RESEARCH PAPER
https://doi.org/10.1071/IS23018

Invertebrate Systematics

Novel molecular resources for single-specimen barcoding of enigmatic crustacean y-larvae

Niklas Dreyer[A,B,C,D,J], Jørgen Olesen[A,*], Mark J. Grygier[E], Danny Eibye-Jacobsen[A], Alexandra S. Savchenko[F], Yoshihisa Fujita[G], Gregory A. Kolbasov[H], Ryuji J. Machida[B], Benny K. K. Chan[B,*] and Ferran Palero[I,*]

For full list of author affiliations and declarations see end of paper

***Correspondence to:**
Jørgen Olesen
Natural History Museum of Denmark, University of Copenhagen, Denmark
Email: jolesen@snm.ku.dk

Benny K. K. Chan
Biodiversity Research Center, Academia Sinica, Taipei, Taiwan
Email: chankk@gate.sinica.edu.tw

Ferran Palero
Institut Cavanilles de Biodiversitat i Biologia Evolutiva (ICBIBE), University of Valencia, Paterna, Spain
Email: ferran.palero@uv.es

Handling Editor:
Shane Ahyong

Received: 25 April 2023
Accepted: 8 January 2024
Published: 8 March 2024

Cite this: Dreyer N *et al.* (2024) Novel molecular resources for single-specimen barcoding of enigmatic crustacean y-larvae. *Invertebrate Systematics* **38**, IS23018. doi:10.1071/IS23018

ABSTRACT

Despite discovery more than 100 years ago and documented global occurrence from shallow waters to the deep sea, the life cycle of the enigmatic crustacean y-larvae isincompletely understood and adult forms remain unknown. To date, only 2 of the 17 formally described species, all based on larval stages, have been investigated using an integrative taxonomic approach. This approach provided descriptions of the morphology of the naupliar and cyprid stages, and made use of exuvial voucher material and DNA barcodes. To improve our knowledge about the evolutionary history and ecological importance of y-larvae, we developed a novel protocol that maximises the amount of morpho-ecological and molecular data that can be harvested from single larval specimens. This includes single-specimen DNA barcoding and daily imaging of y-nauplii reared in culture dishes, mounting of the last naupliar exuviae on a slide as a reference voucher, live imaging of the y-cyprid instar that follows, and fixation, DNA extraction, amplification and sequencing of the y-cyprid specimen. Through development and testing of a suite of new primers for both nuclear and mitochondrial protein-coding and ribosomal genes, we showcase how new sequence data can be used to estimate the phylogeny of Facetotecta. We expect that our novel procedure will help to unravel the complex systematics of y-larvae and show how these fascinating larval forms have evolved. Moreover, we posit that our protocols should work on larval specimens from a diverse array of moulting marine invertebrate taxa.

Keywords: barcoding, biodiversity, Facetotecta, larval barcoding, molecular resources, new primers, systematics, y-larvae.

Introduction

Increased availability of diverse phylogenomic resources has transformed pancrustacean (Hexapoda and Crustacea) systematics, yet numerous clades remain recalcitrant to analysis due to a critical lack of molecular resources (Regier *et al.* 2010; von Reumont *et al.* 2012; Oakley *et al.* 2013; Schwentner *et al.* 2017; Lozano-Fernández *et al.* 2019; Bernot *et al.* 2022, 2023). This particularly applies to the so-called 'dark taxa' (Hartop *et al.* 2022), of which the enigmatic crustacean y-larvae (Pancrustacea: Thecostraca: Facetotecta) constitute one of the most remarkable examples.

Y-larvae are fascinating invertebrates exclusively known from larval specimens and stages (Glenner *et al.* 2008; Dreyer *et al.* 2023*a*). These larvae develop through a series of dispersive stages (y-nauplii: Itô 1986; Kolbasov *et al.* 2021; Olesen *et al.* 2022; Dreyer *et al.* 2023*a*) and a single putative attachment stage (y-cyprid: Kolbasov *et al.* 2022; Dreyer *et al.* 2023*a*). A subsequent stage called the ypsigon, which lacks segmented appendages, a gut and compound eyes, has been induced through *in vivo* exposure of y-cyprids to the crustacean moulting hormone (Glenner *et al.* 2008; Pérez-Losada *et al.* 2009; Dreyer *et al.* 2023*a*). Having shed the y-cyprid's cuticle, the ypsigon is tentatively regarded as an early instar or juvenile of a hypothetical endoparasitic adult stage (Pérez-Losada *et al.* 2009; Dreyer *et al.* 2023*a*). The y-larval life cycle thus envisaged is reminiscent of that

observed in parasitic barnacles (Cirripedia: Rhizocephala; Pérez-Losada *et al.* 2009; Dreyer *et al.* 2023*a*).

Few studies have informed the evolutionary history of Facetotecta, despite discovery more than a century ago and global distribution from surface waters to 6000 m (Hansen 1899; Grygier 1987; Kolbasov *et al.* 2022; Dreyer *et al.* 2023*a*, 2023*b*). Morphological studies based on wild-caught specimens have proven unsuccessful in delimiting species within geographically widespread 'types' such as type IV (Hansen 1899; Schram 1972) and 'Pacific Type I' (Itô 1986) that are each practically identical wherever found. Formal cladistic analyses using morphological characters have not resolved the facetotectan phylogeny (Pérez-Losada *et al.* 2009; Kolbasov *et al.* 2022), and the need for further DNA sequences of live-imaged and properly vouchered material is inescapable (Olesen *et al.* 2022; Dreyer *et al.* 2023*b*; Olesen and Grygier, in press). Phylogenetic analyses show that y-larvae form a distinct and monophyletic group (Grygier 1987; Chan *et al.* 2021; Dreyer *et al.* 2022*a*, 2023*b*) but conflicting datasets currently place Facetotecta as sister either to Ascothoracida (Petrunina *et al.* 2014; Dreyer *et al.* 2022*a*) or all other thecostracans (Ascothoracida and Cirripedia; see Pérez-Losada *et al.* 2009). Previous studies have mostly used unvouchered 'ghost' sequences of y-larvae or only a few relatively conservative molecular markers (Pérez-Losada *et al.* 2009; Petrunina *et al.* 2014; Dreyer *et al.* 2022*a*), which is problematic for an array of reasons (Olesen *et al.* 2022; Dreyer *et al.* 2023*a*, 2023*b*). For example, no facetotectan species description has been supplemented by both nuclear and mitochondrial sequence data, and a complete or near-complete series of naupliar and cyprid instars has been described for only three species to date (*viz*. *Hansenocaris furcifera* Itô, 1989, *H. itoi* Kolbasov & Høeg, 2003, and *H. demodex* Olesen, Dreyer, Palero & Grygier, 2022). The incompletely known life cycle of y-larvae, their small sizes and a dearth of molecular data have hampered an accurate assessment of their evolution and systematics.

The phylogenetic resolution of y-larvae has until recently been unsatisfactory and problematic for several reasons. Considering that the local and global diversity may be much greater than that described to date (Hansen 1899; Glenner *et al.* 2008; Dreyer *et al.* 2023*a*, 2023*b*), sequences of vouchered and named species are essential for biodiversity inventories. Currently, only 17 species are formally described, all within the single genus *Hansenocaris* Itô, 1985 (Olesen and Grygier 2022; Olesen *et al.* 2022; Dreyer *et al.* 2023*a*). Solving phylogenetic relationships is, moreover, crucial for robust macroevolutionary modelling of life histories and morpho-ecological traits, especially when compelling life-history evidence suggests that y-larvae attain maturity as endoparasites in unknown hosts (Glenner *et al.* 2008; Pérez-Losada *et al.* 2009; Dreyer *et al.* 2023*a*, 2023*b*).

Destructive DNA-extraction methods often result in complete digestion of the specimen, preventing adequate morphological species delimitation and museum storage of vouchers. Here, we present novel molecular resources and optimised protocols for successful DNA extraction, voucher cuticle retention and PCR-amplification of single-specimen y-larvae of less than 500 µm in size. This is intended to supplement the single-specimen rearing and live-imaging protocol previously developed by our team (Olesen *et al.* 2022, 2024). Through extensive laboratory experimentation, we developed an optimal DNA-extraction and PCR-amplification protocol for Facetotecta and compared this to two other methods that can also yield single-specimen nucleotide sequences. To this end, and as part of a larger campaign investigating the phylogeny and evolution of Facetotecta (Olesen *et al.* 2022; Dreyer *et al.* 2023*b*), we designed novel oligonucleotide primers to amplify nuclear and mitochondrial genes, and tested 28 primer pairs in more than 2000 PCR reactions. We hereby show that up to 6700 aligned basepairs can be obtained from a single y-specimen. Using material obtained from Pacific, north-east Atlantic, Antarctic (NCBI data) and Arctic waters, we demonstrate and confirm the wide utility of our protocol by showcasing a phylogeny estimate of Facetotecta that complements previous analyses. Finally, we discuss the importance of our protocols for further advancing knowledge of y-larval systematics and evolution.

