Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 February 2021 doi:10.20944/preprints202102.0226.v1

RNA-sequencing indicates high hemocyanin expression as a key strategy for cold adaptation in the Antarctic amphipod *Eusirus* cf. *giganteus* clade g3

SAMUELE GRECO[1§]; ELISA D'AGOSTINO[2§], CHIARA MANFRIN[1]; ANASTASIA SERENA GAETANO[1], GAEL FURLANIS[1], FRANCESCA CAPANNI[1], GIANFRANCO SANTOVITO[2]; PAOLO EDOMI[1]; PIERO GIULIANINI[1]; MARCO GERDOL[1*]

[1]Department of Life Sciences, University of Trieste, 34127 Trieste, Italy

[2]Department of Biology, University of Padova, 35131 Padova, Italy

[§]These authors contributed equally to this work

[*]Address correspondence to mgerdol@units.it

Key words: Antarctica, cold adaptation, hemocyanin, amphipod, transcriptome

Abstract: We here report the *de novo* transcriptome assembly and functional annotation of *Eusirus* cf. *giganteus* clade g3, providing the first database of expressed sequences from this giant Antarctic amphipod. RNA-sequencing, carried out on the whole-body of a single juvenile individual likely undergoing molting, revealed the dominant expression of hemocyanins. The mRNAs encoding these oxygen-binding proteins cumulatively accounted for about 40% of the total transcriptional effort, highlighting the key biological importance of high hemocyanin production in this Antarctic amphipod species. We speculate that this observation may mirror a strategy previously described in Antarctic cephalopods, which compensate the decreased ability to release oxygen to peripheral tissues at sub-zero temperatures by massively increasing total blood hemocyanin content compared with temperate species. These preliminary results will undoubtedly require confirmation through proteomic and biochemical analyses aimed at characterizing the oxygen-binding properties of *E.* cf. *giganteus* clade g3 hemocyanins, and at investigating whether other Antarctic arthropod species exploit similar adaptations to cope with the challenges posed by the extreme conditions of the polar environment.

Introduction

Crustaceans are the most species-rich group of metazoans in Antarctic benthic communities (Arntz et al., 1994) and amphipods largely contribute to this biodiversity, with over 800 different species described to date (De Broyer and Jazdzewski, 1996, 1993). Moreover, several Antarctic crustacean species are believed to comprise complexes of cryptic or species that still remain to be formally described (Brandão et al., 2010; Held and Wägele, 2005; Loerz et al., 2009; Raupach and Wägele, 2006). Thanks to the high oxygen availability of Antarctic waters, amphipods occupy all available micro habitats of the Southern Hemisphere and can therefore be considered among the most successful colonizers of these marine environments (Levin and Gage, 1998). Even though Antarctic amphipods most certainly occupy a key position in polar trophic chains (Dauby et al., 2001, 2002), studies focused on these widespread metazoans are still relatively scarce, and molecular or genetic data are nearly entirely missing for several relevant genera commonly found in polar waters.

Amphipods belonging to the genus *Eusirus* (Krøyer, 1845), and part of the suborder Amphilochidea, superfamily Eusiroidea, have been long known to have a broad circum-Antarctic distribution. According to the Register of Antarctic Marine Species, eight different species belonging to the *Eusirus* genus have been described to date in the Antarctic continent: *Eusirus antarcticus* Thomson, 1880, *Eusirus bouvieri* Chevreux, 1911, *Eusirus giganteus* Andres, Lörz & Brandt, 2002, *Eusirus laevis* Walker, 1903, *Eusirus laticarpus* Chevreux, 1906, *Eusirus microps* Walker, 1906, *Eusirus perdentatus* Chevreux, 1912 and *Eusirus propeperdentatus* Andres, 1979. *E. perdentatus*, the most widespread and the largest among these species, shows typical features of Antarctic gigantism, with adult individuals reaching and often exceeding the size of 70 mm (Chapelle, 2002). However, the marginal morphological differences between *E. perdentatus* and the congeneric *E. giganteus*, first described in 1972, have been the source of taxonomical uncertainties over the years (Andres et al., 2002), until 2012, when molecular approaches were applied for the first time to study Antarctic giant amphipod populations. The results provided by these studies have challenged previously accepted taxonomy, strongly hinting the presence of multiple cryptic species with limited gene flow among each other. In particular, the authors found that *E. perdentatus* harbored two previously undetected cryptic species and *E. giganteus* at least three, defined as clade g1, g2 and g3, some of which occur in sympatry (Baird et al., 2011).

Therefore, although specimens collected in several different locations across all Antarctica have been previously classified as belonging to *E. giganteus* (Gutt, 2008; OBIS, 2020; Verheye et al., 2016), it is likely that these represent multiple cryptic species. Hence, all data collected for this species should be referred to *E.* cf. *giganteus*.

Despite the prevalence of Eusiridae in Antarctic high latitude waters, molecular data are nearly not existing for this taxon, with a total of 220 nucleotide sequences deposited in GenBank (as of January 31st, 2021). At the same time, studies on *Eusirus* spp. remain very limited and, to the best of our knowledge, the molecular and genetic bases of its adaptation to cold have not been explored in detail. This aspect may be of particular interest due to the possible threats giant amphipod species could face due to climate change along with the progressive decline of ocean oxygen availability expected to occur in the next few decades (Spicer and Morley, 2019).

