 4	**0.227**	**0.198**	**0.169**	**0.193**	**0.004**	**0.008**	**0.012**	**0.193**	**0.188**	**0.176**	

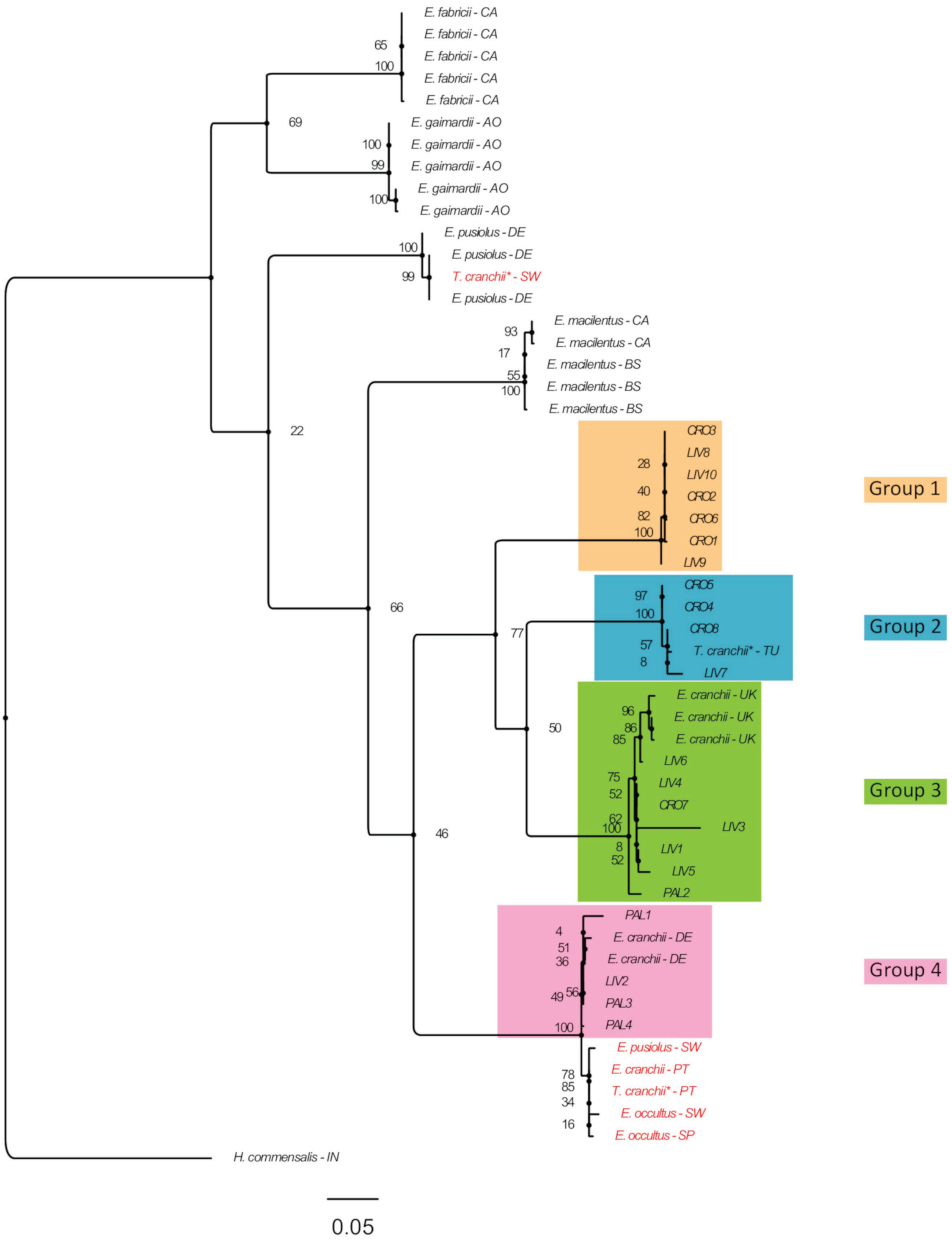

Fig. S. 7: Maximum likelihood (ML) phylogenetic reconstruction of *Eualus* specimens based on COI sequences. The ML method was based on the HKY+G+I model (Hasegawa et al., 1985). The tree was rooted using the outgroup *Hippolyte commensalis* (Kemp, 1925). The initial tree for the heuristic search was 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.*