Materials and methods

Collections, rearing, imaging and digital sorting

Approximately 11 000 y-larvae and 100 *Baccalaureus* nauplius larvae representing Ascothoracida were collected with 75-µm-mesh plankton nets from surface waters (0–5 m) at three sites: the pier outside of the University of the Ryukyus Tropical Biosphere Research Center Sesoko Laboratory on Sesoko Island, Okinawa, Japan (JA); Gongguan Harbour on Green Island (Green Island Marine Research Station, Academia Sinica, Taiwan, TA); and Wanghaixiang Fishing Harbour in Keelung, Taiwan (TA) (Olesen *et al.* 2022; Fig. 1, Supplementary Table S1). In total, 35 of these y-larvae and 1 *Baccalaureus* nauplius specimen were sequenced and included in this work. Supplementary y-larva material came from Tioman Island (Malaysia; MAL), Piscinas do Pesquerio in Ponta Delgada, Azores, Portugal (AZ, $n = 10$) and waters off the White Sea Marine Biological Station, Russian Federation (RU; $n = 14$) between 2017 and 2021. All larvae were sorted from plankton samples with glass Pasteur pipettes under dissection microscopes and either (1) photographed and video-recorded while alive with a Nikon ECLIPSE 80i compound microscope (JA), an Olympus IX70 inverted compound microscope (TA) or a Zeiss AX10 light microscope (TA), both of the latter being equipped with Nomarski (DIC) optics and a Canon EOS 5D Mark IV digital camera, or (2) bulk-fixed in 95% ethanol for

subsequent molecular work (TA, RU, AZ). All images were sorted digitally in Corel Draw to identify groups of morphologically similar specimens (i.e. morphospecies; Olesen *et al.* 2022; Dreyer *et al.* 2023*b*) and at least two or three specimens from each such group were used for Sanger sequencing.

Primer design, DNA extraction, PCR and sequencing

Reference nucleotide sequences of selected loci for Ascothoracida and Cirripedia were downloaded from NCBI GenBank and aligned with in-house (see below) or NCBI Facetotecta sequences using MAFFT (ver. 7.450 (--auto --maxiterate 1000), see https://mafft.cbrc.jp/alignment/software/; Katoh and Standley 2013). The alignments were imported, visualised and used to design primers in Geneious Prime (ver. 2022.0.2, see https://www.geneious.com/; Kearse *et al.* 2012). To enhance specificity and amplification success, we strived to design longer (>20 bp) primers with GC-contents between 30 and 60%, a GC-clamp of 2 (3′ terminating in at least one G/C to promote optimal binding), a melting temperature within the range of 55–68°C and no homopolymeric regions of >4 bases (e.g. TGGGGG). We tested a total of 28 primer pairs (23 newly designed) and performed 2056 PCR reactions using these. Tables 1 and 2 (and Supplementary Table S2) provide an overview of the primers, sequences and amplification success using our single-specimen approach.

Our goal was to develop and test the efficacy of fast 'filter-free' DNA-extraction methods that amplify single-specimen y-larvae while retaining the 'exuviae' as vouchers. (N.B. We use 'exuviae' with quotation marks to refer to cuticular remains after tissue digestion during DNA extraction and without quotation marks to refer to shed cuticles resulting from successful larval moults.) To this end and as a part of a larger phylogenetic and barcoding campaign, we extracted DNA from 421 single y-larva specimens, of which nine were from the Azores, 326 from Sesoko Island (Japan), 4 from Malaysia, 105 from Green Island and Keelung (Taiwan) and 17 from the White Sea (Russian Federation). Using this material, we tested and compared DNA extraction using the GeneReleaser kit (BioVentures, Murfreesboro, TN, USA) that has previously proven to be an efficient protocol for retrieving DNA from zooplankton (Schizas *et al.* 1997; Böttger-Schnack and Machida 2011; Watanabe *et al.* 2016; Olesen *et al.* 2022) and a simplified method using a subset of the reagents in the DNeasy Blood and Tissue kit (QIAGEN, CA, USA). DNA extracts were kept cold in a 'freezing' Eppendorf tube rack during preparation of the PCR assay. We pipetted 12.2 µL of filtered ddH_2O, 5 µL of 'hot start' HOT FirePol polymerase master mix (Solis BioDyne, Tartu, Estonia), 0.4 µL of each oligonucleotide primer (usually ordered as 'desalted') and 3 µL of gDNA template to each vial. For multiplex PCR reactions that use two or more primer pairs in combination, we subtracted the equivalent amount of added primer from the amount of ddH_2O.

Table 1. Amplification success of all primers with single larval specimens using Protocols 1 and 2.

Locus	*n* tested	*n* amplified	Amplification success (%)
COX1	495	296	59.79
18S	589	515	87.43
28S	426	361	84.74
12S	139	73	52.51
16S	300	168	56
H3	107	70	65.42
Total	**2056**	**1483**	**67.65**

Statistics are summarised from both the GeneReleaser and SimplifiedDNeasy extraction methods. Bold numbers indicate the sum of tested and amplified loci as well as the average amplification success (%).

Table 2. Amplification success of preferred 'legacy' loci for Facetotecta phylogenetics.

Locus	Primers	*n* tested	*n* amplified	Amplification success (%)
COX1	Flnew/R1 + mlCOIInF/jgHCO2198	126	117	92.85
18S	Face1F/Face1R	132	110	83.33
18S	Face2F/Face2R	109	107	98.16
18S	Face1F/Face2R	76	65	85.52
28S	Face3F/Face3R	154	145	94.15
28S	83F/875R	111	82	73.87
12S	3F/281R	12	10	83.33
16S	137F/548R	181	115	63.53
16S	SF/SR	68	41	60.29
H3	AF/AR	96	67	69.79
Total		**1412**	**1163**	**79.71**

Statistics are summarised from both the GeneReleaser and SimplifiedDNeasy extraction methods.

To increase primer annealing sensitivity, specificity to complex templates (e.g. >60% G-C content) and yield we generally used a Touchdown PCR profile under the following conditions: initial denaturation for 15 min at 95°C; followed by 10 cycles of denaturation at 95°C for 1 min, annealing at T_m (i.e. melting temperature) +10°C decreasing by 1°C per cycle for 30 s, and extension at 72°C for 1 kb per 60 s; 30 cycles of denaturation at 95°C for 1 min, annealing at T_m −2°C for 30 s and extension at 72°C for 1 kb per 60 s; and final extension at 72°C for 10 min followed by 20 min at 4°C to halt the reaction. Primer details including sequences and annealing temperatures are listed in Supplementary Table S1.

All PCR reactions were conducted in a DNA Engine Thermal Cycler (Bio-Rad, Richmond, CA, USA) and PCR products were visualised in agarose gel with varying concentrations depending on fragment size. DNA sequencing was performed by Genomics BioSci & Tech Ltd (New Taipei City, Taiwan) and Macrogen Europe BV (Amsterdam, Netherlands).

We sequenced templates of three ribosomal (r) loci (mitochondrial *16S*, nuclear *18S* and *28S* rRNA) and two protein-coding genes (mitochondrial *COX1* and nuclear histone-3, *H3*) of 74 specimens and added new sequences to NCBI GenBank (Tables 1 and 2, and Supplementary Table S2). PCR amplicons that yielded clear gel electrophoresis bands and clean chromatograms that aligned with our in-house database when sequenced were defined as 'successful amplifications'. Less than 10% of amplicons were not sequenced but nonetheless yielded clear bands and these were also counted as successful amplicons.

Protocol 1. Complete information of live-imaged specimens with preserved voucher exuviae

Culturing and morphological data

Plankton-collected y-larvae were sorted into plastic Petri dishes of 35-mm diameter and reared further in mini-cultures of up to five morphologically distinct types or forms as described in Olesen *et al.* (2022). Some of these were followed alive daily in culture by video under a compound microscope (LM), living specimens being transferred between the Petri dishes and depression slides in a glass Pasteur pipette. The final naupliar exuviae of the developing specimens were mounted in glycerin jelly on glass slides (GL; Grygier *et al.* 2019) or (although only attempted for a few specimens) SEM stubs for high-resolution morphological investigations (Grygier *et al.* 2019).

DNA extraction and PCR

Specimens were subjected to DNA extraction using both the GeneReleaser and DNeasy kits. For GeneReleaser, we follow a slightly modified protocol from that outlined in Schizas *et al.* (1997). In total, 1 µL of 10× PCR buffer and 9 µL of ddH_2O were mixed in a 250-µL Eppendorph tube, to which a single y-larva was added. Protein denaturation was initiated by incubating the tubes at 94–95°C for 2 min and the specimens were subsequently transferred to ice. We added 1 µL of protein kinase K (QIAGEN, Chatsworth, CA, USA) and vortexed the samples thoroughly before incubating them at 55°C for 15 min and 70°C for 10 min. Subsequently, we added 10 µL of GeneReleaser (BioVentures, Inc., Murfreesboro, TN, USA) and finally incubated the samples under the following thermal program: 65°C for 15 s, 8°C for 15 s, 65°C for 45 s, 97°C for 90 s, 8°C for 30 s, 65°C for 90 s, 97°C for 30 s, 65°C for 30 s, 80°C for 3 min and lastly 4°C for 10 min to halt the reaction. The tubes were centrifuged for 1 min, after which 15 µL of supernatant and 10 µL of AE-buffer (QIAGEN, Chatsworth, CA, USA) were transferred to fresh Eppendorf tubes. To locate and retrieve vouchers' 'exuviae', samples were spun down and inspected under a dissection microscope. Any 'exuviae' found were recovered and mounted as described below.