With this work, we tried to fill this knowledge gap, providing a reference annotated transcriptome assembly for *E.* cf. *giganteus* clade g3 (the first resource of this kind in

*Address correspondence to: Marco Gerdol, mgerdol@units.it
Received: xx Month Year; Accepted: xx Month Year

the infraorder Amphilochida), which may serve as a reference for future studies and could provide a useful resource to investigate molecular strategies of cold adaptation in Antarctic amphipods.

Results

Species identification

The aforementioned presence of multiple *Eusirus* spp. cryptic species in Antarctica, whose recognition is not possible by morphological examination, implies the need to use genetic markers for a proper taxonomical identification. The general external morphology of an adult *Eusirus* cf. *giganteus* individual, sampled in the same geographical location where the juvenile specimen used for RNA-sequencing was collected (see the Materials and Methods section), is exemplified in **Figure 1**. Although molecular data available for this genus are very scarce, the screening of the transcriptomes for previously studied molecular markers allowed us to unequivocally place the specimen target of this study within the clade g3 of *Eusirus* cf. *giganteus* identified by Baird and colleagues (Baird et al., 2011). Indeed, the sequences of cytochrome oxidase 1 (JN001762.1), cytochrome B (JN001729.1) and of the Internal Transcribed Spacer 2 (JN001807.1) perfectly matched, without gaps and mismatches, those reported in this study. Clade g3 has been previously linked with a primary distribution in the Ross Sea along with another *E.* cf. *giganteus* clade (i.e. clade g2) and two *E.* cf. *perdentatus* clades (i.e., p1 and p3), which is consistent with the placement of the Mario Zucchelli Antarctic Base.

FIGURE 1. External morphology of an adult *Eusirus* cf. *giganteus* individual, collected close to the Italian Mario Zucchelli Antarctic Base (Ross Sea 74° 38′ 402″ S, 164° 39′ 281″ E), in November 2017.

Transcriptome assembly and annotation

The *de novo* transcriptome assembly, generated starting from a total of 101,310,170 trimmed paired-end Illumina reads, comprised 57,607 contigs, 4,369 of which exceeded 1Kb in length (**Table 1**). 11,217 contigs (19.47%) were annotated based on significant BLASTx matches in UniProtKB, leading to the association of 10,247, 10,239 and 10,343 contigs with Gene Ontology Biological Process, Molecular Function and Cell Component terms, respectively. Moreover, 10,420 contigs (18.09% of the total) were associated to one or more Pfam conserved domains. The observed mapping rate was 78.26%.

Table 1

Assembly and read mapping statistics for the *de novo* transcriptome assembly of *Eusirus* cf. *giganteus* clade g3

parameter	value
Number of contigs	57,607
GC content	0.41
Contig N50	519
Total annotated contigs	13,874
BLASTx annotated contigs	11,217
Contigs with GO Biological Process annotations	10,247
Contigs with GO Molecular Function annotations	10,239
Contigs with GO Cellular Component annotations	10,343
Contigs with Pfam annotations	10,420

High expression of hemocyanins

The *E.* cf. *giganteus* clade g3 transcriptome displayed an unusual skewed distribution of read mapping, with nearly 30 million reads being captured by a relatively high number of fragmented contigs sharing the same functional annotation as hemocyanins. Cumulatively, these contigs accounted for 411,000 Transcripts Per Million (TPM), indicating that an extraordinarily high transcriptional effort (i.e. over 40% of the total) was employed in the synthesis of mRNAs encoding these oxygen-transporting proteins (**Figure 2**).

Several of the other contigs achieving TPM values > 10,000, and expressed at levels higher than housekeeping genes such as ribosomal structural proteins and components of the mRNA transcription machinery, shared close homology with trypsins, chymotrypsins, chitinases or brachyurins (**Figure 2**). All these proteins play a fundamental role in the molting process in crustaceans (Gao et al., 2017; Van Wormhoudt et al., 1995), and this is consistent with the developmental stage of the sample individual. Indeed, juvenile amphipods would be expected to undergo frequent molting with relatively short intermolt periods (Chang and Mykles, 2011). Overall, these molting-related enzymes achieved a cumulative expression level of ~110,000 TPMs. Mitochondrial RNAs accounted for an additional ~50,000 TPMs, implying that less than 45% of the global transcriptional effort in *E.* cf. *giganteus* clade g3 was put in the synthesis of other nuclear mRNAs.

In line with such an unbalanced transcript representation, the *de novo* assembly process led to a collection of expressed transcripts characterized by a low level of completeness. This was clearly evidenced by the presence of just 29.6% complete arthropod BUSCOs, accompanied by a high number of fragmented (22.0%) and missing (48.4%) orthologs. The duplication rate, on the other hand, was rather low, standing at just 2.5%

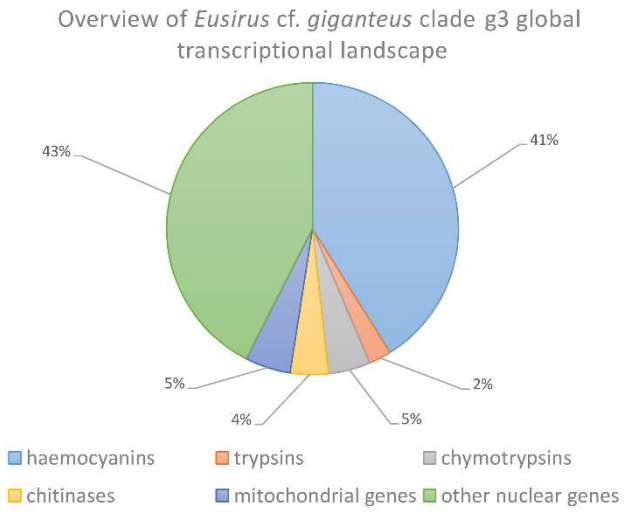

FIGURE 2. Overview of the transcriptional landscape of *E.* cf. *giganteus* clade g3, with details about the relative contribution to global transcription of contigs encoding heamocyanins, trypsins, chymotripsins and chitinases. The relative contribution of non-nuclear (i.e. mitochondrial) genes is shown in a separate category.