For the simplified DNeasy method, we transferred individual y-larval specimens from vials into small Petri dishes. After gentle swirling of a dish by hand, the specimen was often easily located at the centre of the dish and transferred to a 250-µL Eppendorf tube in a 0.5–1.0-µL droplet of ethanol using a 2.5-µL pipette fitted with a sterile filter tip. The vial(s), each containing a single y-larva specimen, were incubated for 2 min at 36°C to evaporate the remaining ethanol. Volumes of 40 µL of AE buffer and 4 µL of Protease K (Qiagen, DNeasy kit) were subsequently added to each vial, which were then incubated for 1 h at 56°C followed by enzyme inactivation at 72°C for 12 min. The vials were spun down and placed in a rack. The 'exuviae', each located at the bottom of a vial, were removed in a 1-µL droplet using a 2.5-µL pipette fitted with a filter tip and either placed directly in preheated glycerin jelly on a glass slide or in a droplet in a Petri dish for later mounting. A note was written on the lid above the droplet indicating the specimen's ID code. We usually performed PCR of 'legacy' markers (as described above; *12S*, *16S*, *18S* and *28S* rRNA, *COX1* and *H3*) before mounting the exuviae to avoid potential contamination prior to PCR of these standard markers. Using the pipette tip or an eyelash taped to a stick, y-naupliar 'exuviae' (those of the last y-nauplius or an earlier stage) were rotated so the dorsal sides faced upward and y-cyprids were placed laterally. We found intact 'exuviae' in vials that had been stored at −18°C for up to 3 years, suggesting that continuous freezing does not affect the quality of prospective vouchers.

We also tested the non-destructive protocol provided by Cornils (2015), which uses a slightly modified extraction method of the QIAamp DNA Mini Kit, including the use of ATL-buffer. This protocol consistently worked for a range of copepod species but for y-larvae yielded faint bands, fragmented 'exuviae' and noisy chromatograms (i.e. low quality DNA template amplicons).

Vials were sequenced with the PCR-identifier, complete specimen-designation and gene listed sequentially (e.g. 1-JA-2019-001-Hansenocaris-demodex_COX1, 2-JA-2019-100-Hansenocaris-demodex_COX1).

Protocol 2. Assorted information of specimens with or without voucher 'exuviae'

Culturing and morphological data

Although admittedly artificial, we decided to lump all protocols that provide different degrees of morphological information into one category. Protocol 2 therefore differs from Protocol 1 in the retention of notably fewer data layers. Protocol 2 was applied to the three live-imaged specimens of *H. demodex* from Green Island (TA) (reported in Olesen *et al.* 2022), the specimens from the Azores (AZ) and specimens of *H. itoi* from the White Sea (WS). In the latter species, only non-sequenced specimens have been live-imaged. For *H. itoi,* all other data layers exist, including confocal images (see Fig. 1*a*).

DNA extraction and PCR

We tested three extraction procedures: that used exclusively for Protocol 1, the DNeasy kit (as described in detail above) following the manufacturer's specifications and the GeneReleaser kit (as described in detail above). PCR reactions were performed as described above.

Protocol 3. Anecdotal information on non-imaged, non-vouchered specimens

Culturing and morphological data

This protocol differs from the other two in the lack of any image data for the specimens, whether living or dead. Information about single or bulk-sequenced specimens is therefore anecdotal at best. We did not apply this method but this has been used by others to generate the published y-larva sequences of 'Facetotecta spp.' and *H. itoi* (Pérez-Losada *et al.* 2002, 2009; Gallego *et al.* 2015) in NCBI GenBank. These authors used different DNA extraction kits but successfully amplified long and short ribosomal (*16S*), protein-coding mitochondrial (*COX1*) and nuclear (*18S*, *28S* and *H3*) markers.

Phylogeny estimation and barcoding of individual y-larvae

Chromatograms of new forward and reverse sequences were trimmed and assembled in Geneious Prime. Protein-coding genes were translated (invertebrate mitochondrial code for *COX1* and universal code for *H3*) to proteins, aligned and manually checked for stop codons. Using MAFFT (--auto --maxiterate 1000), nucleotides of each locus and specimen were aligned with outgroup sequences from Ascothoracida and previously sequenced material of Facetotecta (Pérez-Losada *et al.* 2009). Alignments for each locus were initially inspected by eye and subsequently curated by assessing ambiguities with the less stringent options in GBLOCKS (ver. 0.91b, see http://phylogeny.lirmm.fr/phylo_cgi/one_task.cgi?task_type=gblocks; Castresana 2000) ($-t = p - b1 = 24 - b2 = 39 - b3 = 4 - b4 = 10 - b5 = n - b6 = y$).

The nucleotide alignments were concatenated in Geneious Prime (6730 bp before and 3415 bp after GBLOCKS). The concatenated alignment was partitioned (each codon position in protein-coding genes) and used to estimate the best-fit substitution model with ModelFinder (Kalyaanamoorthy *et al.* 2017); partitions were allowed to be merged if this increased model fit (-m MFP + MERGE). Maximum Likelihood (ML) phylogenetic trees were built with IQ-TREE multicore (ver. 2.0.3, see http://www.iqtree.org/; Minh *et al.* 2020), and node support was assessed with 10 000 ultrafast bootstrap (UFBoot) replicates and 10 000 replicates of the SH-like approximate likelihood ratio test (SH-aLRT; --alrt 10 000). To alleviate issues inherent to the UFBoot approximation method, we specified the '-bnni' option. We allowed each partition to have a unique substitution rate (-spp) and performed a more thorough nearest-neighbour interchange search (-allnni). We also estimated a phylogeny with Bayesian inference using the concatenated dataset in MrBayes (ver. 3.2, see https://github.com/NBISweden/MrBayes/; Ronquist *et al.* 2012). We estimated posterior probabilities through six independent runs with four Markov chains for 10 000 000 generations and with sampling every 1000 generations. We used the 'invgamma' model as the lset rate and computed the potential scale reduction factor with the 'sump' command. We estimated a consensus phylogenetic tree with the first 25% of the iterations discarded as burn-in.

The unrooted tree estimates were visualised in Geneious Prime, rooted in Ascothoracida and further annotated in Corel Draw with morphological information from live and 'exuvial' specimens. These illustrations were adjusted individually for brightness, contrast and background colour.

Results

Loci selection, DNA extraction and primer efficacy

We tested 28 primer-pair combinations by performing 2056 individual PCR reactions and provided 74 new sequences (73 Facetotecta, 1 Ascothoracida) that have been deposited in NCBI GenBank with accession numbers PP336629–PP336655 (*28S* rRNA), PP336577–PP336602 (*18S* rRNA), PP336558–PP336576 (*16S* rRNA) and PP337082 (*COX1*) (Supplementary Table S1).

We also tested the efficacy and amplification success of the two DNA-extraction protocols (Fig. 2*a–c*, Table 1, Supplementary Table S2). We performed 815 PCR reactions with the GeneReleaser kit and 1241 reactions with the simplified DNeasy kit (Supplementary Table S2). Minimum and maximum amplification rates varied between 0% (DNeasy method, 16S 1471F/1427R, COI_Palerointernal, both $n = 6$; Supplementary Table S2) and ~98% (Dneasy method, 18Sface2F/2R, $n = 109$; Supplementary Table S2). Both DNA extraction protocols yielded nearly identical overall amplification rates (~68%; Fig. 2*a*, *c*, Supplementary

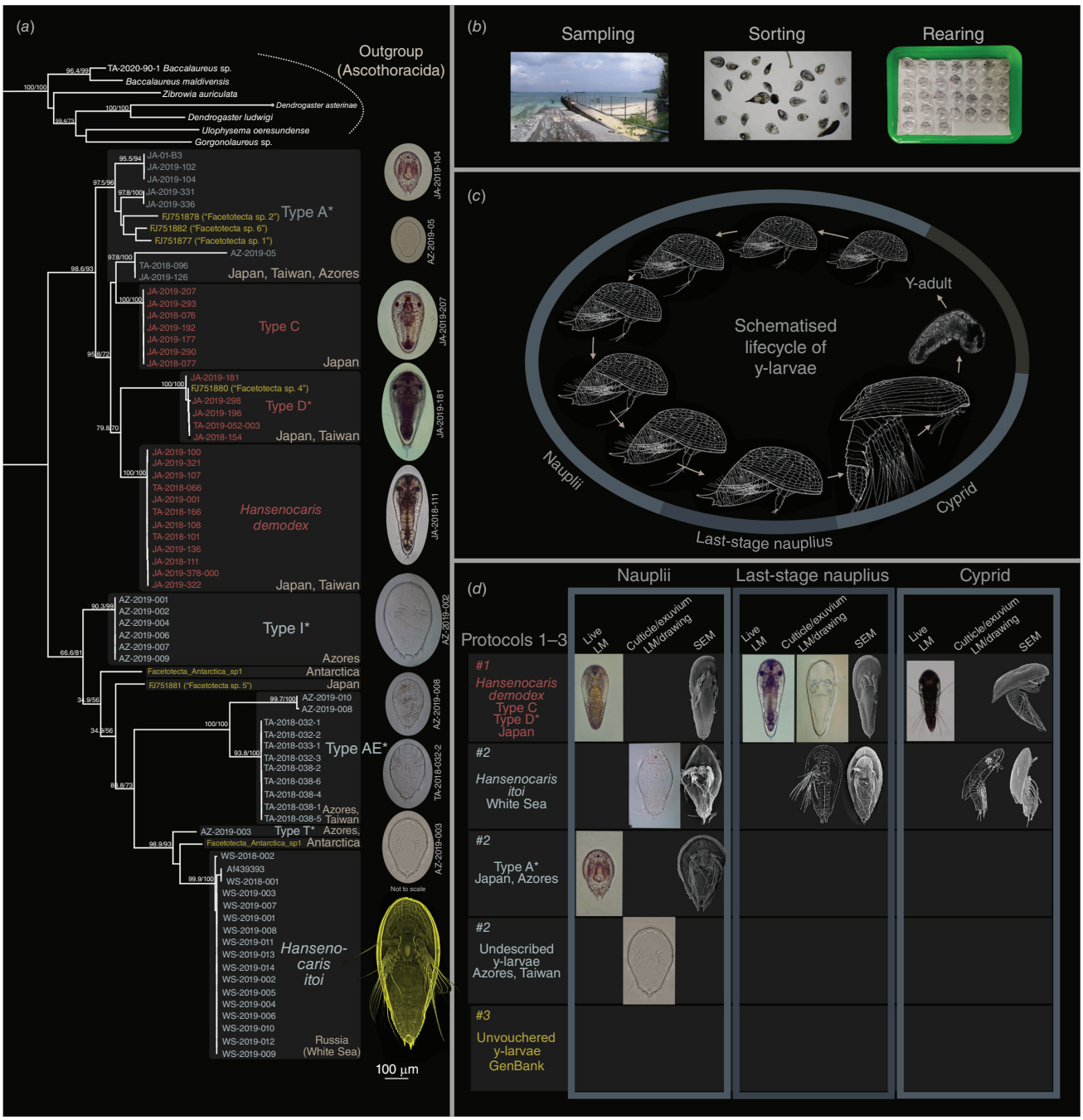

Fig. 1. Multilocus phylogeny estimation for Facetotecta with single-specimen resolution. (*a*) Maximum-likelihood tree from an alignment spanning 6730 bp and 6 loci with Ascothoracida selected as outgroup, showing relationships of Facetotecta and division into at least two major clades. (*b*) Putative life cycle of crustacean y-larvae consisting of seven y-nauplius stages, one y-cyprid stage, one or more ypsigons and as yet unknown adults. (*c*) Examples of morphological information extracted from Protocols 1, 2 and 3, represented by branch colours (red, grey and yellow respectively). Protocol 2 lumps different subprotocols that lack live image data, illustrated by three examples.

Table S2). The GeneReleaser protocol yielded at least ~42% and at most ~90% amplification success across loci (Fig. 2; Supplementary Table S2). Extraction and PCR amplification with the simplified DNeasy method showed lower variance, yielding a slightly higher minimum (~47%) and slightly lower maximum (~88%) amplification success (Fig. 2*a*, Supplementary Table S2). These results may be biased because different primer combinations were tested with the two protocols. Combining the data from the two extraction methods, we computed an overall amplification success rate of 68% across all primer combinations (2096 successfully amplified fragments; Table 1, Supplementary Table S2, Fig. 2*a*).

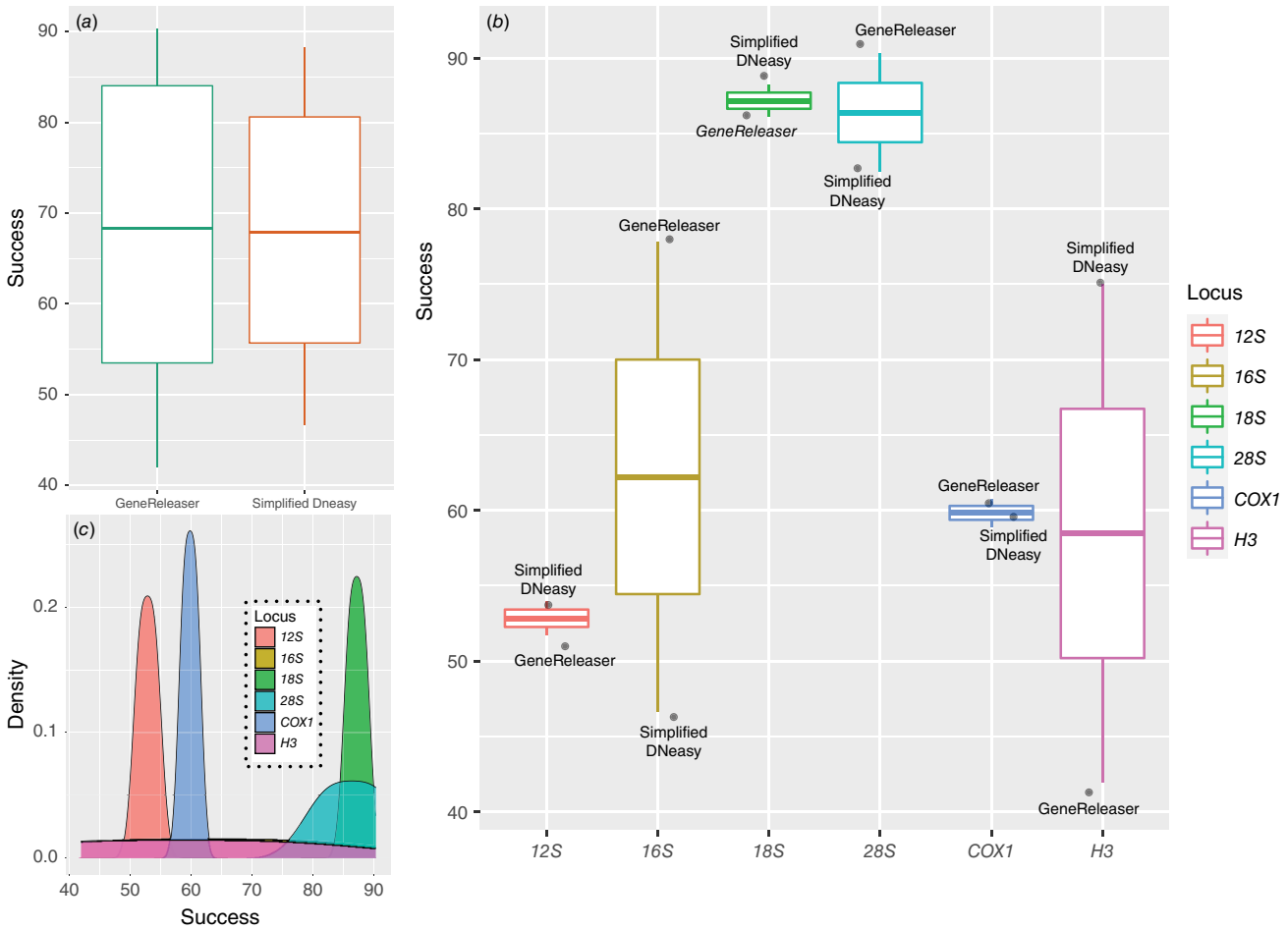

Fig. 2. Comparison of primer amplification efficacy with two filter-free DNA extraction protocols for single-larva systematics. (*a*) Amplification success (%) of all primers using GeneReleaser and Simplified DNeasy extraction kits. (*b*) Amplification success (%) of all primers partitioned to locus using GeneReleaser and the Simplified DNeasy extraction kits. The large difference for *16S* and *H3* may be caused by the different primers used for each extraction protocol (Supplementary Table S1). (*c*) Density plot of amplification success (%) of all primers for both extraction kits combined, showing uniformly higher amplification rates for *12S*, *18S* and *COX1*, and uniformly lower amplification rates for *16S* and *H3*.

Average (mean) successful amplification rates were ~60% for *COX1* (296 of 495 reactions; Fig. 2, Table 1, Supplementary Table S2), ~87% for *18S* (515 of 589 reactions; Fig. 2*c*, Table 1, Supplementary Table S2), ~85% for *28S* (361 of 426 reactions; Fig. 2*c*, Table 1, Supplementary Table S2), ~52% for *12S* (73 of 139 reactions; Fig. 2*c*, Table 1, Supplementary Table S2) and 56% for *16S* (168 of 300 reactions; Fig. 2*c*; Table 1, Supplementary Table S2). The amplification rate difference (variance) between the two extraction methods was notably higher for *16S* and *H3* than any other locus (Fig. 2*b*, *c*, Table 1, Supplementary Table S2), with *12S*, *COX1* and *18S* yielding more similar amplification success rates between the two methods (Fig. 2*b*, *c*, Table 1, Supplementary Table S2). For all primer combinations, *18S* and *28S* expectedly had the highest amplification rates but we did not observe this pattern for *H3* (~40% with GeneReleaser and ~75% with the simplified DNeasy method; Table 1, Supplementary Table S2).