Characterization of a highly expressed full-length hemocyanin sequence

Although 280 contigs could be linked with heamocyanins based on BLAST annotation, the vast majority of them were small fragments, measuring just a few hundred nucleotides, whereas the complete length of the coding sequence of arthropod hemocyanins was expected to be ~2 Kb long, with a complete Open Reading Frame of ~700 codons.

The *de novo* sequencing strategy we applied, which combined three different transcriptome assembly algorithms (Trinity, SPAdes and TransABySS) allowed us to explore in depth the outputs of each single process, with the aim to recover the most likely full-length mRNA precursors. Significant fragmentation of hemocyanin mRNAs was evidenced in all assembly methods, which could be indicative of the presence of several nearly-identical paralogous genes in this species. However, we recovered a single contig which included a complete Open Reading Frame of 673 codons. This sequence, capturing ~27.5 million reads, reached an expression level equal to ~118,550 TPMs, thereby accounting for over 10% of the global transcriptional effort of *E.* cf. *giganteus* clade g3.

This hemocyanin sequence found the best BLASTx match with the Hc A and B chains from the decapods *Panulirus interruptus* and *Panulirus vulgaris*, among those deposited in UniProtKB (64% pairwise sequence identity) (Bak and Beintema, 1987; Jekel et al., 1996, 1988), and with the Hc subunit 1 from *Gammarus roeseli* among those deposited in nr (82% pairwise sequence identity) (Hagner-Holler et al., 2005).

The *E.* cf. *giganteus* clade g3 Hc displayed the three expected canonical domains found in arthropod hemocyanins, i.e. the all N-terminal *all alpha* domain (PF03722), the central copper-containing domain (PF00372) and the C-terminal Ig-like domain (PF03723). Moreover, the copper-containing domain presented all the six highly conserved histidine residues involved in copper binding, confirming its identification as a *bona fide* Hc (**Figure 3**).

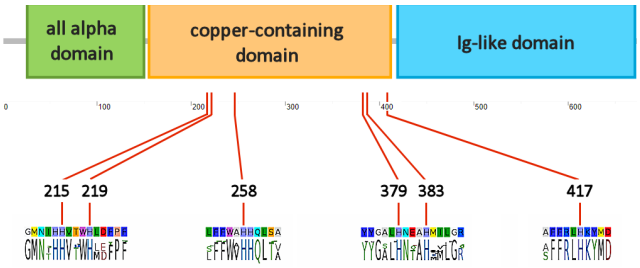

FIGURE 3. Schematic overview of the domain organization of *Eusirus* cf. *giganteus* clade g3 hemocyanin. A detail of the six histidine residues expected to be involved in copper binding is provided (His215, His219, His258, His379, His383, His417), along with the consensus sequence of the neighboring residues in crustacean hemocyanin sequences deposited in UniprotKB.

Discussion

Hemocyanins, found only in Arthropoda and Mollusca, are only second in abundance to hemoglobins as molecules capable of temporarily binding oxygen in organs devoted to gaseous exchange, transporting it to peripheral tissues, and releasing it to these locations due to lower partial pressure (Redmond, 1955). In polar regions, and in particular in the Southern Ocean, where seawater temperature is constantly below zero, the increased solubility of oxygen potentially offers the opportunity for astounding adaptations that involve the circulatory system and oxygen-carrying molecules, such as in the case of icefish (Beers et al., 2010). However, this increased oxygen solubility may paradoxically result in a lower bioavailability due to the difficulties linked with its release in tissues.

The few studies carried out so far on hemocyanins in Antarctica have revealed that some invertebrates developed interesting molecular strategies to overcome this issue. For example, the Antarctic octopus *Pareledone charcoti* compensates a lower Hc oxygen-binding affinity with a higher hemocyanin content in hemolymph compared with species living in temperate waters (Oellermann et al., 2015). Although no proteomic data are available for *E.* cg. *giganteus* clade g3 to evaluate the actual relative abundance of hemocyanins in the hemolymph, we here provide evidence suggesting that a similar strategy might have been employed by this Antarctic amphipod.

We show that hemocyanin-encoding mRNAs cumulatively accounted for nearly 40% of the total transcriptional effort of the juvenile individual subjected to RNA-sequencing, suggesting that Hcs are by far the most abundant plasma proteins in this species. Transcriptome data strongly suggest that multiple highly similar hemocyanin mRNAs are simultaneously expressed, which would point towards the existence of several co-regulated paralogous Hc gene copies. This would be in line with the organization of Hc genes found in other amphipods, such as *Hyalella azteca*, which harbors 9 Hc genes, seven of which display very high pairwise sequence homology (Poynton et al., 2018). Likewise, high sequence similarity between different crustacean hemocyanin chains has also been previously shown in *P. interruptus*, where Hc A and B chains share 96% primary sequence identity (Jekel et al., 1988).