Based on the highest observed primer efficacy (i.e. amplification success), overlap with publicly available primers, fragment lengths obtained and quality of sequences (e.g. the presence of clear and distinct chromatogram peaks), we recommend sequencing the following six 'legacy' loci in future investigations of facetotectan systematics. ***COX1***: F1new and R1 with the internal 'Leray'-segment mlCOIInF/jgHCO2198 (~93% amplification success; ~750 bp; Leray *et al.* 2013; Fig. 3, Table 2, Supplementary Table S2). ***18S***: initially with Face1F/2R (~85% success; ~1800 bp; Fig. 3, Table 2, Supplementary Table S2); subsequently, depending on the amplification rate of the primer, Face1F/1R (~83% success; ~300 bp; Fig. 2, Table 2, Supplementary Table S2) and Face 2F/2R (~98% success; ~300 bp; Fig. 3, Table 2, Supplementary Table S2). ***28S***: 83F/875R

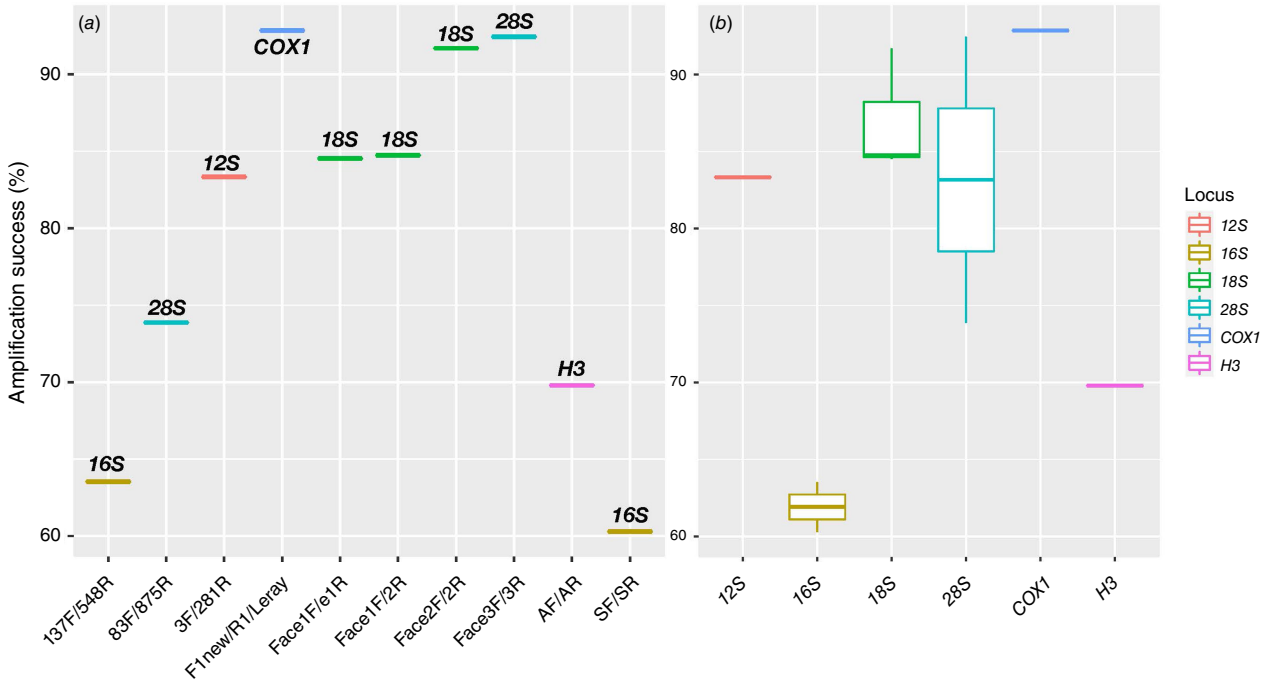

Fig. 3. Amplification success (%) and efficacy of 'legacy' primers. (*a*) Combined amplification success (%) using both GeneReleaser and Simplified DNeasy extraction kits. (*b*) Summary of (*a*) but with amplification success (%) partitioned to locus, hence the larger difference for *28S* and *18S* primers.

(~74% success; ~880 bp; Fig. 3, Table 2, Supplementary Table S2) with the internal primer Face3F/3R (~94% success; ~350 bp; Fig. 3, Table 2, Supplementary Table S2). ***12S***: 3F/281R (~83% success; ~300 bp; Fig. 3, Table 2, Supplementary Table S2). ***16S***: SF/SR (~60% success; 700 bp; Tsang *et al.* 2009; Fig. 3, Table 2, Supplementary Table S2) with 137F/548R (~50% success; ~400 bp; Fig. 3, Table 2, Supplementary Table S2). ***H3***: AF/AR (~70% success; Colgan *et al.* 1998; ~330 bp; Fig. 2, Table 2, Supplementary Table S2).

Morphospecies assignment and phylogeny estimation

Based on LM images we recovered two named species and six undescribed morphospecies (*Hansenocaris demodex* from Green Island, Taiwan and Sesoko Island, Okinawa; *H. itoi* from the White Sea, Russian Federation; Type A* from Green Island, Sesoko Island and the Azores; Type C from Sesoko Island; Type D* from Wanghaixiang Harbour, Taiwan and Sesoko Island; Type I* (note that the Latin letter Type I* is distinguished from both Roman numeral Type I *sensu* Hansen and later authors, and the Pacific Type I *sensu* Itô and later authors) from the Azores; and Type AE* from Taiwan), with morphospecies designations following Dreyer *et al.* (2023*b*). We added four specimens that could not be assigned to morphospecies prior to sequencing due to either a lack of images or voucher material (Facetotecta sp. 1 from Antarctica and Facetotecta sp. 5 from Sesoko Island; Fig. 1) to this.

The final alignment contained 81 specimens (of which 7 are non-imaged and voucherless 'Facetotecta sp.' sequences), 6730 bp, 1680 distinct patterns, 1443 parsimony-informative sites, 384 singleton sites and 4903 constant sites (Table 3). Following the best substitution-model fit inferred by ModelFinder, we merged the partitions as follows: TNe + R2 (*18S* rRNA; BIC-score 14 930.890, logL: −7446.021), TIM2 + F + R2 (*H3*, all codons; BIC-score 6798.637, logL: −3373.182), TN + F + I + G4 (*28S* rRNA; BIC-score 12 621.319; logL −6280.758), TPM2 + F + I (*12S* rRNA; 3435.374, logL −1693.851) and GTR + F + G4 (*16S* rRNA, *COX1*, all codons; BIC-score 11 902.261, logL: −5915.054) (Table 3). We recovered two clades (I and II), although the monophyly of the second clade was not well-supported (Clade I: 98.6% Ultrafast Bootstrap (UFBoot), 93% SH-aLRT support; Clade II: 66.6% UFBoot, 81% SH-aLRT; Fig. 1).

In Clade I the planktotrophic morphospecies A* is paraphyletic (97.5% UFBoot/96% SH-aLRT). Three Type A* specimens from the Pacific (TA-2018-096, JA-2019-126) and the Azores (AZ-2019-5) are nested within a highly supported, monophyletic subclade comprising pigmented and yolky lecithotrophs (95.8% UFBoot/71% SH-aLRT support). The relationship between Types C, D* and *H. demodex* remains unresolved in our dataset (Fig. 1; 69.2% UFBoot/32% SH-aLRT).

In our dataset, all Clade II specimens and morphospecies are planktotrophic, although the feeding types of Facetotecta sp. 5 and Antarctic species 1 and 2 are unknown. All multi-specimen morphospecies in the clade are monophyletic with

Table 3. Alignment and nucleotide substitution models selected for phylogeny estimation after GBLOCKS.

Locus	n	Length (bp)	Informative (bp)	Invariable (bp)	Model
12S	21	404	122	250	TVM + F + I + G4
16S	52	321	160	138	TVM + F + I + G4
18Sface1	46	1383	299	1037	Tne + R2
18Sface2	70	240	36	196	K2P + I + G4
28S	72	648	76	552	K2P + I + G4
COX1	33	419	173	214	TVM + F + I + G4
H3	25	260	73	169	TIM2 + F + I + I + R2
Total	**294**	**3415**	**866**	**2387**	

Bold numbers indicate sum of fragment lengths, informative sites and invariable sites. bp, base pairs.

high support, including an unnamed morphospecies from the Azores ('Unnamed y-nauplius'; AZ-2019-1, 2, 4, 6, 7, 9; 90.3% UFBoot/99% SH-aLRT), Type AE* (100 UFBoot/100% SH-aLRT) and *H. itoi* (99.9% UFBoot/100% SH-aLRT).

Discussion

The enigmatic facetotectans (commonly referred to as y-larvae) are a severely understudied crustacean subclass, and the notable dearth of molecular resources and published morphological voucher data concerning them has impeded assessments of their diversity, ecosystem functions and evolutionary history (Pérez-Losada *et al.* 2002, 2009; Kolbasov *et al.* 2022; Dreyer *et al.* 2023*a*). As with any diverse animal group, such assessments cannot be reliably made based on unvouchered, unphotographed specimens such as those of Pérez-Losada *et al.* (2002, 2009) and Gallego *et al.* (2015). To facilitate rigorous yet easy single-specimen barcoding and phylogenetics, we developed a cost-effective, rapid voucher-preserving protocol that capitalises on the widely used DNeasy extraction kit. Our aim was to develop a rearing and extraction protocol that maximises morphological and molecular information for single y-larval specimens spanning what is emerging as a complex, wide phylogenetic range (Olesen *et al.* 2022; Kolbasov *et al.* 2022; Dreyer *et al.* 2023*b*; data herein). By using a suite of light-microscopic images of hundreds of living larvae (Olesen *et al.* 2022; Olesen and Grygier, in press) and by developing primers targeting both nuclear and mitochondrial genomic markers, we demonstrate a way, based on larval instars, to infer the diversity and phylogenetic relationships of a crustacean taxon with a complicated and partly unknown life cycle. We believe that our protocols should also work if applied to other moulting invertebrate larvae including larger thoracican barnacle larvae (also see Chen *et al.* 2013).

In the following we discuss the three central topics related to the development of these new molecular resources: (1) the efficacy and usefulness of the new primers and the two new DNA extraction methods; (2) how the three protocols and associated differing yields of morphological resolution may affect taxonomic treatments and future classifications; and (3) interesting features of the preliminary phylogeny, for example the molecular and biogeographical extent of morphospecies, and the use of such information to build a classification that must currently rely entirely on larval stages.