In absence of genomic data, this remains a working hypothesis, which would however provide a reasonable explanation for the extremely high levels of expression of *E.* cf. *giganteus* clade g3 Hcs and for the observed severe fragmentation of Hc-coding contigs. Indeed, the low pairwise sequence divergence among paralogous gene copies might have hampered the possibility to obtain a full-length assembly for the other isoforms.

The single complete Hc mRNA sequence we retrieved, which presumably represents the most highly expressed isoform, encodes a full-length protein with all the expected structural features of functional hemocyanins, i.e. the presence of the three characterizing *all alpha*, copper-containing, and Ig-like domains, as well as of the six conserved residues involved in copper binding (**Figure 3**). Bayesian phylogeny placed this sequence with high confidence within known members of the crustacean

hemocyanin superfamily, and close to a cluster of seven highly similar paralogous genes from the amphipod *Hyalella azteca*. At the same time, the *E.* cf. *giganteus* clade g3 sequence only showed distant homology with non-respiratory crustacean cryptocyanins and pseudohemocyanins, which lack two of the histidine residues in the aforementioned array (Burmester, 1999; Terwilliger et al., 1999) (**Figure 4**).

Altogether, these observations suggest that *E.* cf. *giganteus* clade g3, and possibly other Antarctic amphipods, may have evolved a strategy similar to that of cephalopods, maximizing the expression of oxygen-carrying molecules to overcome the decreased ability to unload oxygen in peripheral tissues at sub-zero temperatures. It remains to be determined whether this increased expression is linked with a decreased oxygen-binding affinity, as in cephalopods. Curiously, our findings are in stark contrast with those reported for other Antarctic crustaceans, such as *Glyptonotus antarcticus* (Whiteley et al., 1997), for which very low hemocyanin oxygen-binding capacities and low circulating protein levels have been reported. Further transcriptomic, proteomic and biochemical studies targeting other Antarctic crustacean species may shed some light on this issue.

The preliminary indications deriving from this study have also important implications concerning the proposed relationship between oxygen availability and gigantism in Antarctic waters, which has been subject of intense debate over the past few years (Spicer and Morley, 2019). Indeed, three alternative hypotheses have been proposed to explain the relationship between body size and oxygen bioavailability: (i) the "oxygen limitation hypothesis", according to which gigantism is allowed by the combination between a higher availability of oxygen in freezing waters and the lowest metabolic rates observed in cold-adapted organisms (Chapelle and Peck, 1999); (ii) the "respiratory advantage hypothesis", which is based on the idea that, in spite of a larger solubility, o bioavailability in Antarctic waters is lower, due to a decreased diffusion coefficient, leading organisms with scarce respiratory control to develop gigantism (Verberk and Atkinson, 2013); (iii) the "symmorphosis hypothesis", which postulates that Antarctic organisms have developed evolutionary adaptations to optimize oxygen supply, regardless of their body size (Woods et al., 2009).

In light of the imminent climate changes, whose effects are already visible in some areas of the Antarctica, these alternative interpretations have profound effects on the predicted fate of giant amphipods. Indeed, these animals might be either among the first or among the latest organisms to face serious threats, depending on whether the oxygen limitation or the respiratory advantage hypotheses are correct (Spicer and Morley, 2019).

The exceptionally high expression of hemocyanins in *E.* cf. *giganteus* clade g3 seems to be in contrast with the hypothesis proposed by Chapelle and Peck. In this case, it clearly mirrors the strategy used by cephalopods to overcome the limited ability to unload oxygen in peripheral tissues at low temperatures, which would be consistent with the premises of the "respiratory advantage hypothesis". However, it is noteworthy that, according to the latter hypothesis, gigantism would only be expected to occur in species lacking efficient respiratory pigments or with scarce respiratory control.

Clearly, the collection of additional physiological, biochemical and molecular data from *Eusirus* spp. and other giant Antarctic crustaceans that do not rely on high hemocyanin expression as a strategy for cold adaptation (such as *G. antarcticus*) would be needed to clarify the intricate relationships that exist between oxygen bioavailability, oxygen-carrying molecules and body size in these animals.

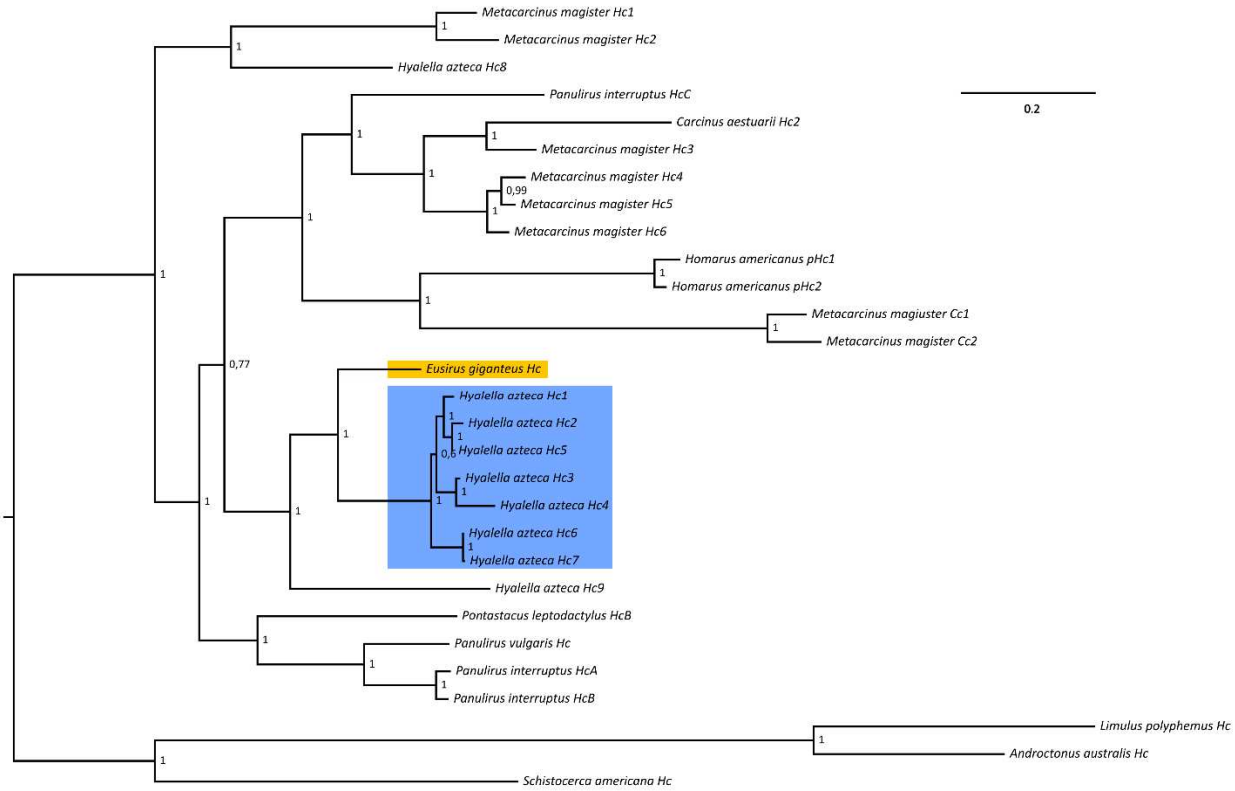

FIGURE 4: Bayesian phylogeny of crustacean hemocyanins. The placement of *Eusirus giganteus* haeomocyanin is highlighted with an orange background. The cluster of paralogous hemocyanin genes from the amphipod *Hyalella azteca* is highlighted with a blue background. Posterior probability values are shown close to each node. Hc: hemocyanin; Cc: cryptocyanin; pHc: pseudohemocyanin.

Materials and Methods

Sample collection and RNA-sequencing

A single *E. giganteus* specimen was collected as a bycatch during the sampling campaign of the intended target species, the amphipod *Pseudorchomene plebs* (Hurley, 1965). Sampling occurred close to the Italian Mario Zucchelli Antarctic Base (Ross Sea 74° 38′ 402″ S, 164° 39′ 281″ E), in November 2017, within the frame of the activities of the Italian XXXIII PNRA expedition. The specimen was a juvenile, approximately 15.8 mm long (from the tip of the head to the base of the telson along the dorsal side), of undetermined sex.

The animal was transferred to the facilities of the base, in a tank with oxygenated running seawater at the same temperature recorded in the natural external environment (-1.9°C). After one week of acclimatization, to the recovery from the stress of sampling, the specimen was sacrificed by immersion in RNAlater (Thermo Fisher Scientific, Waltham, USA), and immediately stored at -80°C.

Following its transportation to the laboratories of the University of Trieste, the specimen was moved into Trizol (Thermo Fisher Scientific, Waltham, USA) and homogenized. RNA extraction was carried out following the manufacturer's instructions, yielding material of sufficient quality and quantity to prepare a mRNA-seq library with the Lexogen SENSE mRNA-seq library prep kit v2 (Lexogen, Wien, Austria), as evaluated by the use of an Agilent 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, USA).

RNA-sequencing was performed at the Genomics and Epigenomics Platform of the Area Science Park (Trieste, Italy), on an Illumina NovaSeq 6000 instrument, with a 2 x 150 bp paired-end strategy.

Ethical statement

The sample collection complied with the regulations provided by the Italian Ministry of Education, University and Research concerning activities and environmental protection in Antarctica and with the Protocol on Environmental Protection to the 137 Antarctic Treaty, Annex II, Art. 3. All the activities on animals performed during the Italian Antarctic Expedition were under the control of a PNRA Ethics Referent, which acts on behalf of the Italian Ministry of Foreign Affairs. In particular, the required data for the project PNRA16_00099 are as follows. Name of the ethics committee or institutional review board: Italian Ministry of Foreign Affairs. Name of PNRA Ethics Referent: Dr. Carla Ubaldi, ENEA Antarctica, Technical Unit (UTA).

Sequencing data processing, de novo transcriptome assembly and annotation

Raw reads were subjected to base-calling and sequencing quality trimming according to the output of FastQC v. 0.11.9 (Andrews, 2010) with fastp v. 0.20.0 (Chen et al., 2018). Trimmed reads were used as an input for a *de novo* transcriptome assembly using the Oyster River Protocol (ORP) v.2.3.1 (MacManes, 2018), a tool which generates a unique, non-redundant assembly by joining the outputs of Trinity (Grabherr et al., 2011), SPAdes (Bankevich et al., 2012) and TransABySS (Robertson et al., 2010). Poorly expressed contigs, which may either derive from exogenous contamination (e.g. gut content) or from transcripts with little biological relevance were excluded from the assembly using the TPM FILT=1 parameter.

The completeness of the transcriptome was evaluated with an analysis carried out with BUSCO v.4.1.4 (Simão et al., 2015) against the OrthoDB v.10 database (Kriventseva et al., 2019), checking the presence of a set of highly conserved single-copy orthologs shared by all arthropods.