Efficacy of DNA extraction protocols and primers

We tested two different DNA extraction methods using a series of novel and published nuclear and mitochondrial primer pairs. Although the success rate or primer amplification varied between 0 and ~98% (Supplementary Table S2), the average amplification success of our 'legacy' primers (Fig. 3, Table 3, Supplementary Table S1) was ~80%. The lowest amplification rates overall were observed for legacy primers *16S* (average ~62%) and *H3* (average ~60%), likely caused by either mismatched priming sites or low copy numbers for individual y-larvae. The highest amplification rates were notably not restricted to nuclear loci; amplification using the *COX1* 'Folmer' (modified F1 and the old R1) and 'Leray' primers also yielded ~93% success. In combination with the high amplification rates of the nuclear ribosomal markers, this offers great promise for future robust species delimitation and phylogeny estimation.

DNA extraction relied on either the GeneReleaser kit that had previously been used to retrieve high-quality DNA templates from minute zooplankton specimens (Schizas *et al.* 1997; Böttger-Schnack and Machida 2011; Watanabe *et al.* 2016), including y-larvae (Olesen *et al.* 2022), or a simplified yet highly effective modification of the QIAGEN DNeasy Blood and Tissue kit. There are important advantages and disadvantages related to both protocols. GeneReleaser yields higher amplification rates for certain loci but the kit only yields ~15–20 µL of DNA extract and the thermal incubation program passes through high temperatures (>90°C). This may have a severe impact on the successful retrieval of

voucher 'exuviae' because (1) in initial trials with various crustacean larvae (data not shown), incubation at such temperatures almost always resulted in the disintegration of the 'exuviae' (i.e. only digested pieces of cuticle were recovered) and (2) these exuviae may easily become trapped in the GeneReleaser powder grains when the tubes are spun down after incubation. Less than 5% of the vials contained 'exuviae' after transferring the GeneReleaser supernatant to new vials. Moreover, the GeneReleaser kit is more expensive (US$125 for a kit on 9 November 2023) and does not have buffers and protease K that have to be procured elsewhere. By contrast, the simplified DNeasy kit contains all relevant products for rapid, efficient DNA extraction and produces bright gel-bands in the expected size ranges, clean chromatograms and high rates of both overall and locus-specific amplification (Fig. 2 and 3, and Supplementary Table S2). For these reasons, we recommend using our extraction method with the simplified DNeasy kit and legacy primers. We are currently designing a series of primers for nuclear protein-coding genes to expand this approach to higher-level phylogenetics.

Comparing protocols and paving the road for comprehensive single-specimen barcoding

We specifically designed our protocols using established 'standard' extraction and PCR methods. As there is no published DNA-based reference database for y-larvae, an essential first step was to ensure primer region overlap with those of published sequences. We therefore designed primers that target the most common regions used for invertebrate barcoding and phylogenetics.

Apart from providing the second multi-locus and vouchered phylogeny estimate for Facetotecta, our preliminary phylogeny of this group (Fig. 1) demonstrates the advantages and disadvantages of the three protocols that we outlined above. Protocol 1 outperformed all previous attempts to study the molecular phylogenetics and evolution of Facetotecta. This protocol retained the information obtained from living specimens (y-nauplii, y-cyprid) throughout larval development and voucher 'exuviae' of single specimens after DNA extraction.

Protocol 1 (Fig. 1) allowed for consistent amplification of long (>1500 bp) and short (<400 bp) DNA templates of protein-coding and ribosomal loci of both the mitochondrial and nuclear genomes (Tables 1, 2, 3, Supplementary Table S2). Protocol 1 therefore ensured maximal gathering of morphological and molecular information from individual specimens and allowed the scoring of both y-naupliar and y-cyprid characters. Given the remarkably high local diversity at, e.g., Sesoko Island (Glenner *et al.* 2008; Olesen *et al.* 2022; Dreyer *et al.* 2023*b*), this protocol will surely become an important asset for understanding the overall diversity of Facetotecta and has already proven invaluable for species descriptions when dealing with sympatric distributions of multiple species (Olesen *et al.* 2022). The last-stage nauplius (LSN) is an unambiguously homologous, easily recognisable stage, of which there is never more than one for a given taxon (Olesen *et al.* 2022). The exuvium of the LSN (shed during metamorphosis to the y-cyprid) can serve as either a primary or complementary voucher specimen, or even as part of a type series, whether this is mounted on a glass slide or fixed and preserved in a liquid storage medium.

For example, the integrative description of *Hansenocaris demodex* was based on a series of nearly identical specimens that were initially reared and live-imaged, and subsequently amplified and sequenced individually with a short *18S* fragment (Olesen *et al.* 2022). Crucial details of the y-cyprid as observed by both LM and SEM were added to this, the latter admittedly not relying on sequenced material but specimens confidently assigned to the species due to similarities of the preceding nauplii (in particular the LSN) to those of sequenced individuals. In future work, several data layers including live images, LM images of voucher exuviae (of y-nauplii), SEM images (of y-cyprids and some LSN voucher exuviae, either naturally molted or 'exuviae' remaining after DNA extraction) and DNA sequence data, will likely be harvested from a single specimen (Grygier *et al.* 2019). LSN specimens may also be imaged by confocal laser-scanning microscopy (CLSM), although we have not attempted this for DNA-extracted specimens yet (for information on mounting exuviae for CLSM and SEM, see Grygier *et al.* 2019; Kolbasov *et al.* 2022; Olesen *et al.* 2022, 2024; Dreyer *et al.* 2023*a*).

Protocol 2 lumps a motley set of culturing and sequencing strategies together, all of which produce a less informative and less complete assortment of visual information than Protocol 1. This protocol is typically applied to y-larvae with planktotrophic nauplii, such as y-naupliar Types A* and AE*, and the planktotrophic y-nauplii from the Azores and the White Sea, as these cannot yet be reared in the lab. The morphological information thus obtained may be sufficient to assign specimens to morphospecies established with molecular support but usually does not provide a comprehensive basis for species descriptions, except at localities where only one species occurs (e.g. *H. itoi* in the White Sea). At localities such as Sesoko Island with many sympatric morphospecies, different instars in the life cycle (i.e. y-nauplii and y-cyprids) cannot easily be linked together, rendering this protocol insufficient for species descriptions. We have nonetheless been able to demonstrate that Protocol 2 works on other Thecostraca larvae, e.g. an ascothoracid larva identified as *Baccalaureus* sp. nests as a sister species to *B. maldivensis* (Fig. 1). Therefore, we anticipate that Protocols 1 and 2 should work on other marine invertebrate larvae that moult and leave exuviae.

Protocol 3 involves the use of non-cultured y-larvae and has been used for all y-larvae with sequences currently deposited in GenBank except for those of Olesen *et al.*

(2022). This protocol does not produce any images and thus no vouchered data, and is therefore impractical for all levels of y-larval phylogenetics. At best, anecdotal information about the specimens may be obtained prior to sequencing. Despite these limitations, our work has allowed us to determine the identity of a least some of the y-larval material ('Facetotecta sp. 1–6') used by Pérez-Losada *et al.* (2009), as some matches are apparent among the vouchered material sequenced herein from the same locality (Sesoko Island). For example, in Pérez-Losada *et al.* (2009), 'Facetotecta spp. 1, 2 and 6' cluster within our Type A* and 'Facetotecta sp. 4' is molecularly identical to our Type D*. However, 'Facetotecta sp. 5' (Pérez-Losada *et al.* 2009) and the two sequences of Gallego *et al.* (2015) from Antarctica remain as 'ghost' sequences, unlinked to known, imaged Facetotecta. These specimens, although unvouchered and lacking any kind of published morphological information, have played a critical role at higher taxonomic levels, helping to systematically place Facetotecta within Thecostraca and Pancrustacea (Pérez-Losada *et al.* 2002, 2009; Petrunina *et al.* 2014).

Olesen *et al.* (2022) also used a single-specimen culturing approach for y-larvae. Our study expands and refines this protocol significantly. Olesen *et al.* (2022) did not use the culturing approach in conjunction with exclusively non-destructive DNA extraction, nor was a comparison of available extraction and amplification methods provided. The resulting phylogeny rests entirely on a short, conservative nuclear fragment of *18S* rRNA. This fragment can distinguish morphospecies but can likely not provide the same degree of species-level resolution as the *COX1* fragment. Our new *COX1* primer pairs yielded a slightly higher amplification rate than that of the *18S* rRNA fragment (~92 *v.* ~90%), and therefore our Protocol 1 can be applied directly to barcoding, accurate species delimitation and robust phylogeny estimation. Our protocol prescribes best practices for extracting maximum information from specimens that cannot be cultured, i.e. planktotrophic y-larvae that in the literature harbour a considerable diversity especially in temperate waters in Europe and North America, and the Arctic (Dreyer *et al.* 2023*a*).

Recent advancements in non-destructive DNA extraction have enabled integrative barcoding of complex, diverse invertebrate communities (Hunter *et al.* 2008; Cornils 2015; Carew *et al.* 2018). While adopting critical improvements that have been made to preserve the integrity of specimen cuticles, our protocols differ from those developed in other studies in several important ways. While ours rely on a strict culturing and live-video-recording approach at the single-specimen level, these also allow for initial mass-culturing of the single specimens that are used and the scoring of diverse morpho-ecological data layers throughout development, including between different larval instars (early/late nauplii) and phases (nauplius, cyprid and ypsigon). Our protocols offer a means of testing novel primer combinations that have high annealing potential for single y-larva DNA templates. Although we argue that single-specimen culturing, recording and amplification are essential to avoid biases in species delimitation and diversity estimation, other published protocols may also prove useful at least for y-larva barcoding.

Uncovering the evolutionary history of y-larvae

Our sequence data matrix recovered the same two clades found earlier (Pérez-Losada *et al.* 2009; Olesen *et al.* 2022) but, compared to these studies, we provide enhanced phylogenetic resolution by sequencing more markers for several new molecular operational taxonomic units (MOTUs), some of which are putative species awaiting taxonomic description (Fig. 1). Most morphospecies are notably supported by both molecular and morphological data in concert, testifying to the usefulness of combining these two data layers for phylogenetic inference. Given the molecular diversity in the phylogeny and the morphological disparity of the larval forms, these two clades may represent high-level taxonomic units (e.g. families) but further study is required to confirm this.

Facetotecta is in critical need of a new higher-level classification scheme, as all species are currently placed in the single genus *Hansenocaris*. The morphospecies from Sesoko Island alone, not to mention those from other localities, are both morphologically and phylogenetically so diverse that retaining these within the same genus is highly unsatisfactory. We expect that single-specimen culturing, live-imaging, amplification and sequencing will provide data in support of a new classification of Facetotecta (Olesen *et al.* 2022; Dreyer *et al.* 2023*a*, 2023*b*). Planktotrophic specimens cannot yet be cultured *in vitro,* posing a significant challenge that may delay detailed taxonomic descriptions. Morphologically separating Types A* and AE*, for example, into what are likely multiple 'species' (or MOTUs) that together occupy a geographical range extending from Japan and Taiwan to the Azores (Fig. 1) is exceedingly difficult. It is unlikely that each type constitutes one large breeding population. Wide geographical ranges of morphospecies could occur in some planktotrophic y-nauplii, however, because of their long developmental times and putative long-distance dispersal capabilities. For example, completion of the larval development of the Arctic facetotectan *Hansenocaris itoi* requires 3 months (April–June: Kolbasov *et al.* 2021) and this species may have a wide distribution.

Hansen (1899) identified y-nauplii that are virtually identical to *H. itoi* in the Baltic Sea (Kiel Bay), a place separated from the type locality in the White Sea by thousands of kilometres of open ocean. The lack of sequence data for the Baltic population, if still extant, illustrates the importance of Protocol 1, in which molecular sequence data are linked to live image data.

In conclusion, the easy-to-follow and inexpensive Protocol 1 we describe allows the extraction of maximal molecular and morphological information from single

y-larval specimens and we expect this method to be applicable to any invertebrate taxon with free-living larvae that moult. This should be used in future efforts, for example, to model life-history evolution in the Facetotecta or other groups. A growing inventory of sequence data from cultured, imaged and vouchered y-larva specimens or species from geographically diverse locations may allow for the evolution of Facetotecta and a comprehensive new taxonomic scheme for this group to finally be within reach.

Supplementary material

Supplementary Table S1 lists an overview of the specimens, collection data and GenBank accession numbers of the y-nauplii used in the study. Supplementary Table S2 shows the primer amplification success for the two DNA extraction methods. Supplementary material is available online and original figures and supplementary tables are available here https://figshare.com/s/9a08fa4635dd4f5c4b14

References

Bernot JP, Avdeyev P, Zamyatin A, Dreyer N, Alexeev N, Pérez-Losada M, Crandall KA (2022) Chromosome-level genome assembly, annotation, and phylogenomics of the gooseneck barnacle *Pollicipes pollicipes*. *GigaScience* **11**, giac021. doi:10.1093/gigascience/giac021

Bernot JP, Owen CL, Wolfe JM, Meland K, Olesen J, Crandall KA (2023) Major revisions in Pancrustacean phylogeny and evidence of sensitivity to taxon sampling. *Molecular Biology and Evolution* **40**, msad175. doi:10.1093/molbev/msad175

Böttger-Schnack R, Machida RJ (2011) Comparison of morphological and molecular traits for species identification and taxonomic grouping of oncaeid copepods. *Hydrobiologia* **666**, 111–125. doi:10.1007/s10750-010-0094-1

Carew ME, Coleman RA, Hoffmann AA (2018) Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ* **6**, e4980. doi:10.7717/peerj.4980

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540–552. doi:10.1093/oxfordjournals.molbev.a026334

Chan BKK, Dreyer N, Gale AS, Glenner H, Ewers-Saucedo C, Pérez-Losada M, Kolbasov GA, Crandall KA, Høeg JT (2021) The evolutionary diversity of barnacles, with an updated classification of fossil and living forms. *Zoological Journal of the Linnean Society* **193**, 789–846. doi:10.1093/zoolinnean/zlaa160

Chen HN, Høeg JT, Chan BKK (2013) Morphometric and molecular identification of individual barnacle cyprids from wild plankton: an approach to detecting fouling and invasive barnacle species. *Biofouling* **29**, 133–145. doi:10.1080/08927014.2012.753061

Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* **46**, 419–437. doi:10.1071/ZO98048

Cornils A (2015) Non-destructive DNA extraction for small pelagic copepods to perform integrative taxonomy. *Journal of Plankton Research* **37**, 6–10. doi:10.1093/plankt/fbu105

Dreyer N, Tsai PC, Olesen J, Kolbasov GA, Høeg JT, Chan BKK (2022*a*) Independent and adaptive evolution of phenotypic novelties driven by coral symbiosis in barnacle larvae. *Evolution* **76**, 139–157. doi:10.1111/evo.14380

Dreyer N, Palero F, Grygier M, *et al.* (2022*b*) Novel molecular resources for single-larva bar-coding of enigmatic crustacean y-larvae. *Authorea* **2022**, 6 December. [Preprint] doi:10.22541/au.167032805.55973749/v1

Dreyer N, Olesen J, Grygier MJ, Eibye-Jacobsen D, Høeg JT, Kerbl A, Fujita Y, Kolbasov GA, Savchenko AS, Worsaae K, Palero F, Chan BKK (2023*a*) The biology and life cycle of enigmatic crustacean y-larvae: a review. *Oceanography and Marine Biology: An Annual Review* **61**, 81–126.

Dreyer N, Palero F, Grygier MJ, Chan BKK, Olesen J (2023*b*) Single-specimen systematics resolves the phylogeny and diversity conundrum of enigmatic crustacean y-larvae. *Molecular Phylogenetics and Evolution* **184**, 107780. doi:10.1016/j.ympev.2023.107780

Gallego R, Heimeier D, Lavery S, Sewell MA (2015) The meroplankton communities from the coastal Ross Sea: a latitudinal study. *Hydrobiologia* **761**, 195–209. doi:10.1007/s10750-015-2487-7

Glenner H, Høeg JT, Grygier MJ, Fujita Y (2008) Induced metamorphosis in crustacean y-larvae: towards a solution to a 100-year-old riddle. *BMC Biology* **6**, 21. doi:10.1186/1741-7007-6-21

Grygier MJ (1987) New records, external and internal anatomy, and systematic position of Hansen's y-larvae (Crustacea: Maxillopoda: Facetotecta). *Sarsia* **72**, 261–278. doi:10.1080/00364827.1987.10419723

Grygier MJ, Høeg JT, Dreyer N, Olesen J (2019) A new internal structure of nauplius larvae: a "ghostly" support sling for cypris y left within the exuviae of nauplius y after metamorphosis (Crustacea: Thecostraca: Facetotecta). *Journal of Morphology* **280**, 1222–1231. doi:10.1002/jmor.21026

Hansen HJ (1899) Die Cladoceren und Cirripedien der Plankton-Expedition. *Ergebnisse der in dem Atlantischen Ocean von Mitte Juli bis Angang November 1999 ausgeführten Plankton-Expedition der Humboldt-Stiftung* **2**(G, d), 1–58. [In German]

Hartop E, Srivathsan A, Ronquist F, Meier R (2022) Towards large-scale integrative taxonomy (LIT): resolving the data conundrum for dark taxa. *Systematic Biology* **71**, 1404–1422. doi:10.1093/sysbio/syac033

Hunter SJ, Gooddall T, Walsh K, Owen R, Day JC (2008) Nondestructive DNA extraction from blackflies (Diptera: Simuliidae): retaining voucher specimens for DNA barcoding projects. *Molecular Ecology Resources* **8**(1), 56–61. doi:10.1111/j.1471-8286.2007.01879.x

Itô T (1985) Contributions to the knowledge of cypris y (Crustacea: Maxillopoda) with reference to a new genus and three new species from Japan. *Special Publication of the Mukaishima Marine Biological Station* **1985**, 113–122.