The transcriptome was functionally annotated with AnnotaM (https://gitlab.com/54mu/annotaM), associating each contig to Gene Ontology terms (Ashburner et al., 2000) and Pfam conserved protein domains (Punta et al., 2012). Annotation was based on BLASTx matches against the UniProtKB/Swiss-Prot database (e-value threshold = 1×10^{-5}), and HMMer v.3.1b2 (Finn et al., 2011) matches against the Pfam-A 34.0 database with default settings.

Gene expression levels were calculated as Transcript Per Million (TPM) (Wagner et al., 2012), based on read mapping using the CLC Genomics Workbench v.20 (Qiagen, Hilden, Germany) RNA-seq mapping tool, with stringent parameters (length fraction = 0.75, similarity fraction = 0.98).

Hamocyanin sequence characterization and phylogenetic analysis

We recovered a single Hc-encoding contig bearing a complete Open Reading Frame thanks to the inspection of the preliminary assemblies produced by ORP. This was virtually translated into an amino acid sequence with the Expasy *translate* tool (Gasteiger et al., 2003), and the presence of the six histidine residues expected to be involved in copper binding in arthropod Hcs was evaluated based on the literature data (Burmester, 1999).

The sequences of *C. magistrus* Hc and cryptocyanins (Terwilliger et al., 2006), as well as other phylogenetically informative crustacean sequences from UniProtKB and nr were also included in the creation of a multiple sequence alignment (MSA) with MUSCLE (Edgar, 2004). Namely, *P. vulgaris* Hc (P80888.2), *P. interruptus* Hc chain A (P04254.2), B (P10787.1) and C (P80096.1), *Homarus americanus* pseudocyanin-1 (Q6KF82.1) and -2 (Q6KF81.1), *Carcinus aestuarii* Hc subunit 2 (P84293.1) and *Pontastacus leptodactylus* Hc chain B (P83180.1) were used, along with the Hc sequences obtained from the genome of the amphipod *H. azteca* (Poynton et al., 2018). The Hc sequenced from *Limulus polyphemus* (P04253.2), *Schistocerca americana* (AAC16760.1) and *Androctonus australis* (P80476.1) were used as outgroups for tree rooting purposes.

The MSA was processed with Gblocks (Talavera and Castresana, 2007) to remove phylogenetically non-informative positions. The resulting file was analyzed with Modeltest-ng (Darriba et al., 2020) to assess the best-fitting model of molecular evolution, which was identified as the Le and Gascuel model, with a gamma-distributed rate of variation across sites and a fixed (empirical) prior on state frequencies (LG+G+F) (Le and Gascuel, 2008). Model choice was based on the corrected Akaike Information Criterion (Cavanaugh, 1997).

Phylogenetic inference analysis was carried out with MrBayes v. 3.2.7a (Huelsenbeck and Ronquist, 2001), with two MCMC analyses run in parallel for 200,000 generations, until all estimated parameters of the model reached an ESS >= 200, as estimated with Tracer v.1.7 (Rambaut et al., 2018).

Abbreviations

Hc: hemocyanin
LG: Le and Gascuel model
MSA: Multiple Sequence Alignment
ORP: Oyster River Protocol
TPM: Transcript Per Million

References

Andres, H.G., Lörz, A.-N., Brandt, A., 2002. A common but undescribed huge species of Eusirus from Antarctica (Crustacea, Amphipoda, Eusiridae). Mitteilungen aus dem Hamburgischen zoologische Museum und Institut.

Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data.

Arntz, W.E., Brey, T., Gallardo, V.A., 1994. Antarctic zoobenthos, in: Oceanography and Marine Biology: An Annual Review 1994. Ansell AD, Gibson RN, Barnes M, pp. 241–304.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29. https://doi.org/10.1038/75556

Baird, H.P., Miller, K.J., Stark, J.S., 2011. Evidence of hidden biodiversity, ongoing speciation and diverse patterns of genetic structure in giant Antarctic amphipods. Molecular Ecology 20, 3439–3454. https://doi.org/10.1111/j.1365-294X.2011.05173.x

Bak, H.J., Beintema, J.J., 1987. Panulirus interruptus hemocyanin. The elucidation of the complete amino acid sequence of subunit a. Eur J Biochem 169, 333–348. https://doi.org/10.1111/j.1432-1033.1987.tb13616.x

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Beers, J.M., Borley, K.A., Sidell, B.D., 2010. Relationship among circulating hemoglobin, nitric oxide synthase activities and angiogenic poise in red- and white-blooded Antarctic notothenioid fishes. Comp. Biochem. Physiol., Part A Mol. Integr. Physiol. 156, 422–429. https://doi.org/10.1016/j.cbpa.2010.03.027

Brandão, S.N., Sauer, J., Schön, I., 2010. Circumantarctic distribution in Southern Ocean benthos? A genetic test using the genus Macroscapha (Crustacea, Ostracoda) as a model. Mol Phylogenet Evol 55, 1055–1069. https://doi.org/10.1016/j.ympev.2010.01.014

Burmester, T., 1999. Identification, molecular cloning, and phylogenetic analysis of a non-respiratory pseudo-hemocyanin of Homarus americanus. J Biol Chem 274, 13217–13222. https://doi.org/10.1074/jbc.274.19.13217

Cavanaugh, J.E., 1997. Unifying the derivations for the Akaike and corrected Akaike information criteria. Statistics & Probability Letters 33, 201–208.

Chang, E.S., Mykles, D.L., 2011. Regulation of crustacean molting: A review and our perspectives. General and Comparative Endocrinology 172, 323–330. https://doi.org/10.1016/j.ygcen.2011.04.003

Chapelle, G., 2002. Antarctic and Baikal amphipods: a key for understanding polar gigantism.

Chapelle, G., Peck, L.S., 1999. Polar gigantism dictated by oxygen availability. Nature 399, 114–115. https://doi.org/10.1038/20099

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., Flouri, T., 2020. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Mol Biol Evol 37, 291–294. https://doi.org/10.1093/molbev/msz189

Dauby, P., Scailteur, Y., Chapelle, G., De Broyer, C., 2002. Potential impact of the main benthic amphipods on the eastern Weddell Sea shelf ecosystem (Antarctica), in: Arntz, W.E., Clarke, A. (Eds.), Ecological Studies in the Antarctic Sea Ice Zone: Results of EASIZ Midterm Symposium. Springer, Berlin, Heidelberg, pp. 45–50. https://doi.org/10.1007/978-3-642-59419-9_7

Dauby, P., Scailteur, Y., De Broyer, C., 2001. Trophic diversity within the eastern Weddell Sea amphipod community. Hydrobiologia 443, 69–86. https://doi.org/10.1023/A:1017596120422

De Broyer, C., Jazdzewski, K., 1996. Biodiversity of the Southern Ocean: towards a new synthesis for the Amphipoda (Crustacea). Boll Mus Civ Sta nat Verona 20, 547–568.

De Broyer, C., Jazdzewski, K., 1993. Contribution to the marine biodiversity inventory. A checklist of the Amphipoda (Crustacea) of the Southern Ocean. Documents de Travail de l'Institut Royal des Sciences Naturelles de Belgique 73, 1–174.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with

R

high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucl. Acids Res. 39, W29–W37. https://doi.org/10.1093/nar/gkr367

Gao, Y., Wei, J., Yuan, J., Zhang, X., Li, F., Xiang, J., 2017. Transcriptome analysis on the exoskeleton formation in early developmetal stages and reconstruction scenario in growth-moulting in Litopenaeus vannamei. Scientific Reports 7, 1098. https://doi.org/10.1038/s41598-017-01220-6

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31, 3784–3788. https://doi.org/10.1093/nar/gkg563

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotech 29, 644–652. https://doi.org/10.1038/nbt.1883

Gutt, J., 2008. The Expedition ANTARKTIS-XXIII/8 of the Research Vessel "Polarstern," Reports on Polar and Marine Research. AWI, Alfred-Wegener-Institut für Polar- und Meeresforschung.

Hagner-Holler, S., Kusche, K., Hembach, A., Burmester, T., 2005. Biochemical and molecular characterisation of hemocyanin from the amphipod Gammarus roeseli: complex pattern of hemocyanin subunit evolution in Crustacea. J Comp Physiol B 175, 445–452. https://doi.org/10.1007/s00360-005-0012-4

Held, C., Wägele, J.W., 2005. Cryptic speciation in the giant Antarctic isopod *Glyptonotus antarcticus* (Isopoda, Valvifera, Chaetiliidae). scimar 69, 175–181. https://doi.org/10.3989/scimar.2005.69s2175

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Jekel, P.A., Bak, H.J., Soeter, N.M., Vereijken, J.M., Beintema, J.J., 1988. Panulirus interruptus hemocyanin. The amino acid sequence of subunit b and anomalous behaviour of subunits a and b on polyacrylamide gel electrophoresis in the presence of SDS. Eur J Biochem 178, 403–412. https://doi.org/10.1111/j.1432-1033.1988.tb14464.x

Jekel, P.A., Neuteboom, B., Beintema, J.J., 1996. Primary structure of hemocyanin from Palinurus vulgaris. Comp Biochem Physiol B Biochem Mol Biol 115, 243–246. https://doi.org/10.1016/0305-0491(96)00119-8

Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F.A., Zdobnov, E.M., 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. 47, D807–D811. https://doi.org/10.1093/nar/gky1053

Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320. https://doi.org/10.1093/molbev/msn067

Levin, L.A., Gage, J.D., 1998. Relationships between oxygen, organic matter and the diversity of bathyal macrofauna. Deep Sea Research Part II: Topical Studies in Oceanography 45, 129–163. https://doi.org/10.1016/S0967-0645(97)00085-4

Loerz, A., Maas, E., Linse, K., Coleman, C.O., 2009. Do circum-Antarctic species exist in peracarid Amphipoda? A case study in the genus Epimeria Costa, 1851 (Crustacea, Peracarida, Epimeriidae). ZooKeys 18, 91–128. https://doi.org/10.3897/zookeys.18.103

MacManes, M.D., 2018. The Oyster River Protocol: a multi-assembler and kmer approach for de novo transcriptome assembly. PeerJ 6, e5428. https://doi.org/10.7717/peerj.5428

OBIS, 2020. https://portal.obis.org/taxon/237660 [WWW Document]. Eusirus giganteus Andres, Lörz & Brandt, 2002. URL https://portal.obis.org/taxon/237660

Oellermann, M., Lieb, B., Pörtner, H.-O., Semmens, J.M., Mark, F.C., 2015. Blue blood on ice: modulated blood oxygen transport facilitates cold compensation and eurythermy in an Antarctic octopod. Front Zool 12. https://doi.org/10.1186/s12983-015-0097-x

Poynton, H.C., Hasenbein, S., Benoit, J.B., Sepulveda, M.S., Poelchau, M.F., Hughes, D.S.T., Murali, S.C., Chen, S., Glastad, K.M., Goodisman, M.A.D., Werren, J.H., Vineis, J.H., Bowen, J.L., Friedrich, M., Jones, J., Robertson, H.M., Feyereisen, R., Mechler-Hickson, A., Mathers, N., Lee, C.E., Colbourne, J.K., Biales, A., Johnston, J.S., Wellborn, G.A., Rosendale, A.J., Cridge, A.G., Munoz-Torres, M.C., Bain, P.A., Manny, A.R., Major, K.M., Lambert, F.N., Vulpe, C.D., Tuck, P., Blalock, B.J., Lin, Y.-Y., Smith, M.E., Ochoa-Acuña, H., Chen, M.-J.M., Childers, C.P., Qu, J., Dugan, S., Lee,

S.L., Chao, H., Dinh, H., Han, Y., Doddapaneni, H., Worley, K.C., Muzny, D.M., Gibbs, R.A., Richards, S., 2018. The Toxicogenome of Hyalella azteca: A Model for Sediment Ecotoxicology and Evolutionary Toxicology. Environ Sci Technol 52, 6009–6022. https://doi.org/10.1021/acs.est.8b00837

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A., Finn, R.D., 2012. The Pfam protein families database. Nucleic Acids Research 40, D290–D301. https://doi.org/10.1093/nar/gkr1065

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032

Raupach, M.J., Wägele, J.-W., 2006. Distinguishing cryptic species in Antarctic Asellota (Crustacea: Isopoda) - a preliminary study of mitochondrial DNA in Acanthaspidia drygalskii. Antarctic Science 18, 191–198. https://doi.org/10.1017/S0954102006000228

Redmond, J.R., 1955. The respiratory function of hemocyanin in crustacea. Journal of Cellular and Comparative Physiology 46, 209–247. https://doi.org/10.1002/jcp.1030460204

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S.D., Mungall, K., Lee, S., Okada, H.M., Qian, J.Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, Y.S., Newsome, R., Chan, S.K., She, R., Varhol, R., Kamoh, B., Prabhu, A.-L., Tam, A., Zhao, Y., Moore, R.A., Hirst, M., Marra, M.A., Jones, S.J.M., Hoodless, P.A., Birol, I., 2010. De novo assembly and analysis of RNA-seq data. Nature Methods 7, 909–912. https://doi.org/10.1038/nmeth.1517

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Spicer, J.I., Morley, S.A., 2019. Will giant polar amphipods be first to fare badly in an oxygen-poor ocean? Testing hypotheses linking oxygen to body size. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20190034. https://doi.org/10.1098/rstb.2019.0034

Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564–577. https://doi.org/10.1080/10635150701472164

Terwilliger, N.B., Dangott, L., Ryan, M., 1999. Cryptocyanin, a crustacean molting protein: Evolutionary link with arthropod hemocyanins and insect hexamerins. PNAS 96, 2013–2018. https://doi.org/10.1073/pnas.96.5.2013

Terwilliger, N.B., Ryan, M., Phillips, M.R., 2006. Crustacean hemocyanin gene family and microarray studies of expression change during eco-physiological stress. Integrative and Comparative Biology 46, 991–999. https://doi.org/10.1093/icb/icl012

Van Wormhoudt, A., Sellos, D., Donval, A., Plaire-Goux, S., Le Moullac, G., 1995. Chymotrypsin gene expression during the intermolt cycle in the shrimpPenaeus vannamei (Crustacea; Decapoda). Experientia 51, 159–163. https://doi.org/10.1007/BF01929362

Verberk, W.C.E.P., Atkinson, D., 2013. Why polar gigantism and Palaeozoic gigantism are not equivalent: effects of oxygen and temperature on the body size of ectotherms. Functional Ecology 27, 1275–1285. https://doi.org/10.1111/1365-2435.12152

Verheye, M.L., Martin, P., Backeljau, T., D'Acoz, C.D., 2016. DNA analyses reveal abundant homoplasy in taxonomically important morphological characters of Eusiroidea (Crustacea, Amphipoda). Zoologica Scripta 45, 300–321. https://doi.org/10.1111/zsc.12153

Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285. https://doi.org/10.1007/s12064-012-0162-3

Whiteley, N.M., Taylor, E.W., Clarke, A., Haj, A.J.E., 1997. Haemolymph oxygen transport and acid-base status in Glyptonotus antarcticus Eights. Polar Biol 18, 10–15. https://doi.org/10.1007/s003000050153

Woods, H.A., Moran, A.L., Arango, C.P., Mullen, L., Shields, C., 2009. Oxygen hypothesis of polar gigantism not supported by performance of Antarctic pycnogonids in hypoxia. Proceedings of the Royal Society B: Biological Sciences 276, 1069–1075. https://doi.org/10.1098/rspb.2008.1489

Funding Statement

This work was supported by the Italian Program of Antarctic Research (grant number PNRA16_00099).

Availability of Data and Materials

Raw sequence data was deposited in the NCBI Sequence

R

Read Archive (SRA) database with the accession ID SRR13341529, under the BioProject PRJNA689112. The assembled contigs were deposited in the Transcriptome Shotgun Assembly database under the accession ID GIYD00000000.

Acknowledgements

We would like to thank Simone Canese e Erica Carlig for providing the specimen analyzed in this work.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.