Itô T (1986) Three types of "nauplius y" (Maxillopoda: Facetotecta) from the North Pacific. *Publications of the Seto Marine Biological Laboratory* **31**, 63–73. doi:10.5134/176117

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587–589. doi:10.1038/nmeth.4285

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780. doi:10.1093/molbev/mst010

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**(12), 1647–1649. doi:10.1093/bioinformatics/bts199

Kolbasov GA, Petrunina AS, Dreyer N, Prudkovsky AA, Olesen J, Chan BKK, Høeg JT (2021) Naupliar development of Facetotecta (Crustacea: Thecostraca) and the nature of the first nauplius instar in the Crustacea – pro et contra the Hexanauplia concept. *Organisms Diversity & Evolution* **21**, 209–230. doi:10.1007/s13127-021-00479-y

Kolbasov GA, Savchenko AS, Dreyer N, Chan BKK, Høeg JT (2022) A synthesis of the external morphology of cypridiform larvae of Facetotecta (Crustacea: Thecostraca) and the limits of the genus *Hansenocaris*. *Ecology and Evolution* **12**, e9488. doi:10.1002/ece3.9488

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial *COI* region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* **10**, 34. doi:10.1186/1742-9994-10-34

Lozano-Fernández J, Giacomelli M, Fleming JF, Chen A, Vinther J, Thomsen PF, Glenner H, Palero F, Legg DA, Iliffe TM, Pisani D, Olesen J (2019) Pancrustacean evolution illuminated by taxon-rich

genomic-scale data sets with an expanded remipede sampling. *Genome Biology and Evolution* **11**, 2055–2070. doi:10.1093/gbe/evz097

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530–1534. doi:10.1093/molbev/msaa015

Oakley TH, Wolfe JM, Lindgren AR, Zaharoff AK (2013) Phylotranscriptomics to bring the understudied into the fold: monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. *Molecular Biology and Evolution* **30**, 215–233. doi:10.1093/molbev/mss216

Olesen J, Grygier MJ (2022) Two new species of lecithotrophic nauplius y with remarkable labra from Okinawa, Japan, and a family-group name for y-larvae (Crustacea: Thecostraca: Facetotecta: Hansenocarididae fam. nov.). *Species Diversity* **27**(2), 301–317. doi:10.12782/specdiv.27.301

Olesen J, Grygier MJ (in press) Taxonomic diversity of marine planktonic "y-larvae" (Crustacea: Facetotecta) from a coral reef hotspot locality (Japan, Okinawa), with a key to y-nauplii. *European Journal of Taxonomy*.

Olesen J, Dreyer N, Palero F, Eibye-Jacobsen D, Fujita Y, Chan BKK, Grygier MJ (2022) Integrative taxonomy of crustacean y-larvae (Thecostraca: Facetotecta) using laboratory-rearing and molecular analyses of single specimens, with the description of a new vermiform species. *Zoological Journal of the Linnean Society* **196**, 549–592. doi:10.1093/zoolinnean/zlac020

Olesen J, Grygier MJ, Herranz M (2024) Autofluorescence imaging of exuviae as a tool for studying slide preparations of micro-arthropods, exemplified by a museum collection of the enigmatic crustacean "y-larvae" (Pancrustacea: Facetotecta). *Zoomorphology*. [Published 16 February 2024] doi:10.1007/s00435-024-00641-y

Pérez-Losada M, Crandall KA, Kolbasov GA, Høeg JT (2002) Reanalysis of the relationships among the Cirripedia and the Ascothoracida and the phylogenetic position of the Facetotecta (Maxillopoda: Thecostraca) using *18S* rDNA sequences. *Journal of Crustacean Biology* **22**, 661–669. doi:10.1163/20021975-99990278

Pérez-Losada M, Høeg JT, Crandall KA (2009) Remarkable convergent evolution in specialized parasitic Thecostraca (Crustacea). *BMC Biology* **7**, 15. doi:10.1186/1741-7007-7-15

Petrunina AS, Neretina TV, Mugue NS, Kolbasov GA (2014) Tantulocarida versus Thecostraca: inside or outside? First attempts to resolve phylogenetic position of Tantulocarida using gene sequences. *Journal of Zoological Systematics and Evolutionary Research* **52**, 100–108. doi:10.1111/jzs.12045

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079–1083. doi:10.1038/nature08742

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539–542. doi:10.1093/sysbio/sys029

Schizas NV, Street GT, Coull BC, Chandler GT, Quattro JM (1997) An efficient DNA extraction method for small metazoans. *Molecular Marine Biology and Biotechnology* **6**(4), 381–383.

Schram TA (1972) Further records of nauplius y type IV Hansen from Scandinavian waters. *Sarsia* **50**, 1–24. doi:10.1080/00364827.1972.10411213

Schwentner M, Combosch DJ, Pakes Nelson J, Giribet G (2017) A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. *Current Biology* **27**(12), 1818–1824. doi:10.1016/j.cub.2017.05.040

Tsang LM, Chan BKK, Shih FL, Chu KH, Allen Chen C (2009) Host-associated speciation in the coral barnacle *Wanella milleporae* (Cirripedia: Pyrgomatidae) inhabiting the *Millepora* coral. *Molecular Ecology* **18**(7), 1463–1475. doi:10.1111/j.1365-294X.2009.04090.x

von Reumont BM, Jenner RA, Wills MA, Dell'Ampio E, Pass G, Ebersberger I, Meyer B, Koenemann S, Iliffe TM, Stamatakis A, Niehuis O, Meusemann K, Misof B (2012) Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of *Hexapoda*. *Molecular Biology and Evolution* **29**(3), 1031–1045. doi:10.1093/molbev/msr270

Watanabe HK, Senokuchi R, Shimanaga M, Yamamoto H (2016) Comparison of the efficiency of three methods of DNA extraction for deep-sea benthic copepods. *JAMSTEC Report of Research and Development* **23**, 52–59. doi:10.5918/jamstecr.23.52

Data availability. All new DNA sequences have been deposited in the NCBI GenBank under accession numbers PP336629–PP336655 (*28S* rRNA), PP336577–PP336602 (*18S* rRNA), PP336558–PP336576 (*16S* rRNA) and PP337082 (*COX1*) (Supplementary Table S1). All commands necessary to reproduce computational analyses are listed directly in the manuscript. Vouchers and slides are deposited in the Olesen lab at the Natural History Museum of Denmark. All original figures, tables and supplementary material can be found at: https://figshare.com/s/9a08fa4635dd4f5c4b14. A preprint of an earlier version of this manuscript is available in Dreyer *et al.* (2022*b*).

Conflicts of interest. N. Dreyer is a postdoctoral fellow with the Editor-in-Chief of *Invertebrate Systematics*, Gonzalo Giribet at Harvard University. To mitigate this potential conflict of interest, this paper was handled by another editor while in peer review. The authors declare that they have no other conflicts of interest.

Declaration of funding. This paper was primarily sponsored by a Villum Experiment grant (number 17467) to J. Olesen and a double-degree graduate stipend to N. Dreyer from the Natural History Museum of Denmark and the Taiwan International Graduate Program (TIGP Biodiversity, Academia Sinica, Taiwan). F. Palero acknowledges the project 'CIDEGENT/2019/028 – Biodiversity Patterns of Crustacea from Karstic Systems (BIOPACKS): molecular, morphological and functional adaptations' funded by the Conselleria d'Innovació, Universitats, Ciència i Societat Digital and project CRUSTRESS (code PID2020-112959GB-I00) funded by MCIN/AEI/10.13039/501100011033. A.S. Savchenko and G.A. Kolbasov acknowledge Russian Science Foundation Grant 22-24-00182. B. K. K. Chan acknowledges the grant from National Science and Technology Council, Taiwan (110-2923-B-001-003-MY3). G. A. Kolbasov acknowledges the Russian Foundation for Basic Research Grant 21-54-52003 MNT_a.

Acknowledgements. Members of the Coastal Ecology lab at Academia Sinica, especially Yao-Feng Tsao and Pei-Chen Tsai, are thanked for providing training in molecular biology and assistance with sampling in Taiwan. Dr Ramon Gallego Simon (University of Washington, Seattle, WA, USA) is thanked for generously sharing his DNA extraction and amplification protocols for Antarctic y-larvae during the onset of our study. Two reviewers, the Associate Editor and the Editor-in-Chief provided valuable comments that improved the quality of the manuscript.

Author contributions. N. Dreyer designed and co-conceived the study, participated in sampling, and co-designed and tested primers, extraction methods and protocols. N. Dreyer also performed all sequencing and most phylogenetic and statistical analyses, designed graphical outputs, photographed specimens and wrote the first manuscript draft. J. Olesen co-supervised the project, co-conceived the study, led sampling in Japan, participated in sampling in Taiwan, co-designed protocols, co-funded the project, assisted in phylogenetic analyses, designed graphical outputs, photographed specimens and edited manuscript drafts. M. J. Grygier participated in sampling, co-developed protocols, originally developed the 'exuvium'-voucher strategy and edited manuscript drafts. D. Eibye-Jacobsen assisted in field work and sorted samples. A. S. Savchenko provided the specimens from the Azores and illustrations of *H. itoi*, and edited manuscript drafts. Y. Fujita arranged laboratory space and provided microscopes and other practical support for field work in Japan, and assisted in sampling. G.A. Kolbasov provided the White Sea material and edited manuscript drafts. R. J. Machida co-designed and co-supervised DNA extraction methods, assisted in primer testing and edited manuscript drafts. B. K. K. Chan co-supervised the project, led sampling in Taiwan and co-funded the project. F. Palero co-supervised the project, designed primers, assisted in statistical analyses and participated in sampling. All authors approved the final version.

Author affiliations

[A]Natural History Museum of Denmark, University of Copenhagen, Denmark.

[B]Biodiversity Research Center, Academia Sinica, Taipei, Taiwan.

[C]Department of Life Science, National Taiwan Normal University, Taipei, Taiwan.

[D]Biodiversity Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan.

[E]National Museum of Marine Biology & Aquarium, Checheng, Pingtung, Taiwan.

[F]Invertebrate Zoology Department, Biological Faculty, Moscow State University, Moscow, Russian Federation.

[G]General Education Center, Okinawa Prefectural University of Arts, Naha, Okinawa, Japan.

[H]White Sea Biological Station, Biological Faculty of Moscow State University, Moscow, Russian Federation.

[I]Institut Cavanilles de Biodiversitat i Biologia Evolutiva (ICBIBE), University of Valencia, Paterna, Spain.

[J]Present address: Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA.