Comparative population genomics provide new insight into the evolutionary history and adaptive potential of World Ocean krill

Marvin Choquet[1,2], Felix Lenner[1,3], Arianna Cocco[1], Gaëlle Toullec[4], Erwan Corre[5], Jean-Yves Toullec[6] and Andreas Wallberg[1,*]

[1] Department of Medical Biochemistry and Microbiology, Uppsala University; Husargatan 3, 751 23 Uppsala, Sweden.

[2] Natural History Museum, University of Oslo, Oslo, Norway.

[3] Department of Immunology, Genetics and Pathology, Uppsala University; Husargatan 3, 751 23 Uppsala, Sweden.

[4] Laboratory for Biological Geochemistry, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

[5] CNRS, Sorbonne Université, FR 2424, ABiMS Platform, Station Biologique de Roscoff, Roscoff, France

[6] Sorbonne Université, CNRS, UMR 7144, AD2M, Station Biologique de Roscoff, Roscoff, France

* Correspondence: andreas.wallberg@imbim.uu.se; Tel.: +46184714231

Abstract: Genetic variation is instrumental for adaptation to new or changing environments but it is poorly understood how it is structured and contributes to adaptation in pelagic species without clear barriers to gene flow. Here we use extensive transcriptome datasets from 20 krill species collected across the Atlantic, Indian, Pacific and Southern Oceans and compare genetic variation both within and between species across thousands of genes. We resolve phylogenetic interrelationships and uncover genomic evidence in support of elevating the cryptic *Euphausia similis* var. *armata* into species. We estimate levels of genetic variation and rates of adaptive protein evolution among species and find that these are comparably low in large Southern Ocean species endemic to cold environments, including the Antarctic krill *Euphausia superba*, suggesting their adaptive potential to rapid climate change may also be low. We uncover hundreds of candidate loci with signatures of adaptive divergence between krill native to cold and warm waters and identify candidates for cold-adaptation that have also been detected in Antarctic fish, including genes that govern thermal

reception such as *TrpA1*. Our results suggest parallel genetic responses to similar selection pressures across Antarctic taxa and provide new insights into the adaptive potential of important zooplankton that are already strongly affected by climate change.

Keywords: krill; climate change; population genomics; comparative genomics; genetic adaptation; evolution

Introduction

The world's oceans have absorbed over 90% of anthropogenic greenhouse gas energy released over the last century and warmed by about 1°C (Intergovernmental Panel on Climate Change 2014). This has strongly impacted pelagic species, causing rapid poleward shifts and accelerated phenologies in fish and zooplankton, such as jellyfish, salps, copepods and krill (Richardson 2008; Poloczanska et al. 2013; Poloczanska et al. 2016; Ratnarajah et al. 2023), and threatening to destabilize important food webs and ecosystem services (Doney et al. 2012; Baxter and Laffoley 2016). Genetic adaptation could be key to sustain marine populations, promoting resilience by targeting genes encoding traits such as growth, reproductive timing and thermal tolerance (Hoffmann and Sgrò 2011; Dam 2013). However, despite their pivotal roles in marine ecosystems, insight into what genes and variation that contribute to fitness and adaptation is still missing in zooplankton (Dam 2013; Bucklin et al. 2018).

Krill (Euphausiacea; “euphausiids”; 86 spp.) are crustacean macrozooplankton and grazers of phytoplankton primary production or smaller zooplankton (Russell et al. 1969). As important food for fish, mammals and birds, they play critical roles in transferring nutrients to higher trophic levels in marine ecosystems (Tarling 2010; McBride et al. 2014; Siegel 2016; Johnston et al. 2022) and superabundant species like the Antarctic krill *Euphausia superba* contribute to global biogeochemical systems through the biological carbon pump (Cavan et al. 2019). Different species occur throughout tropical, temperate and arctic ecosystems and their biogeography depend on physiological thermal tolerance, oceanographic conditions and nutrient availability (Russell et al. 1969; Cimino et al. 2020). Across these environments, euphausiids have evolved divergent life cycle strategies: low-latitude species associated with nutrient-rich upwellings or near-shore habitats tend to develop and mature

quickly, be short-lived (<1 year or 1–2 years), breed multiple times or continuously throughout the year and have high productivity rates (Siegel 2000). High-latitude species instead deploy strategies to cope with extended periods of dark, cold and nutritionally adverse conditions, and are characterized by slow rates of development, extended longevity (e.g. 2 years or more), larger body sizes that build long-lasting lipid stores, and have short annual reproductive seasons and low productivity (Falk-Petersen et al. 2000; Siegel 2000; Siegel 2016). Moreover, many krill are stenothermal and occupy habitats with narrow latitudinal and thermal ranges. The Antarctic *E. superba* and *E. crystallorophias* likely split from other Southern Ocean species about 25–27 MYA, during the formation of the Antarctic Circumpolar Current (Zane and Patarnello 2000) and have since adapted to the extreme Antarctic environments. Today, *E. superba* inhabits cold waters ranging between −2.0°C to +4.0°C (Siegel 2016), while the neritic Ice krill *E. crystallorophias* occurs within an even more narrow thermal envelope (−1.8°C to 0°C) (Cuzin-Roudy et al. 2014). Short-term stress tests (e.g. CT50 assays) show they are sensitive to high temperatures (Cascella et al. 2015; Huenerlage et al. 2016), which is common among Antarctic ectotherms (Peck et al. 2014). In contrast, North Atlantic krill such as *Meganyctiphanes norvegica* ("the Northern krill") and several *Thysanoessa* have thermal ranges spanning 2–15°C or more and occur from the colder Arctic Ocean to the warmer Gulf of Maine (Russell et al. 1969; Tarling 2010; Ollier et al. 2018), indicating greater thermal tolerance in these widely distributed euphausiids.

Similar to many other Antarctic species, *E. superba* and *E. crystallorophias* appear to have reduced capacity to upregulate inducible and protective heat shock proteins (Hsps) in response to increased temperature (Cascella et al. 2015; Huenerlage et al. 2016; Peck 2016; Chen et al. 2018; Toullec et al. 2020). Adaptation to cold environments could commonly involve fundamental genetic and physiological alterations and trade-offs that are maladaptive under rapidly warming temperatures (Pörtner et al. 2007), and it is therefore concerning that both Antarctic and Arctic krill are declining or shifting to higher latitudes due to climate change (Atkinson et al. 2019; Edwards et al. 2021). Tropical krill may also live close to their thermal limits and be sensitive to change, as suggested in other marine invertebrates (Peck et al. 2014), although these species are not regularly monitored or studied. The long-term ability of krill, or other planktonic organisms, to resist temperature variation is difficult to

estimate experimentally and it is not clear if any krill has enough physiological tolerance to track continued ocean warming.

Evolutionary adaptation through natural selection on adaptive variation could be crucial for coping with climate change. Knowledge about the genetic mechanisms that contribute to adaptation can be gained through genome-scale comparisons between populations or species adapted to different local conditions (Savolainen et al. 2013; Meek et al. 2023). In marine taxa, scans for adaptive divergence between populations have uncovered loci involved in reproductive timing in Atlantic herring (Han et al. 2020) and bleaching in *Acropora* corals (Fuller et al. 2020), while comparative analyses among *Mytilus* mussel species have detected candidate genes evolving under divergent temperature-dependent selection (Popovic and Riginos 2020). Population genetic theory suggests that zooplankton could have high adaptive potential (Peijnenburg and Goetze 2013): i) they typically have large populations with many reproductive individuals (i.e. a high effective population size, or "N_e") that can maintain or generate much variation to select from; ii) due to large N_e, they are expected to be comparably unaffected by genetic drift, i.e. stochastic changes in allele frequencies between generations that may interfere with selection; iii) many species have short generation times and high reproductive rates, amenable for adapting to rapid changes. However, empirical support for these predictions is still missing as various challenges have hindered molecular research on marine zooplankton.

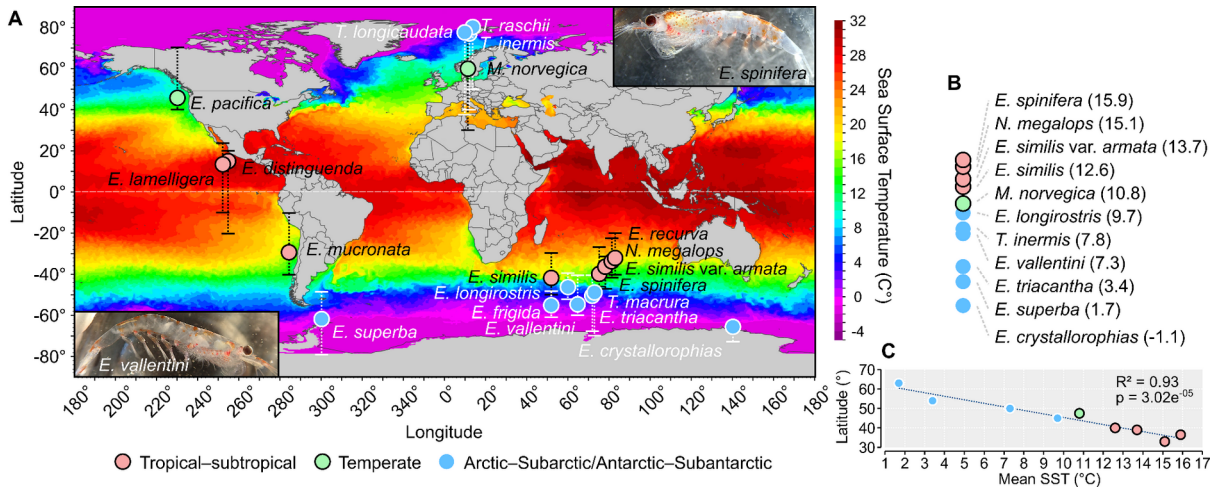
One of these challenges resides in the ambiguity of taxonomy across zooplankton taxa (Bucklin et al. 2007; Bailey et al. 2016; Choquet et al. 2018). Thus, the widespread occurrence of similar-looking or cryptic species has led to under-estimation of zooplankton diversity, although large-scale barcoding and metabarcoding surveys have started to address this issue (Bucklin et al. 2016; Bucklin et al. 2021). In krill, cryptic variation was reported within *Stylocheiron affine* (Wiebe et al. 2016) and the specific status of *Euphausia similis* and its variety *armata* (Hansen 1911) remain unclear. Furthermore, accessing zooplankton species potentially adapted to polar environments can be a logistical challenge when they live in remote locations, like the Antarctic krill, and may therefore discourage attempts at conducting advanced studies of local adaptation (Bucklin et al. 2018). Besides, such studies require genome-wide information and yet only a handful of genomes have been

sequenced so far across zooplankton, mostly due to very large genome sizes common in many marine zooplankton, which, combined with a small body size, prohibit whole genome sequencing (Bucklin et al. 2018). Krill have huge and repetitive genomes that, among six known species, range between 11–48 Gb (Jeffery 2012), making genome-wide analyses very challenging. To date, only the 48 Gb genome of the Antarctic krill has been assembled and population-scale scans recovered few potentially adaptive variants (Shao et al. 2023), possibly due to the highly panmictic life history of this species (Bortolotto et al. 2011; Deagle et al. 2015; Shao et al. 2023). Overcoming these obstacles is critical for monitoring and forecasting zooplankton responses to climate change. While the constant advances of sequencing technologies and protocols promise to ease whole genome sequencing of challenging organisms in the future, transcriptomics offers many possibilities to advance insight into zooplankton ecology (Lenz et al. 2021).

Here we produce, for the first time, a comprehensive comparative dataset at the genome-scale for krill, spanning 20 species from the Atlantic, Indian and Southern Oceans (fig. 1). Euphausiids are an attractive model for zooplankton because of their importance for ocean ecosystems and distribution across all latitudes. Moreover, they belong to a comparatively small and ecologically well-characterized order of zooplankton and are not extremely divergent, facilitating association between genes and phenotypes through comparative analyses. We sidestep the issue of prohibitively large genomes of these non-model organisms by using RNA-seq data to analyze genetic variation in expressed protein-coding genes (De Wit et al. 2015). We aim to elucidate the evolutionary history of krill by comparing variation between species, uncovering candidate genes for adaptation to cold or warm environments and testing if krill have high adaptive potential, as well as identify factors that may determine it. In addition, we provide further evidence for how thermal tolerance varies between species and associate with the environment.

5

Fig. 1.



The geographic distribution of 20 surveyed krill species. (*A*) Circles indicate locations of one or more collected specimens (colors represent the thermal conditions associated with the range of each species). The basin-specific latitudinal range of each species is indicated with bars. Ranges are from Russel & Yonge (Russell et al. 1969). Global Sea Surface Temperature readouts are a daily snapshot from OISST V2 (2022-06-01) from ERDDAP: https://www.ncei.noaa.gov/erddap/griddap/. Photos of Indian Ocean krill collected in 2019. (*B*) Approximate mean SST of the ranges of 11 species used for analyses of thermal tolerance or intraspecific variation. (*C*) Linear regression correlation between habitat latitudinal midpoint and mean SST (n=9 species).

Results

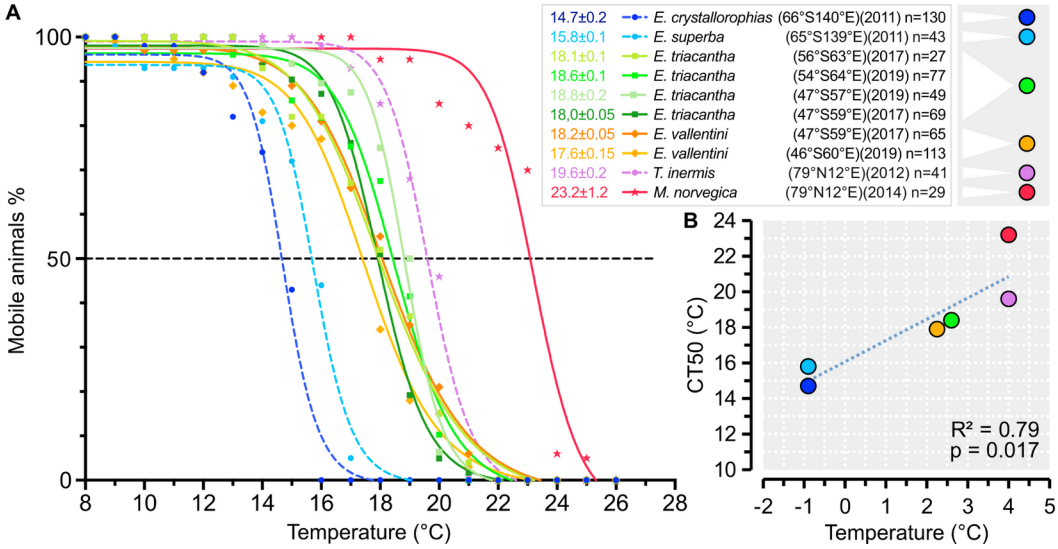
Thermal tolerance in krill is associated with habitat temperature

We used CT50 mobility assays to characterize and compare thermal tolerance in six krill species sampled from the Southern, Atlantic and Indian Oceans. The two Antarctic species—the Ice krill (*E. crystallorophias*) and the Antarctic krill (*E. superba*)—have significantly lower CT50 thresholds compared to the two subantarctic *E. vallentini* and *E. triacantha* (fig. 2*A*). The average CT50 for *E. vallentini* is 17.9±0.42°C. This may be an underestimate because specimens were fragile and appeared impacted by fishing (reflected in the rapidly subsiding high sigmoid plateau and the flatter slope). CT50 for *E. triacantha* was 18.37±0.39, which appeared to be at better health. The widespread Northern krill *Meganyctiphanes norvegica* showed a record CT50 with a value of nearly 23°C for animals that were fished in 4°C waters of the high Arctic Spitsbergen (same locality as *T. inermis*),

6

clearly illustrating extreme tolerance to temperature in a species ranging from the warm Mediterranean Sea to cold Arctic waters (Tarling 2010). CT50s are strongly correlated with ambient temperatures (fig. 2*B*; supplementary fig. S1), underscoring that thermal tolerances of krill are linked to the respective temperatures of their native environments.

Fig. 2.



Thermal tolerance in Southern Ocean and Atlantic Ocean krill. (*A*) The figure shows all the results acquired over several years during campaigns in Antarctica, Arctic and Indian Ocean. Three species (dashed curves) have been published before (Cascella et al. 2015; Huenerlage et al. 2016), while *E. vallentini* (orange), *E. triacantha* (green) and *M. norvegica* (red) are new to this study. The white legend indicates species/experiment-specific calculated CT50 values, fishing coordinates, years and sample sizes. (*B*) Linear regression correlation between CT50 and the measured ambient sea temperature at the experimental locations. The mean CT50 of all experiments was used for *E. triacantha* (n=4) and *E. vallentini* (n=2).

Genome-scale datasets to study variation and adaptation

We compiled RNA-seq data spanning 124 transcriptomes from 20 species (102 transcriptomes from 17 species are new for this study). Morphological and molecular species identifications agreed for all samples with barcodes in the MetaZooGene database (Table S1). After removing redundant and non-coding isoforms, the reference transcriptomes contained 20–53 K coding transcripts per species (Table S2). BUSCO analyses indicated high completeness (median=92.5%) and low duplication levels (median=3.3%), amenable for tracing evolutionary patterns using base-level substitutions and polymorphisms. We measured intraspecific genetic variation in nine species, using up to 20 specimens per species and detected between 260–2,200 K high-quality SNPs across 12–22 K genes (Table S3).

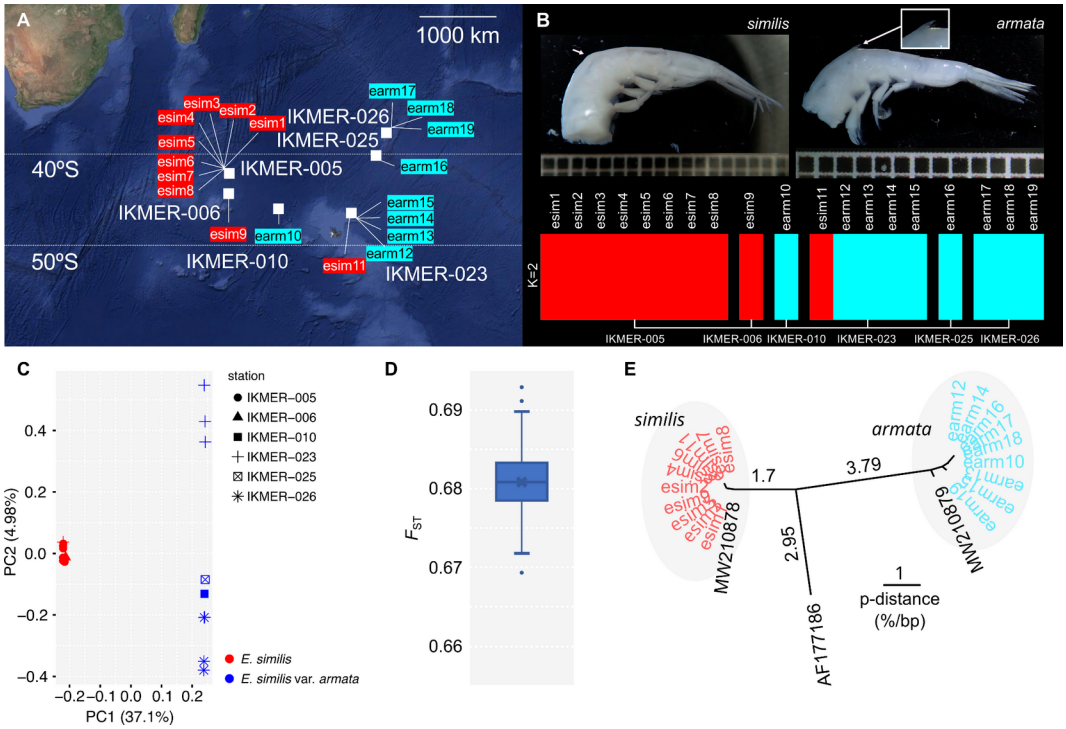
7

Genetic divergence between Indian Ocean *Euphausia similis* and *Euphausia similis* var. *armata* suggests they are different evolutionary lineages

The Indian Ocean sampling region spanned multiple waterfronts and we interrogated samples from this region for population structure. Two out of three temperate–subtropical species showed structure. We sequenced 19 specimens of *E. similis* from the Indian Ocean (fig. 3*A*). Nine specimens collected mostly to the east morphologically matched the *Euphausia similis* var. *armata* variety ("*armata*"), having an accessory abdominal spine of variable size (fig. 3*B*; supplementary fig. S2*A*) (Hansen 1911; Baker et al. 1990). *E. similis* without the spines were mostly found to the west (figs. 3*A–B*; supplementary fig. S2*B*; "*similis*"). We initially detected strong structuring between *similis* and *armata* using ancestry analysis and PCA (fig. 3*B–C*), which was also supported by high F_{ST} (0.68) (fig. 3*D*). We next assembled 500–800 bp mtCOI fragments and assessed divergence. Sequences segregated consistently with the nuclear data and were ~6.2% divergent, clustering with different MetaZooGene *similis* barcodes derived from southern Atlantic Ocean samples (fig. 3*E*; supplementary fig. S3). Because of the consistent molecular and phenotypic differences between *similis* and *armata*, we treated them as separate evolutionary lineages in subsequent analyses. We detected comparably weak north–south structuring in our limited *N. megalops* sample (F_{ST}=0.06), but did not observe any structure in the subtropical *E. spinifera*, nor in the subantarctic *E. vallentini* and *E. triacantha* collected from the Antarctic Polar Front (supplementary fig. S4).

8

Fig. 3.



Genetic structure between *Euphausia similis* ("esim") and *Euphausia similis* var. *armata* ("earm") from the Indian Ocean. (*A*) Map of sampling stations (n=6). (*B*) Top: representative tails of *E. similis* (without a dorsal spine) and *E. similis* var. *armata* (with a spine on the third abdominal segment). Arrows indicate location of the spine (if present). Bottom: genetic ancestry and admixture of samples (K=2 ancestral groups). (*C*) The genetic structure among 19 samples using PCA (n=84,436 unlinked nuclear SNPs). (*D*) Weighted F_{ST} estimated from 1,000 random resamples drawing 1 SNP per gene (n=6,074 genes) (Choquet et al. 2019). (*E*) A mitochondrial COI neighbor-joining gene tree inferred from uncorrected pairwise genetic distances. Distances are shown for the major internal branches (% differences per bp). Sequences from MetaZooGene are: AF177186 (north-east of Japan; Kenji Taki; personal communication); MW210878 and MW210879 (the southern Atlantic Ocean).

Two-fold difference in genetic diversity among nine species of krill

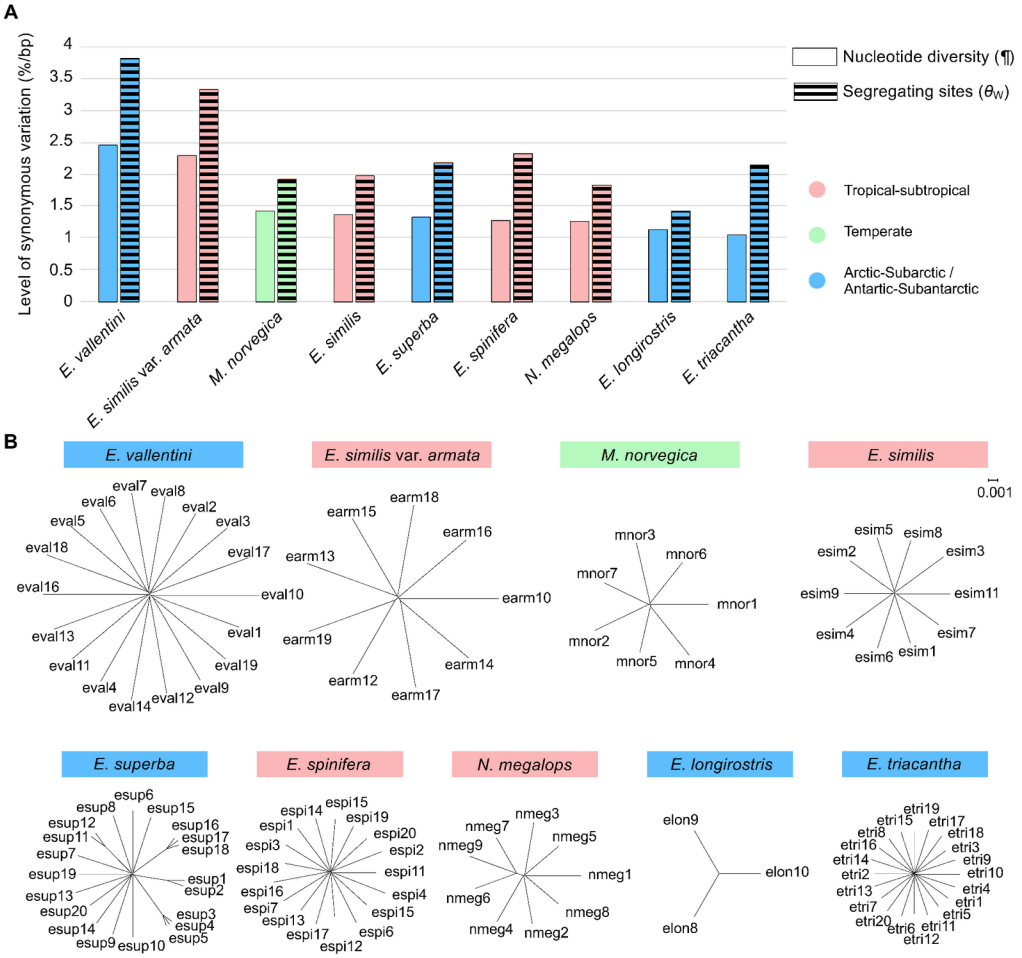
We took transcriptome-wide levels of genetic variation as indicators of genetic diversity and studied how it varied among gene regions and species (supplementary figs. S5–S6). Overall, we detected about 1.6× more variation in untranslated regions (UTRs) compared to coding regions (average θ_w=1.4% vs. 0.9%), and 3.1–5.6× more variation at synonymous sites compared to non-synonymous sites (fig. 4; supplementary fig. S5). These patterns likely reflect both direct and linked purifying selection, reducing variation around functionally important sites (Cvijović et al. 2018). Restricting

9

measurements to only genes with established orthology (see below), we observe lower but strongly correlated levels of variation (~0.78×; r^2=0.99; supplementary fig. S7), indicating the full gene sets are less conserved than orthologs overall but representative of variation among species. We let synonymous variants represent selectively neutral variation and be indicative of effective population size (N_e) and demographic history. All species have variation surpassing 1% per-base, the most diverse being *E. vallentini* (π_S=2.5%; θ_w=3.83%), while *E. triacantha* is the least diverse (π_S=1.1%) (fig. 4). Assuming similar mutation rates among species, estimates of N_e indicate *E. vallentini* (3.6 M) and *E. similis* var. *armata* (3.2 M) have the largest effective population sizes (Table 1). Our *armata* sample has 1.7× as much variation as *similis sensu stricto*, further suggesting these lineages follow different evolutionary trajectories. Tajima's D (D_T) is negative across all nine species, ranging from -1.2 to -1.9 (Table 1), indicating excess low-frequency variants compared to expectations under neutrality, which could result from processes such as selective sweeps or recent population expansions. Neither intrinsic life history traits (e.g. body or larval size) or environmental parameters (e.g. temperature, latitude or range) are strong predictors of π_S or D_T (i.e. no correlation $p<0.1$; see Table S9 for parameters).

10

Fig. 4.



Genetic variation at synonymous sites in nine krill species. (*A*) Nucleotide diversity (π_S) and Watterson's theta (θ_{wS}) estimated from 74 k to 709 k synonymous SNPs per species, after correction for numbers of accessible sites. (*B*) Neighbour-joining trees showing variation and structure across species. Branch lengths are scaled as the average genetic distance per base between samples (π_S), estimated at synonymous loci and corrected for accessible sites and drawn at the same scale.

Comparative analyses establish comprehensive gene orthologies and a robust species tree

To enable direct comparative analyses of molecular evolution, we inferred gene orthology among the 20 krill species and seven outgroups (supplementary materials) and reconstructed a species tree. We detected 103,292 orthogroups (OGs) with genes from two or more krill species, but only 880 OGs contained single-copy orthologs spanning all 20 species. We produced two subsets of OGs, balancing species missingness against the number of OGs. The first dataset spanned 2,280 OGs, >1 million amino acid positions and 18+ krill species (Table S4), and was used to reconstruct a fully supported phylogeny (fig. 5*A*). The topology is mostly consistent with inferences of krill interrelationships by

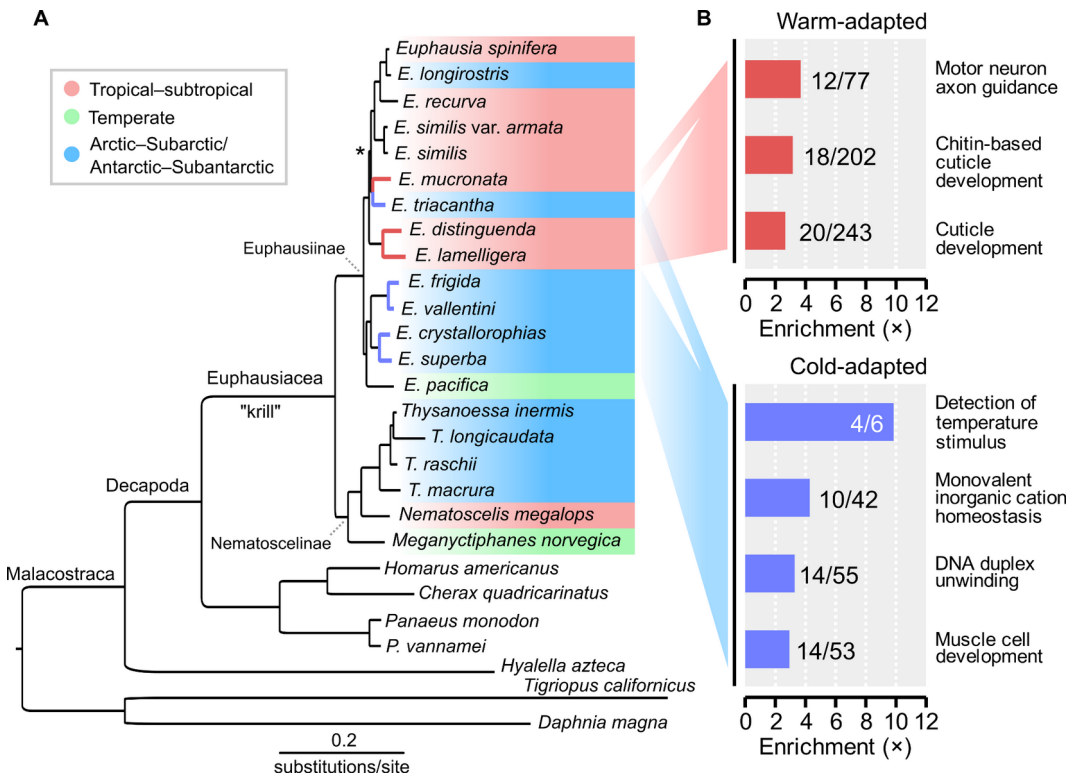
11

Vereshchaka and co-workers based on morphology and four molecular markers: *Meganyctiphanes* + *Nematoscelis* + *Thysanoessa* form the monophyletic clade *Nematoscelinae*, which is the sister taxon of *Euphausiinae* (Vereshchaka et al. 2019). Groupings within *Euphausia* confirm a monophyletic Southern Ocean "*Euphausia superba* group" including *frigida*, *vallentini*, *crystallorophias* and *superba*, while our clade with *spinifera*, *longirostris*, *similis*, *similis* var. *armata* and *triacantha* also include *recurva* and *mucronata*, mixing other species groups.

Resolving phylogenetic interrelationships is necessary for understanding the evolution of euphausiid genes, traits and ecologies and numerous attempts have been made using morphology or sparse molecular datasets such as mitochondrial markers (see Bucklin et al. 2007; Vereshchaka et al. 2019; Schram and Koenemann 2022 for reviews). Our reconstruction is the first phylogenomic species tree for krill and shows that *E. triacantha* is only distantly related to other *Euphausia* of the Southern Ocean. *E. triacantha* and *E. longirostris* have likely expanded into cold environments independently from each other, and from other *Euphausia*. Our tree implies that krill have frequently diversified into different climates, offering the opportunity for studying parallel events to uncover genetic adaptations.

12

250 **Fig. 5.**



251 Phylogenomic inference of interrelationships and molecular evolution in krill. (*A*) A species tree inferred from 2,280
252 orthologous protein alignments and constructed from 1,000 ultrafast bootstrap replicates (JTTDCMut+F+I+G4 model; lnL=-
253 11,173,074) (Hoang et al. 2018). All nodes in the majority-rule consensus tree but one have 100% bootstrap support (*= 99%
254 support). Foreground branches used in branch-site tests to detect positively selected genes are highlighted (blue=five cold-
255 adapted species; red=three warm-adapted species). (*B*) Statistically enriched gene ontologies (GO) among candidate genes
256 for positive selection (False Discovery Rate <0.05) in the "cold-adapted contrast" (n=629 candidates; LRT<0.01 with ≥1
257 BEB95% sites) and "warm-adapted" ontologies (n=383 candidates; LRT<0.05 with ≥1 BEB95% sites). Numbers beside bars
258 indicate genes in the target set (i.e. selection candidates) vs. genes in the background test set. Redundant GO terms were
259 removed with *Revigo* (Supek et al. 2011).

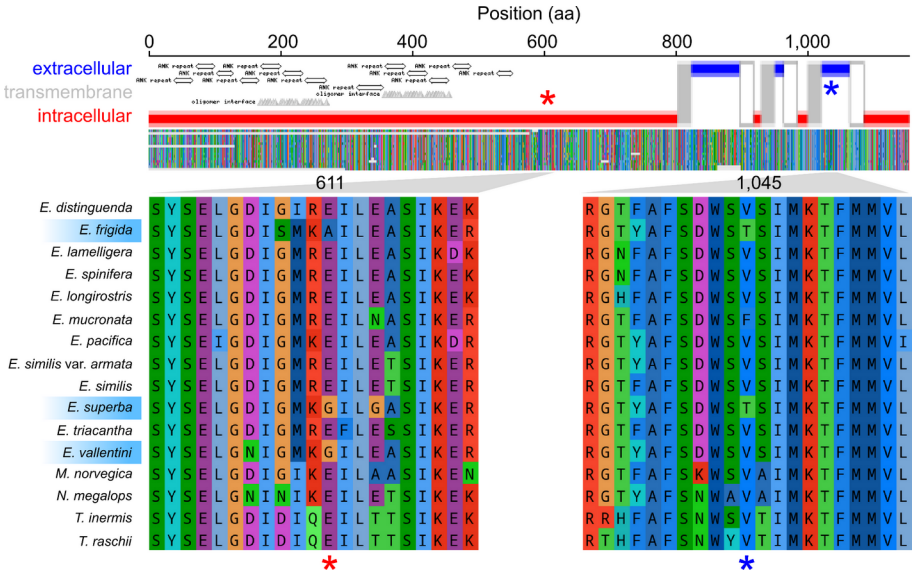
260 **Scans for positive selection reveal candidate genes for environmental adaptation**

261 We aimed to detect genes that may underlie episodic adaptation to cold or warm climates in krill. To
262 this end, we applied branch-site tests to the OG dataset spanning >13,000 OGs (>5 million codon
263 positions; Table 2; Table S4) to detect genes with locally elevated ω between species associated with
264 different environments (Yang and Nielsen 2002). In the first contrast, we compared five Southern
265 Ocean cold-adapted *Euphausia* against the rest of the phylogeny (*E. superba*, *E. crystallorophias*, *E.*
266 *frigida*, *E. vallentini* and *E. triacantha*; fig. 5*A*; Table S4). We identified 832 gene candidates (LRT

13

267 p<0.01; n=629 genes with *Drosophila* homologs; Table S5), containing 1,800 selected codons. A gene
268 ontology (GO) enrichment test found enrichment for genes involved in detecting temperature among
269 candidates (fig. 5*B*; Table S6), including homologs of *Transient receptor potential cation channel A1*
270 (*TrpA1*), *painless*, *subdued* and *straightjacket*. Both *TrpA1* and *painless* encode TRPA ion channels
271 (Hwang et al. 2012; Himmel and Cox 2020), which are activated by temperature and various ligands.
272 *TrpA1* is linked to sensation of either harmfully cold or warm temperatures across the animal tree of
273 life (Zhang et al. 2022), while *painless* is associated with heat sensation (Dhaka et al. 2006; Himmel
274 and Cox 2020). Crustacean TRP evolution is characterized by widespread gene duplication (Kozma et
275 al. 2020), and phylogenetic analysis reveals multiple paralogs also in krill, and suggest the krill *TrpA1*
276 OG is orthologous to decapod *TrpA1-like*, while the *painless* OG could be a "*Pain1*" ortholog
277 (supplementary fig. S8). We detect one intracellular and one extracellular site with significant
278 evidence of positive selection in cold-adapted species in the otherwise conserved *TrpA1* OG
279 (ω=0.171; fig. 6). The *painless* OG contains six significant sites but this gene is much more rapidly
280 evolving overall (ω=0.392), which could indicate erroneous clustering of divergent paralogs in the
281 orthology inference. *subdued* encodes an anoctamin ion channel of the Transmembrane protein 16
282 family (Yuan et al. 2022), which conducts chloride in response to heat and is associated with thermal
283 nociception in *Drosophila* (Jang et al. 2015). It is conserved overall (ω=0.12) and we detect three sites
284 with signatures of positive selection (supplementary fig. S9*A*). In addition, the candidate set contains
285 six heat shock protein or chaperone genes, including homologs of *Hsp110*/*Hsc70Cb*, *Hsp67Bc*,
286 *Hsp70Bbb* and three *Chaperonin containing TCP1* chaperonins (*CCT3*, *CCT7* and *CCT8*), that are
287 involved in folding proteins and protecting them from noxious temperatures and other environmental
288 stressors (Mayer and Bukau 2005; Chen et al. 2018).

Fig. 6.



Signatures of positive selection in a *Transient receptor potential cation channel A1* (*TrpA1*) homolog in cold-adapted krill. Top: the intracellular, extracellular and six expected transmembrane domains of the TRPA1 membrane protein encoded by the *E. superba* ortholog (predicted using *TOPCONS v2.0* (Tsirigos et al. 2015)). Gray transmembrane regions=outbound; white=inbound. Ankyrin-repeats were detected using BLAST Conserved Domain Search (Lu et al. 2020). Middle: the trimmed protein multiple-sequence alignment (MSA) of the krill TRPA1 orthogroup. Bottom: magnified view of the MSA highlighting the sites and species. Two stars indicate putatively selected sites (BEB>95%).

In the second contrast, we compared three tropical–subtropical Pacific Ocean species against the rest ((*E. distinguenda*, *E. lamelligera* and *E. mucronata*; fig. 5*A*; Table S4). In this smaller foreground, we detected fewer candidates and less strong GO enrichments (383 genes with LRT<0.05). We found signatures of positive selection in genes involved in axon guidance and neurodevelopment (fig. 5*B*; Tables S5–S6), including *Neurotrophin 1* (*NT1*) and *Semaphorin 1a* (*Sema1a*) (Cafferty et al. 2006; Shen et al. 2017). In addition, we detect selection signals in two heat shock protein-related genes: *Hsc/Hsp70-interacting protein* (*HIP*) and *Heat shock protein 70 cognate 5* (*Hsc70-5*), the latter of which is linked to heat response in crustaceans (Yuan et al. 2017). This contrast uncovers several candidates for photoreception, including *ninaB*, which synthesizes the retinal chromophore pigments of rhodopsin (Voolstra et al. 2010) and *Photoreceptor dehydrogenase* (*Pdh*), which also processes retinal pigment. We identify an opsin (OG98713; supplementary fig. S9*B*), automatically annotated as Rh5. However, queries against KrillDB[2] (Urso et al. 2022) and NCBI

15

308 Genbank suggest it is a peropsin, a light-sensitive yet non-visual receptor with poorly understood
309 function in crustaceans but with expression patterns that oscillate daily in Antarctic krill (Henze and
310 Oakley 2015; Biscontin et al. 2016; Palecanda et al. 2022).

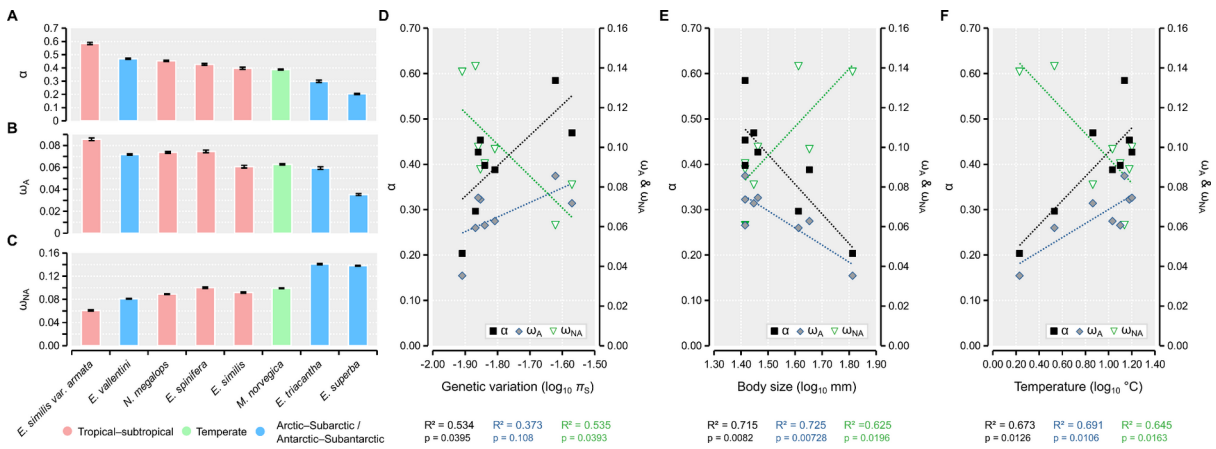
311 In both contrasts, we detect selection signals across multiple genes encoding cuticular proteins
312 (n_{cold}=10; n_{warm}=5) and chitinases (n_{cold}=1; n_{warm}=1). Cuticular protein genes are among the largest gene
313 families in arthropod genomes (Magkrioti et al. 2004), and help build the exoskeleton. In krill,
314 cuticular development and gene expression are closely linked to physiological and seasonal cycles of
315 ecdysis (molting), growth and reproduction (Buchholz and Buchholz 2010; Urso et al. 2022). In krill
316 native to warm waters, we detect potentially selective substitutions in *Ecdysone-induced protein 75B*
317 (*Eip75B*), a gene encoding a nuclear receptor that regulates the timing of molting, reproduction and
318 circadian oscillation in insects and crustaceans (Kumar et al. 2014; Gouveia et al. 2018), that could
319 contribute to adaptation.

320 **Slow rates of adaptive protein evolution in some Southern Ocean species**

321 Theory suggests high efficacy of natural selection and adaptive potential in zooplankton (Peijnenburg
322 and Goetze 2013). This hypothesis predicts that much variation between related zooplankton should
323 have evolved through adaptation. To test this, we estimated how much of divergence in protein-coding
324 genes between species may have been driven by positive selection as opposed to non-adaptive
325 processes (like genetic drift) in krill. We inferred *dN* and *dS* between species (Tables S7–S8) and
326 ancestral nodes and estimated intraspecific polymorphism (*pN/pS*) using unfolded site frequency
327 spectra (supplementary fig. S10). We next estimated the rates of adaptive and non-adaptive amino acid
328 substitutions (ω_A and ω_{NA}), respectively, and the overall proportions of adaptive amino acid
329 substitutions (α) using non-parametric and model-based approaches (Table S7). We used the Gamma-
330 Zero model to compare all species, which fits a gamma distribution for neutral and deleterious
331 mutations to the data. It produced estimates similar to the Displaced-Gamma model, both of which
332 behaved consistently across all datasets, producing estimates correlated with but higher than basic
333 non-parametric methods and were considered accurate in recent tests (Al-Saffar and Hahn 2022). We
334 detected the highest rates of adaptive protein evolution in the subtropical *E. similis* var. *armata* lineage

(α=0.59; ω_A=0.086) (fig. 7*A–C*; Table S8), whereas adaptive evolution appears more limited in the Antarctic krill *E. superba* (α=0.20; ω_A=0.035) and subantarctic *E. triacantha* (α=0.30; ω_A= 0.059). We then analyzed how effective population size (N_e), here approximated by π_S, life history traits and habitat conditions correlate with adaptive rates to identify factors that may influence adaptive potential. Since inference of α depends on π_S, we used estimates of π_S from genes not used to compute α. We find that π_S predicts α (r^2=0.53; p=0.040), but is more strongly associated with ω_{NA} (r^2=0.54; p=0.039) than with ω_A (r^2=0.37; p=0.108) (fig. 7*D*). Our results suggest that genetic variation between krill species may be more strongly driven by increased fixation of non-adaptive variants in small populations than by accelerated positive selection in large populations, a pattern previously observed across animal phyla (Galtier 2016). We find that small body and propagule sizes (i.e. larval sizes) and warmer habitats are strongly associated with high α (fig. 7*E–F*; Table S9), suggesting that life history and habitat characteristics may influence the amount and rates of adaptation in krill.

Fig. 7.



Estimates of adaptive protein evolution in nine krill species (GammaZero model). (*A–C*) The proportions and rates of adaptive and non-adaptive amino acid substitutions (α, ω_A and ω_{NA}), respectively. Whiskers indicate maximum likelihood confidence intervals. (*D–F*) Log_{10} linear regression correlations between genetic variation (π_S), maximum body size and habitat sea-surface temperature and α, ω_A and ω_{NA}, respectively. R^2 Pearson correlation coefficients and significance-values shown below the graph.

Discussion

We here use comparative population genomics to characterize the patterns of genetic variation and

17

evolutionary history of krill on the largest scale to date. Assessments of phylogenetic interrelationships and genetic adaptation among krill species have previously only been based on a few markers. Our analyses of molecular evolution in over 13,000 putative gene orthologs among krill native to cold and warm environments offer new insights into genetic variation, taxonomic diversity and adaptation.

Elusive determinants of genetic variation in krill

Mitochondrial markers have uncovered considerable variability in intraspecific variation among krill species but rarely exceeding 2%/bp (Bucklin et al. 2007; Bortolotto et al. 2011). We here estimate levels of selectively neutral genetic variation across thousands of nuclear genes and find that it varies more than two-fold among nine species (π_S=1.1–2.5%). Large and superabundant species like the Antarctic krill *E. superba* and Northern krill *M. norvegica*, that each have enormous biomass estimated to 100s of Mts (Atkinson et al. 2009; Tarling 2010), only have intermediate levels of variation. The most diverse species in our set is *E. vallentini* (π_S=2.5%), a small (13–28mm) subantarctic, omnivorous and possibly panmictic species with a 2-year lifespan and circumpolar distribution mostly distributed north of the Polar Front (Russell et al. 1969; Ridoux 1988; Mayzaud et al. 2003; Harkins et al. 2013; Cuzin-Roudy et al. 2014). It can be the dominant species in shelf areas, including the Patagonian coast and the Kerguelen archipelago (where we collected it) (Palma and Silva 2004; Koubbi et al. 2011; González et al. 2016). In contrast, we find that *E. triacantha* is the least diverse species (π_S=1.1%). This large (24–41mm) and carnivorous krill has a 3-year lifespan and 2-year generation time and a circumpolar distribution that spans the Polar Front, across which it can dominate the mesopelagic euphausiid biomass (Russell et al. 1969; Siegel 1987; Phleger et al. 2002; Cuzin-Roudy et al. 2014). These two extremes could represent “r-strategist” vs. “K-strategist” species, whose characteristic life history traits (e.g. short vs. long lifespan, small vs. large bodies or propagules) have previously been linked to high vs. low genetic diversity across metazoans (Romiguier et al. 2014; Ellegren and Galtier 2016). However, associated traits like fecundity are not known in either species, and across our dataset we otherwise find no statistically significant correlations between such life history traits and levels of variation. We are therefore unable to identify

18

general determinants of diversity in krill, which may need sampling of more species and ecological information.

Compared to the neutral levels of genetic diversity of other arthropods, we find that krill has intermediary levels of variation (π_S: mean±SD): 1.5±0.5% in krill vs. 1.4±0.9% in 43 species (Leffler et al. 2012) or 1.4%±1.2% in 26 species (Romiguier et al. 2014). Thus, krill are in this sense neither hypo- or hypervariable. Across the tree of life, genetic diversity in natural populations varies much less than census population sizes (N_e), a largely unexplained phenomenon termed "Lewontin's paradox" (Lewontin 1974; Charlesworth and Jensen 2022), and many marine species appear to have much lower N_e than expected from abundance (Hedgecock 1994). We infer long-term N_e for *E. superba* and *M. norvegica* to be only 2 M and 1.8 M, respectively, far below expected N_e of hundreds of trillions of individuals. Low estimates of N_e inferred from limited observable variation could be due to selection for low mutation rates afforded in species with large populations (Sung et al. 2012). For example, mutation accumulation lines in the ubiquitous unicellular phytoplankton *Emiliania huxleyi* recovered a low mutation rate ($\mu=5.6\times10^{-10}$), although it was not considered low enough to explain its modest levels of variation (π_S=0.6%) (Krasovec et al. 2020). Likewise, Shao and co-workers recently estimated low nucleotide diversity in the Antarctic krill ($\pi_{genome}\approx0.25\%$; $\pi_{CDS}\approx0.17\%$) using genome-scale data, and also inferred a low mutation rate ($\mu=6.2\times10^{-10}$) as a means to explain this (Shao et al. 2023). These estimates are extremely low and difficult to reconcile with ours. We measure 5.3× as much variation at synonymous sites than the genome-wide estimate in *E. superba* (1.3% vs. 0.25%), or in coding regions overall (0.88% vs. 0.17%). Insight into genomic variation in euphausiids is still limited, as is baseline data derived from other sources than next-generation sequencing. However, Papot and co-workers studied variation in three PCR-amplified paralogs of the nuclear *hsp70* in *E. superba* and *E. crystallorophias* (Papot et al. 2016), and re-analysis of their *E. superba* data gives π_S ranging between 1.4–5.0% ($\pi_{S(A)}$=1.4%; $\pi_{S(B)}$=5.0%; $\pi_{S(C)}$=1.8%), close to our average estimate across 10,777 genes. Moreover, Shao and co-workers derived low μ assuming a molecular clock based on their estimate of 19.5% divergence from the crab *Eriocheir sinensis*, whereas we estimate distances at synonymous sites (d_s) between *E. superba* and 13 other congeneric *Euphausia* species to be 27±7% (Table S7). This greatly exceeds the distance to the crab, suggesting their estimate may be downward-

19

410 constrained by unusually conserved alignments. While we do not exclude the possibility of low or
411 variable mutation rates among genomic regions and species, we find that comparable statistics for less
412 constrained sites (e.g. π_S and d_s) are not available, and call for additional studies of mutation rate in
413 krill.

414 Other factors may also underlie low N_e / N_c ratios, such as highly skewed reproductive success
415 among individuals, population size changes, linked selection that reduces variation around selected
416 sites or constantly shifting adaptive peaks (Hedgecock and Pudovkin 2011; Charlesworth and Jensen
417 2022; Árnason et al. 2023), all of which suggest that natural populations may be far away from Hardy-
418 Weinberg equilibrium. A recent study in the Atlantic cod found compelling genomic evidence of
419 skewed reproductive success driven by pervasive and recurring selection (Árnason et al. 2023),
420 evidenced for instance by negative Tajima's D (D_T) throughout the genome. We find strongly negative
421 D_T in all krill and that genetic variation is much depleted at UTRs and non-synonymous sites,
422 consistent with pervasive effects of purifying and linked selection across genomes. *Euphausia*
423 *triacantha*, the least diverse species in terms of π_S, has the largest excess of low frequency variants and
424 the lowest D_T (-1.9), indicating recent population expansion in this species. Likewise, genome-wide
425 patterns of variation in *E. superba* suggest it too has expanded from a recent bottleneck (Shao et al.
426 2023). Demographic history and natural selection together shape and limit variation in euphausiids.

427 **Genome scale assays provide new insights into euphausiid biodiversity**

428 Our current understanding of marine zooplankton biodiversity is plagued by widespread cryptic
429 variation, which has called for more systematic use of molecular tools in diversity assessments
430 (Choquet et al. 2018; Bucklin et al. 2021; Bucklin et al. 2021). For instance, genome-wide data
431 recently revealed the existence of three genetically distinct lineages in the pelagic pteropod *Limacina*
432 *bulimoides*, originally considered a single circumglobal species (Choo et al. 2023). Because of their
433 thin shells susceptible to dissolution, pteropods are used as indicators of ocean acidification. Thus, as
434 different species may respond differently to acidification, an accurate definition of species boundaries
435 is critical to assess its impacts.

436 In the krill, there are currently 86 described and recognized species, and no new species has

437 been described since 1987 (Baker et al. 1990). Cryptic variation was reported within *Stylocheiron*
438 *affine*, with a divergence level of 14% (mtCOI) differentiating Red Sea individuals from their Atlantic
439 congeners (Wiebe et al. 2016; Bucklin et al. 2021). This mtCOI barcoding gap was on-par with
440 interspecific levels of divergence in the genus (Bucklin et al. 2007). Here, we detect high
441 mitochondrial divergence between *E. similis* and *E. similis* var. *armata*, that co-segregate with the
442 main diagnostic trait. The *armata* variety was first described by Hansen (1911) as differing from *E.*
443 *similis* by: "[...] *a protruding, acute process on the third abdominal segment*". Later, John (1936)
444 added that "[...] *the process may be variable in size and shape; it varies from being low, rounded and*
445 *inconspicuous to being a large compressed spine pointing backwards over the fourth segment*",
446 described slight differences in the structure of antennules and male copulatory organs, and proposed
447 that *armata* reaches sexual maturity at a smaller length than *E. similis*. Russell et al. (1969)
448 distinguished *armata* as living nearer the surface compared to *E. similis*. Our morphological analysis
449 focused exclusively on the process on the third abdominal segment and we report a distinctive process
450 in *armata* with a variable degree of development, in line with John's findings (1936).
451 We estimate 6.2% of mitochondrial divergence (mtCOI) between *E. similis* and *armata*,
452 above the within-species level of variation reported for the genus (average at 2.5%) (Bucklin et al.
453 2007), but below the between-species level (average 16.5%), making it inconclusive. However, the
454 power of a single mitochondrial gene is limited. Here, we improve on this by also assessing variation
455 across the genome. We found 36,930 fixed SNPs (with a F_{ST}=1) across 4,011 genes. These are
456 enriched for terms relating to roles in oocyte development and chromosome segregation (table S10),
457 which could indicate barriers to gene flow (Hamaguchi and Sakaizumi 1992; Montecinos et al. 2017;
458 Boynton et al. 2018). Our average estimate of genome-wide F_{ST} between *E. similis* and *armata* is
459 0.68, well above the F_{ST} of 0.41 reported between the two most differentiated lineages of the pteropod
460 *Limacina bulimoides*, where Choo et al. (2023) concluded that these are most likely distinct species
461 separated by strong reproductive isolation. Pairwise *dS* between *E. similis* and *E. similis* var. *armata*
462 is 5.24%, slightly higher than between the two recognised species *E. spinifera* and *E. longirostris*
463 (4%) (Table S7) although below the average between other sister species in the genus *Euphausia*
464 (17%), which may suggest recent speciation. Considering all of the above, we propose to give the

465 *armata* variety of *E. similis* the status of distinct species. Failure to recognize *E. similis* and *armata* as
466 distinct evolutionary units or species would undercount krill biodiversity, inflate intraspecific
467 structure and skew our understanding of species biology. In addition, it is essential to accurately
468 describe taxonomy for our objectives of comprehending mechanisms underlying ecological adaptation
469 in krill and assess adaptive potential to climate change.

470 **The genetic basis of ecological adaptation in krill**

471 Polar regions and oceans are heating faster than the global average, impacting habitats and krill stocks
472 in Arctic and Antarctic waters (Flores et al. 2012; Atkinson et al. 2019; Clem et al. 2020; Edwards et
473 al. 2021; Rantanen et al. 2022). These rapid changes necessitate better insight into how thermal
474 tolerance varies among species and what genetic mechanisms may contribute to adaptation to a
475 warmer world. By analyzing CT50 data in six species, including new observations from subantarctic
476 and temperate species, we find that CT50 thresholds appear to be correlated with the ambient
477 temperatures of both sampling sites and habitats of species, likely reflecting adaptation to different
478 thermal environments. However, when accounting for the different ambient starting temperatures, *E.*
479 *crystallorophias*, *E. superba*, *E. vallentini*, *E. triacantha* and *T. inermis* show similar absolute CT50s,
480 indicating comparable thermal tolerances beyond the temperatures experienced in their natural
481 habitats. Only *M. norvegica* stands out with a much higher absolute CT50, which can be expected
482 from its ubiquitous distribution. All the experiments carried out in this work were performed on
483 animals collected at the colder end of their natural range, which may influence these patterns. By
484 sampling each species also from warmer habitats, it would be possible to gain more comprehensive
485 insight into their respective plasticities in thermal tolerance. Our observations suggest that these krill
486 species may not be, at least at the adult stage, as thermally sensitive and stenothermic as other polar
487 invertebrates (Peck et al. 2014). This is consistent with the findings of experiments conducted on *E.*
488 *superba* (Toullec et al., 2020), but it may not apply during other stages of their life cycle, such as
489 larval development, where growth rates and survival were shown to start deteriorating already around
490 +3.0°C (Atkinson et al. 2006; Perry et al. 2020).

491 Daily and seasonal variation in ecological conditions such as light and temperature play

492 critical roles in shaping the physiology and timing of behavior, growth, development and reproduction
493 of zooplankton (Richardson 2008; Häfker et al. 2023), which in turn affect the function of marine
494 ecosystems. The molecular mechanisms that govern important ecophysiological and phenological
495 traits have only been intensely studied in some krill, focusing on species adapted to extreme
496 photoperiodic seasonality at high latitudes (e.g. *E. superba*, *Thysanoessa* spp.) or that occur across
497 wide geographical ranges (e.g. *M. norvegica*). This includes efforts to uncover and functionally
498 characterize the photoreceptor and circadian clock gene repertoires (Biscontin et al. 2016; Christie et
499 al. 2017; Palecanda et al. 2022; Urso et al. 2022) or genes involved in regulating the molting cycle
500 (Seear et al. 2010), seasonal dynamics (Seear et al. 2012; Höring et al. 2021) and heat shock response
501 (Huenerlage et al. 2016; Papot et al. 2016). However, analyses of how specific genes may contribute
502 to environmental adaptation through natural selection have only been performed in heat shock proteins
503 *hsp70* genes of Antarctic species before. Both *E. superba* and *E. crystallorophias* appear to have
504 experienced diversifying selection between *hsp70* paralogs, but *E. superba hsps* are characterized by
505 greater functional diversity and balancing selection associated with a more heterogeneous
506 environment, while *E. crystallorophias hsps* may have evolved under stronger directional selection in
507 a more constantly cold environment (Cascella et al. 2015; Papot et al. 2016). In the Antarctic krill,
508 genes involved in controlling seasonal metabolic, molting, growth and reproductive cycles have been
509 shown to be differentially expressed between summer and winter seasons (Höring et al. 2021; Urso et
510 al. 2022). These cycles are thought to be driven by photoperiodic environmental cues such as day
511 light, temperature and food availability and involve genetic pathways for photoreception and circadian
512 clocks (Teschke et al. 2008; Biscontin et al. 2017; Piccolin et al. 2018; Höring et al. 2021).
513 Environmental conditions vary widely among the habitats occupied by the species in our study and we
514 find signatures of positive selection in hundreds of genes associated with multiple ecophysiological
515 functions. We find selection signals across genes involved in photoreception (e.g. *ninaB*, *pdh* and
516 *peropsin*), phenological physiology (e.g. *Eip75B*) and development (multiple cuticular proteins and
517 chitinases), also in species evolving in warm environments.
518 While our candidate sets relate to diverse physiological functions, evolutionary change in
519 thermosensory genes in particular could be key for adaptation to extreme or changing temperatures. In

our analyses of Southern Ocean krill, we identify multiple candidate genes involved in detecting and responding to temperature, including a *TrpA1* homolog. TRP genes encode Transient receptor potential (TRP) Ca^{2+} ion channels, which are the main thermosensors in sensory neurons of animals (Dhaka et al. 2006; Bandell et al. 2007) and considered a “molecular toolkit for thermosensory adaptations” in animals (Hoffstaetter et al. 2018). TrpA1 is associated with harmful heat signaling and heat avoidance behaviors in both invertebrate and vertebrate ectotherms (Akashi 2021; Xiao and Xu 2021), as well as cold sensation in a wide range of animals (Akashi 2021; Zhang et al. 2022). In Southern Ocean notothenioid fish, *TrpA1* has been hypothesized to evolve through duplication and positive selection and be one of the main thermosensors underlying adaptation to cold Antarctic waters (York and Zakon 2022). In addition, we detect potentially adaptive substitutions across several genes encoding heat shock proteins or chaperonin-containing TCP1 (CCTs), which either assist with general or cytoskeletal protein folding. CCTs are eukaryotic "cold-shock" proteins and overexpressed during cold stress (Somer et al. 2002). In notothenioid fish, they appear to have undergone adaptive evolution to accomplish protein folding in cold and energy-depleted environments but are also upregulated from heat stress (Pucciarelli et al. 2006; Cuellar et al. 2014).

Our results suggest that homologous thermosensory receptors and protective proteins have been targeted by natural selection in both krill and fish, and that this repertoire of genes could underlie thermal adaptation also in other Antarctic or Arctic organisms. However, the specific functional properties of these candidate genes, and impact of their substitutions, are currently unknown in krill. For example, we detect signals of selection in a homolog of *subdued*. In *Drosophila*, the thermosensory Subdued ion channel collaborates with TRPs in heat nociception but activates at high temperatures that are unlikely to be experienced by any krill (+40°C) (Jang et al. 2015). On the other hand, *Drosophila* and krill Subdued homologs are only 43–45% identical and TMEM16 proteins overall are functionally diverse (Picollo et al. 2015). It is possible that krill Subdued requires considerably less activation energy, as seen in other cold-adapted proteins (D’Amico et al. 2002). The genes detected here are promising candidates but further functional validation is necessary to understand their role in adaptation.

24

547 **On adaptive potential in Antarctic krill**

548 The potential for genetic adaptation ultimately depends on access to genetic variation, favorable
549 demographics and life history traits. Therefore, the extent to which selection has shaped genetic
550 variation among species in the past, and the potential they have to adapt to ongoing environmental
551 change, may vary significantly among taxa. Empirical studies of protein evolution have indicated that
552 few substitutions among apes have been fixed through positive selection (α=0–0.3) (Hvilsom et al.
553 2012; Galtier 2016), while high α in organisms like *Drosophila* (α≈0.5) and sea squirts (α≈0.8)
554 indicate strong influence of positive selection (Smith and Eyre-Walker 2002; Tsagkogeorga et al.
555 2012). These analyses typically uncover higher α in invertebrates compared to vertebrates (Galtier
556 2016), find that adaptive evolution is limited by the supply of variation in low-N_e species and
557 negatively correlated with "K-selected" life history strategies such as long lifespans or large offspring
558 (i.e. the propagule size) (Galtier 2016; Rousselle et al. 2020), although how these potential
559 determinants influence adaptation is still not well understood.

560 The hypothesis of high evolutionary potential in zooplankton (Peijnenburg and Goetze 2013)
561 predicts that most variation between species should be shaped by adaptive evolution. This may not
562 hold for krill, at least not judging by how natural selection has driven divergence in protein-coding
563 genes in the past. Among the eight species of krill examined here, we find that the proportion of
564 adaptive protein evolution (α) is only 0.40±0.11 (mean±SD), suggesting the majority of amino acid
565 substitutions between species may have been fixed through processes other than adaptation. Compared
566 to other arthropods, α seems low. For example, Galtier previously estimated α to 0.62±0.13 among 11
567 species (Galtier 2016). Our results indicate that the overall α in krill scales with N_e, but it is more
568 strongly influenced by non-adaptive change than by adaptive change. As seen in other species, the rate
569 of fixation of slightly deleterious variants (for example due to drift in small populations) scales more
570 clearly with N_e than the fixation rate of adaptive variants (Galtier 2016; Moutinho et al. 2020). Non-
571 adaptive processes appear to contribute more to protein evolution in the Southern Ocean species *E.*
572 *triacantha* (α=0.30) and *E. superba* (α=0.20). In the other krill, the ratios between *dN/dS* (ω) and
573 *pN/pS* range from 1.29× to 1.83×, but are in these two species only 1.10× and 1.01×, respectively. For

comparison, across the subset of 566 putatively cold-selected genes that we identify with branch-site tests and where these species have substitutions at selective sites, we find that α is 1.44× higher (0.43) in *E. triacantha* and 1.40× higher in *E. superba* (0.29) than the transcriptome-wide estimates, suggesting these genes have undergone more adaptive evolution overall than the genomic background. While the two species have undoubtedly adapted genetically to their respective environments, the question remains why so many substitutions throughout most genes have not evolved through adaptation. Our results suggest that K-strategists may be associated with low rates of adaptation in krill, as the two species have large bodies and larvae. However, demography and environment may also contribute. *E. triacantha* has the biggest excess of low frequency variants and lowest D_T (-1.9) in our analyses, consistent with recent expansion from a bottleneck of small N_e causing fixation of non-adaptive variants through genetic drift. Genome-wide analyses of *E. superba* suggest it too has recently undergone major population expansion (Shao et al. 2023), although our D_T is less extreme in this species (-1.4). An alternative explanation could be that slow rates of adaptation in the Antarctic krill may follow from comparably flat environmental gradients or slow rates of environmental change in Antarctic habitats. Since its formation, the Antarctic Circumpolar Current has acted as a barrier to warmer waters and helped maintain cold and stable Antarctic conditions (Clarke et al. 1992). Simulations indicate that the rate of recurring environmental change affects adaptive rates more strongly than N_e (Lourenço et al. 2013). Our sample is limited and only by expanding these analyses to small Antarctic species such as *E. crystallorophias* and *E. frigida* would it be possible to disentangle life history from habitat conditions as the main drivers of adaptive rate in Southern Ocean krill.

Krill are genetically adapted to their respective environments, but most divergence among species appears to have evolved through non-adaptive processes, suggesting adaptive traits have emerged over long timescales. What does this mean for their adaptive potential under rapid climate change? The answer lies in species-specific access to functional variation and efficacies of natural selection. We compared non-synonymous variation (π_N) among our candidate genes for warm-adaptation against background π_N to gauge how much standing and assumably functional variation is available in populations to select from. Interestingly, we observe more standing variation for warm-adaptive genes among most species native to warm environments than to cold environments

26

602 (supplementary fig. S12; Table S11). This could mean that more adaptive alleles are present in
603 populations normally challenged by high or variable temperatures, possibly maintained by balancing
604 selection or incomplete selective sweeps over environmental gradients or time. The Antarctic krill has
605 no excess of functional (π_N) variation among genes implied in warm-adaptation compared to
606 background genes, and warm-adaptive genes have been evolving under even weaker influence of
607 positive selection than the background (α=0.17 vs. α=0.20). A combination of long generation time
608 (2–3 years) (Siegel 2000), extensive panmixia impairing local adaptation (Lenormand 2002; Shao et
609 al. 2023) and possibly limited functional variation suggest the Antarctic krill has comparably low
610 adaptive potential to rapidly changing environments.

611 Conclusions

612 This study is the first to use comparative genomics at the genome scale to investigate evolutionary
613 history and adaptation in marine zooplankton. We demonstrate the usefulness of transcriptome data as
614 a means to identify patterns of genetic variation and adaptation among non-model species from
615 different environments. Thorough understanding of taxonomy is essential to accurately interpret
616 molecular signatures of adaptation in species and evaluate their adaptive potential. We identify
617 significant genome-wide divergence between the two variants of *Euphausia similis* and *Euphausia*
618 *similis* var. *armata* from the Indian Ocean. We show that the two forms are likely genetically isolated
619 and evolving independently, with different levels of genetic diversity and rates of adaptive evolution.
620 We argue that *Euphausia similis* var. *armata* should be given species status with the name "*Euphausia*
621 *armata*", resolving a 100-year old issue in krill systematics.

622 Conservation genetics has traditionally used neutral variation to survey population status and
623 risk, but the link between neutral diversity and population viability is disputed and it can be argued
624 that conservation should consider variation that directly promotes viability (Teixeira and Huber 2021).
625 Comparison of new gene candidates for thermal adaptation, like *TrpA1*, *subdued*, *Hsc/Hsp70-*
626 *interacting protein* (*HIP*), between tropical and Antarctic krill species may provide valuable
627 information on functional variation that promotes fitness under warmer conditions. Surveys of genetic
628 variation for these candidates can help uncover to what extent adaptive alleles occur in natural

populations and provide insight into how much genetic change is required to adapt to future climates, i.e. the genetic offset (Fitzpatrick and Keller 2015).

We find that rates of adaptation are not uniformly high in krill, but vary widely among species. In particular, some Southern Ocean species like the Antarctic krill may have low adaptive potential. Just as other research indicates that there may be more structure and reproductive isolation among zooplankton populations than traditionally believed, our results suggest they do not behave like infinitely large effective populations unaffected by genetic drift or with unlimited adaptive potential. Instead, it seems to be determined by interactions between demography, life histories and environmental conditions. Expanding these analyses to more taxa could uncover common indicators of high or low adaptive potential in zooplankton, and allow more accurate forecasts of species distributions under continued change (Razgour et al. 2019; Capblancq et al. 2020).

Materials and Methods

Species identification and life histories

We used dissection microscopes and keys for morphological species identification, including Baker *et al.* (Baker et al. 1990), and *Cytochrome c oxidase I* (*COI*) barcodes for molecular species validation. Sample information is provided in Table S1. We georeferenced species ranges from (Brinton et al. 2000) on a OISST v2 Sea Surface Temperature (SST) map with *QGIS* http://www.qgis.org and computed the average SST for each range from color histograms in *GIMP* https://www.gimp.org/. To associate adaptive rates with life history traits, we compiled information from (John 1936; Russell et al. 1969; Baker et al. 1990; Brinton et al. 2000).

Atlantic Ocean samples

We collected *M. norvegica* (Northern krill) specimens (n=7) from the Swedish Gullmarsfjord in 2018 using a small Isaacs-Kidd midwater trawl (IKMT) and preserved them in RNA*later*™ (Invitrogen™). We used material from three North Atlantic–Arctic *Thysanoessa* species: *T. longicaudata* and *T. raschii* specimens were collected in August 2012 in the Kongsfjorden (Spitzberg) and frozen in liquid nitrogen. We re-used a published RNA-seq library (ENA: SRR2174568) for *T. inermis* from

28

(Huenerlage et al. 2016).

Indian Ocean samples

Specimens of *Nematoscelis megalops* and six *Euphausia* species (*E. recurva*, *E. spinifera*, *E. similis*, *E. longirostris*, *E. triacantha*, *E. vallentini*) were collected in the Indian Ocean in 2019. We used a large IKMT aboard the R/V Marion Dufresne and kept specimens alive for 1–48 hours before preserving them in RNA*later*™ or performing experiments.

Pacific Ocean samples

Four Pacific species of *Euphausia* were provided by Dr N. Tremblay: *E. distinguenda* and *E. lamelligera* (West coast of Mexico, February 2012), *E. pacifica* (West coast of Oregon, September 2011), *E. mucronata* (West coast of Chile, August 2011).

Southern Ocean samples

During a 2017 REPCCOAI cruise, we collected a specimen of *T. macrura* North-East of Kerguelen Archipelago and a *E. frigida* specimen near the Indian Ocean Polar Front and froze them in liquid nitrogen. We downloaded a RNA-seq library for the Ice krill *E. crystallorophias* (ENA: ERR264582) from (Toullec et al. 2013) and 20 libraries for the Antarctic krill *E. superba* from (Höring et al. 2021) (NCBI: PRJEB30084).

CT50 experiments

Experiments were performed using live *E. triacantha* and *E. vallentini* samples from the Indian Ocean and *M. norvegica* from the Arctic Ocean. The experimental protocol was identical to that used for *Euphausia crystallorophias* and *Euphausia superba* (Cascella et al. 2015) or the Arcto-boreal *Thysanoessa inermis* (Huenerlage et al. 2016). After acclimation of 24 h at ambient sea water temperatures, actively swimming animals were transferred to an experimental tank. Temperature was increased by 1°C per 10 min. The animals were maintained in the tank until they no longer responded to tactile stimuli of a probing rod. At this point, it was considered that the critical temperature had been reached. The CT50 was determined through the non-linear curve fitting option in Prism9 (GraphPad Software, LLC). The mobility/survival curve used was: Survival = c/(1 + (T/CT50), where c is the plateau value before the sharp decrease and CT50 is the threshold temperature at which only

29

682 50% of animals are mobile. The program explores the different parameter values and calculates a 95%
683 confidence interval (standard error provided).

684 **Extraction and sequencing of RNA**

685 For *M. norvegica* and samples collected from the Indian Ocean in 2019, we dissected and used
686 abdominal muscle to extract RNA, using the Qiagen RNeasy® Plant Mini Kit and following the
687 manufacturer's protocol. Care was taken to remove the intestine to avoid bacterial or food
688 contaminants. Thermo Scientific NanoDrop™ and Agilent Technologies 2200 TapeStation
689 instruments were used to measure yield, purity and RIN-values. Samples with RIN>8 were provided
690 to Science for Life Laboratory, Sweden, without DNase treatment (the kit protocol proved highly
691 efficient in removing gDNA). Staff prepared Illumina RNA-seq 94 libraries using Illumina TruSeq
692 Stranded mRNA Library Prep kit with polyA selection. Paired-end libraries (2×150 bp) were
693 sequenced on a Illumina NovaSeq 6000 S4 lane and demultiplexed by the platform. Other samples
694 were prepared as in (Huenerlage et al. 2016).

695 **Molecular species validation**

696 For each library, we queried 1 million forward-oriented reads against all krill *COI* reference barcodes
697 in the MetaZooGene database (MZGDB v3; n=3,003 sequences from 64 fully identified species)
698 (Bucklin et al. 2021) using *BLASTN* (Tan et al. 2006; Camacho et al. 2009) and recorded best hits. For
699 barcodes with 20+ hits, we computed the average identity scores. We took the barcode with the
700 highest average score as an indication of species.

701 **RNA trimming, assembly and annotation**

702 We used *Trim Galore!* v0.6.1 https://github.com/FelixKrueger/TrimGalore/ and *Cutadapt* v2.3
703 (Martin 2011) to trim low quality positions from reads (phred<20), only keeping trimmed reads of at
704 least 50 bp. *Trinity* v2.11.0 (Grabherr et al. 2011) was then used with default settings to assemble 19
705 transcriptomes, making one "reference" transcriptome per species. When possible, we pooled reads
706 from up to five specimens to maximize gene completeness. Preliminary tests indicated that pooling
707 data produced on average 1.1–1.3× as many complete BUSCO genes compared to a large (44–133 M
708 read pairs) or single small library (~22 M read pairs). Due to unexpected population structure between

709 *E. similis* and *E. similis* var. *armata*, a separate reference transcriptome was assembled for *armata*.
710 The *T. raschii* transcriptome was assembled similarly to *T. inermis* in (Huenerlage et al. 2016). We
711 used the Trinity script "*get_longest_isoform_seq_per_trinity_gene.pl*" to reduce redundancy by
712 keeping only the longest splice isoform per gene.

713 We used *TransDecoder* v5.5.0 https://github.com/TransDecoder/TransDecoder to identify
714 protein-coding transcripts and *TransDecoder.LongOrfs* to detect open reading frames (ORFs) >300
715 bp. ORFs were queried for domain homology against Swissprot with BLASTP v2.9.0+ (Camacho et
716 al. 2009) (e-value cut-off $1e^{-5}$) and Pfam (release 34.0) with HMMER3 *hmmscan* v3.3
717 http://hmmer.org/. We used *TransDecoder.Predict* to identify and carry forward only protein-coding
718 transcripts. Coordinates for functional regions (UTRs or coding sequence) were kept in GFF files. We
719 used *KaKs_Calculator* (Zhang et al. 2006) to enumerate synonymous and non-synonymous sites,
720 while transcriptome completeness was assessed using *BUSCO* v3.0.2b (Simão et al. 2015) and the
721 arthropod odb9 lineage set from OrthoDB. We annotated the transcripts using queries against the
722 *Drosophila* FlyBase database (dmel_r6.38) (Larkin et al. 2021) with *DIAMOND* v9.9.0 (Buchfink et
723 al. 2015).

724 **SNP calling**

725 We measured genetic variation within nine species/lineages: *M. norvegica*, *E. longirostris*, *E. similis*
726 and its variant form *armata*, *E. spinifera*, *E. triacantha*, *E. vallentini, N. megalops* and *E. superba*
727 using Single Nucleotide Polymorphisms (SNPs) detected in the RNA-seq data.

728 We mapped the trimmed RNA libraries from 113 specimens (3–20 specimens per species) to
729 their respective reference transcriptomes using BWA-MEM v0.7.17-r1188 (Li 2013). Alignments
730 were cleaned with *samtools* v1.14 (Danecek et al. 2021) to only retain reads mapping uniquely,
731 concordantly and in pairs. Duplicates were removed using *Picard* v2.23.4
732 (http://broadinstitute.github.io/picard/). For each species, we called variants with *GATK*
733 HaplotypeCaller v4.2.0.0. Variants were called per-individual and combined into a multi-sample
734 gVCF with CombineGVCFs. Individuals were then jointly genotyped with GenotypeGVCFs. We used
735 *vcftools* v0.1.16 (Danecek et al. 2011) to filter variants, retaining only bi-allelic SNPs genotyped by ≥5

736 reads at sites with a mean read depth of ≥5× across all individuals.

737 **Population structure**

738 We here produced high-quality SNP subsets with vcftools, keeping SNPs with a minimum Phred
739 quality score of 30, a minor allele count of three and less than 20% genotype missingness. We pruned
740 SNPs based on linkage-disequilibrium using *Plink* v.1.90 www.cog-genomics.org/plink/1.9/ (Chang et
741 al. 2015), proceeding in windows of 50 SNPs, sliding by 10 markers at a time and with a r^2 threshold
742 set at 0.8. We used the unlinked SNPs to investigate population structure in the Indian Ocean species
743 by performing Principal Component Analyses (PCA) with Plink and ancestry analyses with
744 *ADMIXTURE* v.1.3.0 (Alexander et al. 2009), using 2–N ancestral clusters (*K*) (N=number of
745 geographic sampling sites).

746 **Divergence between *E. similis* ("*similis*") and *E. similis* var. *armata* ("*armata*")**

747 Material from *similis* and *armata* was initially mapped to the same *E. similis* reference. In a joint SNP
748 call, we uncovered extreme structure and used *vcftools* v0.1.16 to estimate the F_{ST} fixation index
749 between them. We also observed that *similis* and *armata* samples matched different MetaZooGene
750 *COI* barcodes. To further characterize this discrepancy, we assembled the *COI* mitochondrial gene
751 (mtCOI) from the RNA-seq data. Reads with best hits against *E. similis* accessions (AF177186,
752 MW210878 or MW210879) were extracted and assembled *de novo* into *COI* fragments using *SPAdes*
753 v3.15.5 (Prjibelski et al. 2020). The fragments and the barcodes were aligned using *MAFFT* v7.453
754 (Katoh and Standley 2013). A pairwise distance matrix and neighbor-joining tree was inferred using
755 *SplitsTree* v4.18.2 (Huson and Bryant 2006).

756 **Levels of intraspecific genetic variation**

757 The full biallelic SNP datasets (above) were annotated with *SNPeff* v4.3T (Cingolani et al. 2012), i.e.
758 to be synonymous, non-synonymous or UTR variants. We estimated variation using the population
759 mutation rate Watterson's theta (θ_w) (Watterson 1975) separately at 5'-UTR, 3'-UTR, synonymous
760 and non-synonymous sites, or jointly across full genes while accounting for accessible sites. We then
761 used SplitsTree to calculate transcriptome–wide pairwise genetic distance matrices among individuals
762 and deduce π (nucleotide diversity) and BioPerl to calculate Tajima's D (D_T) as an indicator of

763 demographic history (Stajich et al. 2002). The effective population size (N_e) was inferred from N_e =
764 $\theta_w/4\mu$, where μ is the mutation rate per base and generation. Because mutation rates are unknown in
765 krill, we used μ from snapping shrimp (Silliman et al. 2021), the closest species with a known rate.

766 **Orthology and phylogenetic analyses**

767 We analyzed the 20 krill transcriptomes and seven outgroup datasets for gene orthology with
768 *ProteinOrtho* v6.0.14 (Lechner et al. 2011), using *DIAMOND* to detect similarities. For each set of
769 orthologous genes (an "orthogroup" or "OG") with 10 or more krill species, we produced protein-level
770 alignments. MAFFT was run using the G-INSI-I algorithm, using the "--allowshift --unalignlevel 0.8
771 --leavegappyregion" variable scoring matrix settings to reduce the risk of over-aligning non-
772 homologous regions (Katoh and Standley 2016).

773 Single-copy OGs with at least 18 krill species and two outgroups were used to make a species
774 tree. First, we trimmed unreliably aligned positions using *Gblocks* (Castresana 2000) (settings: "-t=p -
775 b1=N -b2=N -b4=5 -b5=h -b6=y, where N was 50% + 1 of the number of sequences). We then
776 performed phylogenetic inference of concatenated OGs using the maximum likelihood (ML) method
777 in *IQ-TREE* v2.1.0 (Minh et al. 2020) and the JTTDCMut+F+I+G4 model (chosen using BIC)
778 (Kalyaanamoorthy et al. 2017). We additionally produced individual gene trees using IQ-TREE
779 (JTT+I+G4 model).

780 **Detecting positive selection and candidate genes for adaptation**

781 We prepared OGs with 10+ krill species for analysis of adaptive evolution. The OGs were aligned
782 without outgroups. For OGs with duplicate sequences (e.g. species-specific paralogs), we first
783 produced gene trees with *Fasttree* (Price et al. 2010) and then used *OrthoSNAP* (Steenwyk et al. 2022)
784 to split alignments into single-copy ortholog subsets, only keeping subsets with 10+ species.
785 Nucleotide sequences were fitted to the protein alignments using *PAL2NAL* (Suyama et al. 2006).
786 Spuriously aligned sequences may produce false-positive selection signals. We used *Gblocks* to trim
787 unreliably aligned codons (settings: "-t=c -b5=h -b6=y"). In addition, we masked sequences around
788 internal indels, replacing 4 codons with gaps around each indel and removing positions where >2

789 species had missing data. We also masked alignment fragments <15bp. For each OG, we pruned the
790 species tree using *Phyutility* (Smith and Dunn 2008) to make unrooted species-matched subtrees.
791 We sought to uncover genes evolving under positive selection in krill adapted to either cold or
792 warm climates and designed two tests. In the "cold" contrast, we took the terminal branches leading to
793 five Antarctic–subantarctic *Euphausia* species as foreground branches, while assigning all other
794 branches to the background. In the "warm" contrast, we took three Pacific tropical–subtropical
795 equatorial/near-equatorial species as foreground branches. The ratio between non-synonymous
796 substitutions per non-synonymous site (*dN*—a proxy for selection) over synonymous substitutions per
797 synonymous site (*dS*—a proxy for unconstrained or neutral evolution), termed *dN/dS* or ω, compares
798 the rate of amino acid substitutions between species against expectations under neutrality. Elevated ω
799 (ω>1) can be used as an indication of positive selection, while ω=1 and ω<1 indicate absence of
800 selection or purifying selection (Yang 2006), respectively. However, pairwise estimates for whole
801 genes are only summary statistics that represent the long-term average mode of selection but do not
802 localize selection to specific codons or lineages. Branch-site models leverage multiple sequence
803 alignments and phylogenetic species trees and allow variable *dN/dS* ratios that can detect episodic
804 positive selection involving only some branches (or species) and nucleotide sites (Yang and Nielsen
805 2002). We therefore performed branch-site tests with *codeml* in *PAML* v4.9j (Yang 2007) via the *ETE*
806 *Toolkit* v3 (Huerta-Cepas et al. 2016), providing the codon-level alignment and subtree. We scanned
807 for episodic positive selection by testing whether models allowing positive selection at one or more
808 sites and branches fit the observed data and tree significantly better than null models that only allow
809 neutral or purifying evolution (Yang and Nielsen 2002). For each OG, we ran *codeml/ETE* three times:
810 i) using the bsA1 null model; ii) using the alternative bsA model that incorporates positive selection
811 (Zhang et al. 2005); iii) performing a likelihood ratio test (LRT) to test if the null model could be
812 rejected, taking the log-likelihood difference between models ($-2\Delta\ln L$) against a χ^2 distribution (one
813 df). For genes to be considered candidates, we required both a significant LRT in favor of the selective
814 model and the detection of at least one positively selected site among foreground species, using Bayes
815 Empirical Bayes posterior probability >95% (Yang et al. 2005).
816 To test for shared biological properties among the candidate genes, we performed Gene

817 Ontology (GO) enrichment tests of *Drosophila* gene homologs using *ShinyGO* (Ge et al. 2020). We
818 took the candidate genes as the target set and all analyzed genes as the background set.

819 **Estimating rates of adaptive evolution**

820 McDonald & Kreitman introduced the idea to compare *dN/dS* (ω) between species against non-
821 synonymous and synonymous polymorphisms segregating within species (*pN/pS*) to study selection at
822 individual genes (McDonald and Kreitman 1991). Under the assumption that non-synonymous
823 variants are either neutral, or deleterious and removed from populations via purifying selection, *pN/pS*
824 should equal ω for genes evolving neutrally. Adaptive non-synonymous variants would be expected to
825 be quickly fixed by positive selection and contribute to divergence (*dN*) but not polymorphism (*pN*),
826 such that $\omega \gg pN/pS$ can be taken as evidence for positive selection. By comparing ω to *pN/pS* across
827 multiple genes, it is possible to estimate the overall proportion of adaptive amino acid substitutions
828 (α), in which $\alpha = 1 - (pN/pS / \omega)$, and infer how much of the observed variation between species may
829 have been shaped by positive selection as opposed to neutral evolution (Smith and Eyre-Walker 2002).
830 However, estimates of adaptive rates can be distorted by slightly deleterious non-synonymous
831 variation that are not efficiently removed by purifying selection but segregate at low frequencies due
832 to genetic drift, leading to overestimation of *pN/pS* and underestimation of *α* (Charlesworth and Eyre-
833 Walker 2008; Moutinho et al. 2020), particularly when populations undergo bottlenecks. Current
834 methods therefore interrogate whole site frequency spectra (SFS) of neutral and selective variants
835 while aiming to control for confounding processes such as population structure, bottlenecks and
836 interference between linked sites (Eyre-Walker and Keightley 2009; Messer and Petrov 2013; Galtier
837 2016; Al-Saffar and Hahn 2022). Therefore the site frequency spectrum (SFS) of synonymous variants
838 is used to model demographic history, in order to more accurately model the distribution of fitness
839 effects (DFE) from the non-synonymous SFS and derive more realistic estimates of *α* (reviewed in
840 Galtier 2016; Moutinho et al. 2020; Al-Saffar and Hahn 2022), as well as the respective rates of
841 adaptive or non-adaptive protein evolution (ω_A or ω_{NA}).

842 For eight species with SNP data, we aimed to estimate the overall α, ω_A or ω_{NA} across
843 thousands of genes. To estimate ω, we made 4-way datasets and pruned species trees that included the

844 focal species and three additional species at different distances. We re-estimated branch lengths using
845 the YN98+F1X4 model in *Bio++ bppML* v2.3.1 (Nielsen and Yang 1998; Guéguen et al. 2013),
846 selected an ancestral node at appropriate distance and used *Bio++ bppancestor* v2.3.1 to infer the gene
847 sequence at the ancestral node. The node was selected from overall pairwise genetic distances between
848 species, avoiding very short distances as shared ancestral polymorphism may distort substitution-
849 estimates (Mugal et al. 2020).

850 We then estimated synonymous/non-synonymous substitutions and sites between the focal
851 and ancestral sequences using KaKs_Calculator and the "YN" method to account for unequal base
852 frequencies, transition/transversion rate biases and multiple substitutions (Yang and Nielsen 2000).
853 We compiled transcriptome-wide unfolded non-synonymous and synonymous site frequency spectra
854 (SFS) from SNPs. We here used the ancestral sequence to polarize the SNP alleles: the allele shared
855 with this sequence was considered ancestral, while the other was considered derived. This was
856 performed with *basefinder* (a novel tool implemented here; commit adb820c) that cross-references
857 SNP and alignment positions on either plus or minus strands. SFSs were produced from derived allele
858 frequencies using genes with at least one polarized SNP. To have similar resolution of the SFSs
859 between species and facilitate unbiased comparison, we downsampled the original population datasets
860 to six or eight individuals. For each gene, individuals were selected to maximize the overall coverage
861 and genotyped SNPs.

862 We used the model-based program *grapes* (Galtier 2016) to estimate α, ω_A and ω_{NA} in each
863 species. The program infers fitness-effects and fits either non-parametric methods or models from
864 *DoFE* (Eyre-Walker et al. 2006; Eyre-Walker and Keightley 2009) to the SFS + divergence data under
865 maximum likelihood criteria. The "Basic" method derives the statistics by comparing *dN/dS* to *pN/pS*
866 without correction, while "FWW" removes non-synonymous polymorphisms segregating at low
867 frequencies (<15%) (Fay et al. 2001). In addition, *grapes* implements six model-based methods that
868 estimate population demographics and fitness effects to correct for selectively neutral or fitness-
869 altering non-synonymous polymorphisms (see Galtier 2016 and Al-Saffar and Hahn 2022 for details).
870 For each species and method, we ran *grapes* 10 times with random initial values ("-nb_rand_start
871 10").

872 **TABLE 1**

873 Synonymous variation and demographics in nine krill species.

Ocean	Species	N	Genes[1]	SNPs	Total length[2]	Tajima's D (D_T)	N_e[3]
Atlantic	*M. norvegica*	7	15,445	236,943	3,883,522	-1,153	1,816,807
Indian	*E. longirostris*	3	9,535	74,381	2,278,357	-1,402	1,353,964
	E. similis	10	12,731	216,303	3,072,702	-1,289	1,878,999
	E. similis var. *armata*	9	16,165	442,254	3,849,450	-1,335	3,163,058
	E. spinifera	17	11,804	254,630	2,678,672	-1,741	2,201,560
	E. triacantha	20	13,707	301,721	3,294,790	-1,911	2,038,746
	E. vallentini	18	19,509	709,573	4,470,717	-1,362	3,624,473
	N. megalops	9	12,156	170,844	2,697,988	-1,367	1,743,387
Southern	*E. superba*	20	10,777	205,319	2,203,203	-1,488	2,074,722

874 [1]Expressed genes with at least one synonymous SNP.
875 [2]Counting only accessible sites covered by at least 5× depth of coverage.
876 [3]N_e= effective population size.

877 **TABLE 2**

878 Signatures of positive selection on protein-coding genes.

Nr	Contrast	Analyzed OGs (n)[1]	LRT level	LRT significant OGs (n)	LRT significant + BEB95% OGs (n)	Significant OGs with *Drosophila* orthologs
1	Five cold-adapted species vs. rest	13,198	$p<0.05$[2]	1,794	1,349	1,020 of 8,570 (12%)
			$p<0.01$[3]	997	832	629 of 8,570 (7%)
2	Three warm-adapted species vs. rest	13,004	$p<0.05$[2]	1,039	630	383 of 8,455 (5%)
			$p<0.01$[3]	482	338	191 of 8,455 (2%)

879 [1]OGs that include at least one focal species
880 [2]$\chi 2$ critical value=3.841
881 [3]$\chi 2$ critical value=6.635

37

References

Akashi H. 2021. Thermal Sensitivity of Heat Sensor TRPA1 Correlates With Temperatures Inducing Heat Avoidance Behavior in Terrestrial Ectotherms. *Front. Ecol. Evol.* [Internet] 9. Available from: https://www.frontiersin.org/articles/10.3389/fevo.2021.583837

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1655–1664.

Al-Saffar SI, Hahn MW. 2022. Evaluating methods for estimating the proportion of adaptive amino acid substitutions. :2022.08.15.504017. Available from: https://www.biorxiv.org/content/10.1101/2022.08.15.504017v1

Árnason E, Koskela J, Halldórsdóttir K, Eldon B. 2023. Sweepstakes reproductive success via pervasive and recurrent selective sweeps.Gagnaire P-A, Przeworski M, Freund F, Gagnaire P-A, editors. *eLife* 12:e80781.

Atkinson A, Hill SL, Pakhomov EA, Siegel V, Reiss CS, Loeb VJ, Steinberg DK, Schmidt K, Tarling GA, Gerrish L, et al. 2019. Krill (Euphausia superba) distribution contracts southward during rapid regional warming. *Nat. Clim. Change* 9:142.

Atkinson A, Shreeve RS, Hirst AG, Rothery P, Tarling GA, Pond DW, Korb RE, Murphy EJ, Watkins JL. 2006. Natural growth rates in Antarctic krill (Euphausia superba): II. Predictive models based on food, temperature, body length, sex, and maturity stage. *Limnol. Oceanogr.* 51:973–987.

Atkinson A, Siegel V, Pakhomov EA, Jessopp MJ, Loeb V. 2009. A re-appraisal of the total biomass and annual production of Antarctic krill. *Deep Sea Res. Part Oceanogr. Res. Pap.* 56:727–740.

Bailey J, Durbin EG, Rynearson T. 2016. Species composition and abundance of copepods in the morphologically cryptic genus Pseudocalanus in the Bering Sea. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 134:173–180.

Baker A de C, Boden BP, Brinton E. 1990. A Practical Guide to the Euphausiids of the World. Natural History Museum

Bandell M, Macpherson LJ, Patapoutian A. 2007. From chills to chilis: mechanisms for thermosensation and chemesthesis via thermoTRPs. *Curr. Opin. Neurobiol.* 17:490–497.

Baxter JM, Laffoley D. 2016. Explaining ocean warming : causes, scale, effects and consequences. Gland : IUCN

Biscontin A, Frigato E, Sales G, Mazzotta GM, Teschke M, De Pittà C, Jarman S, Meyer B, Costa R, Bertolucci C. 2016. The opsin repertoire of the Antarctic krill Euphausia superba. *Mar. Genomics* 29:61–68.

Biscontin A, Wallach T, Sales G, Grudziecki A, Janke L, Sartori E, Bertolucci C, Mazzotta G, De Pittà C, Meyer B, et al. 2017. Functional characterization of the circadian clock in the Antarctic krill, Euphausia superba. *Sci. Rep.* 7:17742.

Bortolotto E, Bucklin A, Mezzavilla M, Zane L, Patarnello T. 2011. Gone with the currents: lack of genetic differentiation at the circum-continental scale in the Antarctic krill Euphausia superba. *BMC Genet.* 12:32.

Boynton PJ, Janzen T, Greig D. 2018. Modeling the contributions of chromosome segregation errors and aneuploidy to Saccharomyces hybrid sterility. *Yeast Chichester Engl.* 35:85–98.

Brinton E, Ohman MD, Townsend AW, Knight MD, Bridgeman AL. 2000. Euphausiids of the World Ocean. Springer Berlin Heidelberg

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12:59–60.

Buchholz F, Buchholz C. 2010. Chapter Six - Growth and Moulting in Northern Krill (Meganyctiphanes norvegica Sars). In: Tarling GA, editor. Advances in Marine Biology. Vol. 57. Academic Press. p. 173–197. Available from: http://www.sciencedirect.com/science/article/pii/B9780123813084000066

Bucklin A, DiVito KR, Smolina I, Choquet M, Questel JM, Hoarau G, O'Neill RJ. 2018. Population Genomics of Marine Zooplankton. In: Oleksiak MF, Rajora OP, editors. Population

Genomics: Marine Organisms. Population Genomics. Cham: Springer International Publishing. p. 61–102. Available from: http://link.springer.com/10.1007/13836_2017_9

Bucklin A, Lindeque PK, Rodriguez-Ezpeleta N, Albaina A, Lehtiniemi M. 2016. Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *J. Plankton Res.* 38:393–400.

Bucklin A, Peijnenburg Katja T C A, Kosobokova K, Machida RJ. 2021. New insights into biodiversity, biogeography, ecology, and evolution of marine zooplankton based on molecular approaches. *ICES J. Mar. Sci.* 78:3281–3287.

Bucklin A, Peijnenburg Katja T. C. A., Kosobokova KN, O'Brien TD, Blanco-Bercial L, Cornils A, Falkenhaug T, Hopcroft RR, Hosia A, Laakmann S, et al. 2021. Toward a global reference database of COI barcodes for marine zooplankton. *Mar. Biol.* 168:78.

Bucklin A, Questel Jennifer M, Blanco-Bercial L, Frenzel A, Smolenack SB, Wiebe PH. 2021. Population connectivity of the euphausiid, Stylocheiron elongatum, in the Gulf Stream (NW Atlantic Ocean) in relation to COI barcode diversity of Stylocheiron species. *ICES J. Mar. Sci.* 78:3464–3476.

Bucklin A, Wiebe PH, Smolenack SB, Copley NJ, Beaudet JG, Bonner KG, Färber-Lorda J, Pierson JJ. 2007. DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). *J. Plankton Res.* 29:483–493.

Cafferty P, Yu L, Long H, Rao Y. 2006. Semaphorin-1a Functions as a Guidance Receptor in the Drosophila Visual System. *J. Neurosci.* 26:3999–4003.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.

Capblancq T, Fitzpatrick M, Bay R, Exposito-Alonso M, Keller S. 2020. Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. *Annu. Rev. Ecol. Evol. Syst.* 51.

Cascella K, Jollivet D, Papot C, Léger N, Corre E, Ravaux J, Clark MS, Toullec J-Y. 2015. Diversification, Evolution and Sub-Functionalization of 70kDa Heat-Shock Proteins in Two Sister Species of Antarctic Krill: Differences in Thermal Habitats, Responses and Implications under Climate Change. *PLOS ONE* 10:e0121642.

Castresana J. 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol. Biol. Evol.* 17:540–552.

Cavan EL, Belcher A, Atkinson A, Hill SL, Kawaguchi S, McCormack S, Meyer B, Nicol S, Ratnarajah L, Schmidt K, et al. 2019. The importance of Antarctic krill in biogeochemical cycles. *Nat. Commun.* 10:4742.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7.

Charlesworth B, Jensen JD. 2022. How Can We Resolve Lewontin's Paradox? *Genome Biol. Evol.* 14:evac096.

Charlesworth J, Eyre-Walker A. 2008. The McDonald–Kreitman Test and Slightly Deleterious Mutations. *Mol. Biol. Evol.* 25:1007–1015.

Chen B, Feder ME, Kang L. 2018. Evolution of heat-shock protein expression underlying adaptive responses to environmental stress. *Mol. Ecol.* 27:3040–3054.

Choo LQ, Spagliardi G, Malinsky M, Choquet M, Goetze E, Hoarau G, Peijnenburg KTCA. 2023. Genome-wide phylogeography reveals cryptic speciation in the circumglobal planktonic calcifier Limacina bulimoides. *Mol. Ecol.* [Internet] n/a. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.16931

Choquet M, Kosobokova K, Kwaśniewski S, Hatlebakk M, Dhanasiri AKS, Melle W, Daase M, Svensen C, Søreide JE, Hoarau G. 2018. Can morphology reliably distinguish between the copepods Calanus finmarchicus and C. glacialis, or is DNA the only way? *Limnol. Oceanogr. Methods* 16:237–252.

Choquet M, Smolina I, Dhanasiri AKS, Blanco-Bercial L, Kopp M, Jueterbock A, Sundaram AYM, Hoarau G. 2019. Towards population genomics in non-model species with large genomes: a case study of the marine zooplankton Calanus finmarchicus. *R. Soc. Open Sci.* [Internet] 6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6408391/

Christie AE, Yu A, Pascual MG. 2017. Circadian signaling in the Northern krill Meganyctiphanes norvegica: In silico prediction of the protein components of a putative clock system using a

989 publicly accessible transcriptome. *Mar. Genomics*.
990 Cimino MA, Santora JA, Schroeder I, Sydeman W, Jacox MG, Hazen EL, Bograd SJ. 2020. Essential
991 krill species habitat resolved by seasonal upwelling and ocean circulation models within the
992 large marine ecosystem of the California Current System. *Ecography* 43:1536–1549.
993 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A
994 program for annotating and predicting the effects of single nucleotide polymorphisms,
995 SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*
996 *(Austin)* 6:80–92.
997 Clarke A, Crame JA, Stromberg J-O, Barker PF. 1992. The Southern Ocean Benthic Fauna and
998 Climate Change: A Historical Perspective [and Discussion]. *Philos. Trans. Biol. Sci.*
999 338:299–309.
1000 Clem KR, Fogt RL, Turner J, Lintner BR, Marshall GJ, Miller JR, Renwick JA. 2020. Record
1001 warming at the South Pole during the past three decades. *Nat. Clim. Change* 10:762–770.
1002 Cuellar J, Yébenes H, Parker SK, Carranza G, Serna M, Valpuesta JM, Zabala JC, Detrich HW. 2014.
1003 Assisted protein folding at low temperature: evolutionary adaptation of the Antarctic fish
1004 chaperonin CCT and its client proteins. *Biol. Open* 3:261–270.
1005 Cuzin-Roudy J, Irisson J-O, Penot F, Kawaguchi S, Vallet C. 2014. Southern Ocean Euphausiids. In:
1006 De Broyer C, Koubbi P, Griffiths HJ, Raymond B, d'Udekem d'Acoz C, Van de Putte AP,
1007 Danis B, David B, Grant S, Gutt J, et al., editors. Biogeographic Atlas of the Southern Ocean.
1008 Cambridge UK: Scientific Committee on Antarctic Research. p. 309–320.
1009 Cvijović I, Good BH, Desai MM. 2018. The Effect of Strong Purifying Selection on Genetic
1010 Diversity. *Genetics* 209:1235–1278.
1011 Dam HG. 2013. Evolutionary Adaptation of Marine Zooplankton to Global Change. *Annu. Rev. Mar.*
1012 *Sci.* 5:349–370.
1013 D'Amico S, Claverie P, Collins T, Georlette D, Gratia E, Hoyoux A, Meuwis M-A, Feller G, Gerday
1014 C. 2002. Molecular basis of cold adaptation. *Philos. Trans. R. Soc. B Biol. Sci.* 357:917–925.
1015 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G,
1016 Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics*
1017 27:2156–2158.
1018 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
1019 McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools.
1020 *GigaScience* 10:giab008.
1021 De Wit P, Pespeni MH, Palumbi SR. 2015. SNP genotyping and population genomics from expressed
1022 sequences – current advances and future possibilities. *Mol. Ecol.* 24:2310–2323.
1023 Deagle BE, Faux C, Kawaguchi S, Meyer B, Jarman SN. 2015. Antarctic krill population genomics:
1024 apparent panmixia, but genome complexity and large population size muddy the water. *Mol.*
1025 *Ecol.* 24:4943–4959.
1026 Dhaka A, Viswanath V, Patapoutian A. 2006. Trp ion channels and temperature sensation. *Annu. Rev.*
1027 *Neurosci.* 29:135–161.
1028 Doney SC, Ruckelshaus M, Duffy JE, Barry JP, Chan F, English CA, Galindo HM, Grebmeier JM,
1029 Hollowed AB, Knowlton N, et al. 2012. Climate change impacts on marine ecosystems.
1030 *Annu. Rev. Mar. Sci.* 4:11–37.
1031 Edwards M, Hélaouët P, Goberville E, Lindley A, Tarling GA, Burrows MT, Atkinson A. 2021.
1032 North Atlantic warming over six decades drives decreases in krill abundance with no
1033 associated range shift. *Commun. Biol.* 4:1–10.
1034 Ellegren H, Galtier N. 2016. Determinants of genetic diversity. *Nat. Rev. Genet.* 17:422.
1035 Eyre-Walker A, Keightley PD. 2009. Estimating the Rate of Adaptive Molecular Evolution in the
1036 Presence of Slightly Deleterious Mutations and Population Size Change. *Mol. Biol. Evol.*
1037 26:2097–2108.
1038 Eyre-Walker A, Woolfit M, Phelps T. 2006. The Distribution of Fitness Effects of New Deleterious
1039 Amino Acid Mutations in Humans. *Genetics* 173:891–900.
1040 Falk-Petersen S, Hagen W, Kattner G, Clarke A, Sargent J. 2000. Lipids, trophic relationships, and
1041 biodiversity in Arctic and Antarctic krill. *Can. J. Fish. Aquat. Sci.* 57:178–191.
1042 Fay JC, Wyckoff GJ, Wu CI. 2001. Positive and negative selection on the human genome. *Genetics*
1043 158:1227–1234.

Fitzpatrick MC, Keller SR. 2015. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* 18:1–16.
Flores H, Atkinson A, Kawaguchi S, Krafft BA, Milinevsky G, Nicol S, Reiss C, Tarling GA, Werner R, Rebolledo EB, et al. 2012. Impact of climate change on Antarctic krill. *Mar. Ecol. Prog. Ser.* 458:1–19.
Fuller ZL, Mocellin VJL, Morris LA, Cantin N, Shepherd J, Sarre L, Peng J, Liao Y, Pickrell J, Andolfatto P, et al. 2020. Population genetics of the coral Acropora millepora: Toward genomic prediction of bleaching. *Science* [Internet] 369. Available from: https://science.sciencemag.org/content/369/6501/eaba4674
Galtier N. 2016. Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis. *PLOS Genet.* 12:e1005774.
Ge SX, Jung D, Yao R. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36:2628–2629.
González HE, Graeve M, Kattner G, Silva N, Castro L, Iriarte JL, Osmán L, Daneri G, Vargas CA. 2016. Carbon flow through the pelagic food web in southern Chilean Patagonia: relevance of Euphausia vallentini as a key species. *Mar. Ecol. Prog. Ser.* 557:91–110.
Gouveia D, Bonneton F, Almunia C, Armengaud J, Quéau H, Degli-Esposti D, Geffard O, Chaumot A. 2018. Identification, expression, and endocrine-disruption of three ecdysone-responsive genes in the sentinel species Gammarus fossarum. *Sci. Rep.* 8:3793.
Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29:644–652.
Guéguen L, Gaillard S, Boussau B, Gouy M, Groussin M, Rochette NC, Bigot T, Fournier D, Pouyet F, Cahais V, et al. 2013. Bio++: Efficient Extensible Libraries and Tools for Computational Molecular Evolution. *Mol. Biol. Evol.* 30:1745–1750.
Häfker NS, Andreatta G, Manzotti A, Falciatore A, Raible F, Tessmar-Raible K. 2023. Rhythms and Clocks in Marine Organisms. *Annu. Rev. Mar. Sci.* 15:509–538.
Hamaguchi S, Sakaizumi M. 1992. Sexually differentiated mechanisms of sterility in interspecific hybrids between Oryzias latipes and O. curvinotus. *J. Exp. Zool.* 263:323–329.
Han F, Jamsandekar M, Pettersson ME, Su L, Fuentes-Pardo AP, Davis BW, Bekkevold D, Berg F, Casini M, Dahle G, et al. 2020. Ecological adaptation in Atlantic herring is associated with large shifts in allele frequencies at hundreds of loci.Flint J, Wittkopp PJ, editors. *eLife* 9:e61076.
Hansen HJ. 1911. The genera and species of the Order Euphausiacea, with account of remarkable variation. *Bull. Inst. Océan. Monaco* 210:1–54.
Harkins GW, D'Amato ME, Gibbons MJ. 2013. Self-maintaining or continuously refreshed? The genetic structure of Euphausia lucens populations in the Benguela upwelling ecosystem. *J. Plankton Res.* 35:982–992.
Hedgecock D. 1994. Does variance in reproductive success limit effective population size of marine organisms. In: Genetics and Evolution of Aquatic Organisms. Beaumont A, editors. London: Chapman & Hall. p. 122–134.
Hedgecock D, Pudovkin AI. 2011. Sweepstakes Reproductive Success in Highly Fecund Marine Fish and Shellfish: A Review and Commentary. *Bull. Mar. Sci.* 87:971–1002.
Henze MJ, Oakley TH. 2015. The Dynamic Evolutionary History of Pancrustacean Eyes and Opsins. *Integr. Comp. Biol.* 55:830–842.
Himmel NJ, Cox DN. 2020. Transient receptor potential channels: current perspectives on evolution, structure, function and nomenclature. *Proc. R. Soc. B Biol. Sci.* 287:20201309.
Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35:518–522.
Hoffmann AA, Sgrò CM. 2011. Climate change and evolutionary adaptation. *Nature* 470:479–485.
Hoffstaetter LJ, Bagriantsev SN, Gracheva EO. 2018. TRPs et al.: a molecular toolkit for thermosensory adaptations. *Pflugers Arch.* 470:745–759.
Höring F, Biscontin A, Harms L, Sales G, Reiss CS, De Pittà C, Meyer B. 2021. Seasonal gene expression profiling of Antarctic krill in three different latitudinal regions. *Mar. Genomics*

56:100806.
Huenerlage K, Cascella K, Corre E, Toomey L, Lee C-Y, Buchholz F, Toullec J-Y. 2016. Responses of the arcto-boreal krill species Thysanoessa inermis to variations in water temperature: coupling Hsp70 isoform expressions with metabolism. *Cell Stress Chaperones* 21:969–981.
Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* 33:1635–1638.
Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254–267.
Hvilsom C, Qian Y, Bataillon T, Li Y, Mailund T, Sallé B, Carlsen F, Li R, Zheng H, Jiang T, et al. 2012. Extensive X-linked adaptive evolution in central chimpanzees. *Proc. Natl. Acad. Sci.* 109:2054–2059.
Hwang RY, Stearns NA, Tracey WD. 2012. The Ankyrin Repeat Domain of the TRPA Protein Painless Is Important for Thermal Nociception but Not Mechanical Nociception. *PLOS ONE* 7:e30090.
Intergovernmental Panel on Climate Change ed. 2014. Climate Change 2013 - The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press Available from: http://ebooks.cambridge.org/ref/id/CBO9781107415324
Jang W, Kim JY, Cui S, Jo J, Lee B-C, Lee Y, Kwon K-S, Park C-S, Kim C. 2015. The Anoctamin Family Channel Subdued Mediates Thermal Nociception in Drosophila*. *J. Biol. Chem.* 290:2521–2528.
Jeffery NW. 2012. The first genome size estimates for six species of krill (Malacostraca, Euphausiidae): large genomes at the north and south poles. *Polar Biol.* 35:959–962.
John DD. 1936. The southern species of the genus Euphausia. *Discov. Rep.* 14:393–324.
Johnston NM, Murphy EJ, Atkinson A, Constable AJ, Cotté C, Cox M, Daly KL, Driscoll R, Flores H, Halfter S, et al. 2022. Status, Change, and Futures of Zooplankton in the Southern Ocean. *Front. Ecol. Evol.* [Internet] 9. Available from: https://www.frontiersin.org/articles/10.3389/fevo.2021.624692
Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14:587–589.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
Katoh K, Standley DM. 2016. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* 32:1933–1942.
Koubbi P, Hulley P, Raymond B, Penot F, Gasparini S, Labat J-P, Pruvost P, Mormede S, Irisson J-O, Duhamel G, et al. 2011. Estimating the biodiversity of the subantarctic Indian part for ecoregionalisation: Part I. Pelagic realm of CCAMLR areas 58.5.1. and 58.6. Report of the CCAMLR WS-MPA, Brest, France.
Kozma MT, Ngo-Vu H, Wong YY, Shukla NS, Pawar SD, Senatore A, Schmidt M, Derby CD. 2020. Comparison of transcriptomes from two chemosensory organs in four decapod crustaceans reveals hundreds of candidate chemoreceptor proteins. *PLOS ONE* 15:e0230266.
Krasovec M, Rickaby REM, Filatov DA. 2020. Evolution of Mutation Rate in Astronomically Large Phytoplankton Populations. *Genome Biol. Evol.* 12:1051–1059.
Kumar S, Chen D, Jang C, Nall A, Zheng X, Sehgal A. 2014. An ecdysone-responsive nuclear receptor regulates circadian rhythms in Drosophila. *Nat. Commun.* 5:5697.
Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, Garapati PV, Goodman JL, Gramates LS, Millburn G, Strelets VB, et al. 2021. FlyBase: updates to the Drosophila melanogaster knowledge base. *Nucleic Acids Res.* 49:D899–D907.
Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124.
Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, Venkat A, Andolfatto P, Przeworski M. 2012. Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLOS Biol.* 10:e1001388.
Lenormand T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183–189.
Lenz PH, Lieberman B, Cieslak MC, Roncalli V, Hartline DK. 2021. Transcriptomics and

metatranscriptomics in zooplankton: wave of the future? *J. Plankton Res.* 43:3–9.
Lewontin RC. 1974. The Genetic Basis of Evolutionary Change. 1st edition. New York: Columbia Univ Pr
Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio* [Internet]. Available from: http://arxiv.org/abs/1303.3997
Lourenço JM, Glémin S, Galtier N. 2013. The Rate of Molecular Adaptation in a Changing Environment. *Mol. Biol. Evol.* 30:1292–1301.
Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, et al. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48:D265–D268.
Magkrioti CK, Spyropoulos IC, Iconomidou VA, Willis JH, Hamodrakas SJ. 2004. cuticleDB: a relational database of Arthropod cuticular proteins. *BMC Bioinformatics* 5:138.
Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12.
Mayer MP, Bukau B. 2005. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* 62:670–684.
Mayzaud P, Boutoute M, Alonzo F. 2003. Lipid composition of the euphausiids Euphausia vallentini and Thysanoessa macrura during summer in the Southern Indian Ocean. *Antarct. Sci.* 15:463–475.
McBride MM, Dalpadado P, Drinkwater KF, Godø OR, Hobday AJ, Hollowed AB, Kristiansen T, Murphy EJ, Ressler PH, Subbey S, et al. 2014. Krill, climate, and contrasting future scenarios for Arctic and Antarctic fisheries. *ICES J. Mar. Sci.* 71:1934–1955.
McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. *Nature* 351:652–654.
Meek MH, Beever EA, Barbosa S, Fitzpatrick SW, Fletcher NK, Mittan-Moreau CS, Reid BN, Campbell-Staton SC, Green NF, Hellmann JJ. 2023. Understanding Local Adaptation to Prepare Populations for Climate Change. *BioScience* 73:36–47.
Messer PW, Petrov DA. 2013. Frequent adaptation and the McDonald–Kreitman test. *Proc. Natl. Acad. Sci.* 110:8615–8620.
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. *Mol. Biol. Evol.* 37:1530–1534.
Montecinos AE, Guillemin M-L, Couceiro L, Peters AF, Stoeckel S, Valero M. 2017. Hybridization between two cryptic filamentous brown seaweeds along the shore: analysing pre- and postzygotic barriers in populations of individuals with varying ploidy levels. *Mol. Ecol.* 26:3497–3512.
Moutinho AF, Bataillon T, Dutheil JY. 2020. Variation of the adaptive substitution rate between species and within genomes. *Evol. Ecol.* 34:315–338.
Mugal CF, Kutschera VE, Botero-Castro F, Wolf JBW, Kaj I. 2020. Polymorphism Data Assist Estimation of the Nonsynonymous over Synonymous Fixation Rate Ratio ω for Closely Related Species. *Mol. Biol. Evol.* 37:260–279.
Nielsen R, Yang Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148:929–936.
Ollier A, Chabot D, Audet C, Winkler G. 2018. Metabolic rates and spontaneous swimming activity of two krill species (Euphausiacea) under different temperature regimes in the St. Lawrence Estuary, Canada. *J. Crustac. Biol.* 38:697–706.
Palecanda S, Iwanicki T, Steck M, Porter ML. 2022. Crustacean conundrums: a review of opsin diversity and evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 377:20210289.
Palma S, Silva N. 2004. Distribution of siphonophores, chaetognaths, euphausiids and oceanographic conditions in the fjords and channels of southern Chile. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 51:513–535.
Papot C, Cascella K, Toullec J-Y, Jollivet D. 2016. Divergent ecological histories of two sister Antarctic krill species led to contrasted patterns of genetic diversity in their heat-shock protein (hsp70) arsenal. *Ecol. Evol.* 6:1555–1575.
Peck LS. 2016. A Cold Limit to Adaptation in the Sea. *Trends Ecol. Evol.* 31:13–26.

Peck LS, Morley SA, Richard J, Clark MS. 2014. Acclimation and thermal tolerance in Antarctic marine ectotherms.Davies SA, Dow JAT, Lukowiak K, editors. *J. Exp. Biol.* 217:16–22.
Peijnenburg KTCA, Goetze E. 2013. High evolutionary potential of marine zooplankton. *Ecol. Evol.* 3:2765–2781.
Perry FA, Kawaguchi S, Atkinson A, Sailley SF, Tarling GA, Mayor DJ, Lucas CH, King R, Cooper A. 2020. Temperature–Induced Hatch Failure and Nauplii Malformation in Antarctic Krill. *Front. Mar. Sci.* [Internet] 7. Available from: https://www.frontiersin.org/article/10.3389/fmars.2020.00501
Phleger CF, Nelson MM, Mooney BD, Nichols PD. 2002. Interannual and between species comparison of the lipids, fatty acids and sterols of Antarctic krill from the US AMLR Elephant Island survey area. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 131:733–747.
Piccolin F, Suberg L, King R, Kawaguchi S, Meyer B, Teschke M. 2018. The Seasonal Metabolic Activity Cycle of Antarctic Krill (Euphausia superba): Evidence for a Role of Photoperiod in the Regulation of Endogenous Rhythmicity. *Front. Physiol.* [Internet] 9. Available from: https://www.frontiersin.org/articles/10.3389/fphys.2018.01715
Picollo A, Malvezzi M, Accardi A. 2015. TMEM16 Proteins: Unknown Structure and Confusing Functions. *J. Mol. Biol.* 427:94–105.
Poloczanska ES, Brown CJ, Sydeman WJ, Kiessling W, Schoeman DS, Moore PJ, Brander K, Bruno JF, Buckley LB, Burrows MT, et al. 2013. Global imprint of climate change on marine life. *Nat. Clim. Change* 3:919–925.
Poloczanska ES, Burrows MT, Brown CJ, García Molinos J, Halpern BS, Hoegh-Guldberg O, Kappel CV, Moore PJ, Richardson AJ, Schoeman DS, et al. 2016. Responses of Marine Organisms to Climate Change across Oceans. *Front. Mar. Sci.* [Internet] 3. Available from: http://journal.frontiersin.org/article/10.3389/fmars.2016.00062/abstract
Popovic I, Riginos C. 2020. Comparative genomics reveals divergent thermal selection in warm- and cold-tolerant marine mussels. *Mol. Ecol.* 29:519–535.
Pörtner HO, Peck L, Somero G. 2007. Thermal limits and adaptation in marine Antarctic ectotherms: an integrative view. *Philos. Trans. R. Soc. B Biol. Sci.* 362:2233–2258.
Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments.Poon AFY, editor. *PLoS ONE* 5:e9490.
Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinforma.* 70:e102.
Pucciarelli S, Parker SK, Detrich HW, Melki R. 2006. Characterization of the cytoplasmic chaperonin containing TCP-1 from the Antarctic fish Notothenia coriiceps. *Extremophiles* 10:537–549.
Rantanen M, Karpechko AY, Lipponen A, Nordling K, Hyvärinen O, Ruosteenoja K, Vihma T, Laaksonen A. 2022. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* 3:1–10.
Ratnarajah L, Abu-Alhaija R, Atkinson A, Batten S, Bax NJ, Bernard KS, Canonico G, Cornils A, Everett JD, Grigoratou M, et al. 2023. Monitoring and modelling marine zooplankton in a changing climate. *Nat. Commun.* 14:564.
Razgour O, Forester B, Taggart JB, Bekaert M, Juste J, Ibáñez C, Puechmaille SJ, Novella-Fernandez R, Alberdi A, Manel S. 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proc. Natl. Acad. Sci.* 116:10418–10423.
Richardson AJ. 2008. In hot water: zooplankton and climate change. *ICES J. Mar. Sci.* 65:279–295.
Ridoux V. 1988. Subantarctic krill, Euphausia vallentini Stebbing, preyed upon by penguins around Crozet Islands (Southern Indian Ocean): population structure and annual cycle. *J. Plankton Res.* 10:675–690.
Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, Faivre N, et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515:261–263.
Rousselle M, Simion P, Tilak M-K, Figuet E, Nabholz B, Galtier N. 2020. Is adaptation limited by mutation? A timescale-dependent effect of genetic diversity on the adaptive substitution rate in animals. *PLOS Genet.* 16:e1008668.
Russell FS, Yonge CM, Mauchline J, Fisher LR. 1969. Adv. Mar. Biol. 7: The biology of

euphausiids. *Adv. Mar. Biol.*
Savolainen O, Lascoux M, Merilä J. 2013. Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14:807–820.
Schram FR, Koenemann S. 2022. Euphausiacea. In: Schram FR, Koenemann S, editors. Evolution and Phylogeny of Pancrustacea: A Story of Scientific Method. Oxford University Press. p. 0. Available from: https://doi.org/10.1093/oso/9780195365764.003.0035
Seear P, Goodall-Copestake WP, Fleming AH, Rosato E, Tarling GA. 2012. Seasonal and spatial influences on gene expression in Antarctic krill Euphausia superba. *Mar. Ecol. Prog. Ser.* 467:61–75.
Seear PJ, Tarling GA, Burns G, Goodall-Copestake WP, Gaten E, Özkaya Ö, Rosato E. 2010. Differential gene expression during the moult cycle of Antarctic krill (Euphausia superba). *BMC Genomics* 11:582.
Shao C, Sun S, Liu K, Wang Jiahao, Li S, Liu Q, Deagle BE, Seim I, Biscontin A, Wang Q, et al. 2023. The enormous repetitive Antarctic krill genome reveals environmental adaptations and population insights. *Cell* [Internet]. Available from: https://www.sciencedirect.com/science/article/pii/S0092867423001071
Shen H-C, Chu S-Y, Hsu T-C, Wang C-H, Lin I-Y, Yu H-H. 2017. Semaphorin-1a prevents Drosophila olfactory projection neuron dendrites from mis-targeting into select antennal lobe regions. *PLoS Genet.* 13:e1006751.
Siegel V. 1987. Age and growth of Antarctic Euphausiacea (Crustacea) under natural conditions. *Mar. Biol.* 96:483–495.
Siegel V. 2000. Krill (Euphausiacea) life history and aspects of population dynamics. *Can. J. Fish. Aquat. Sci.* 57:130–150.
Siegel V ed. 2016. Biology and Ecology of Antarctic Krill. Springer International Publishing Available from: //www.springer.com/la/book/9783319292779
Silliman K, Indorf JL, Knowlton N, Browne WE, Hurt C. 2021. Base-substitution mutation rate across the nuclear genome of Alpheus snapping shrimp and the timing of isolation by the Isthmus of Panama. *BMC Ecol. Evol.* 21:104.
Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212.
Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in Drosophila. *Nature* 415:1022–1024.
Smith SA, Dunn CW. 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* 24:715–716.
Somer L, Shmulman O, Dror T, Hashmueli S, Kashi Y. 2002. The eukaryote chaperonin CCT is a cold shock protein in Saccharomyces cerevisiae. *Cell Stress Chaperones* 7:47–54.
Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JGR, Korf I, Lapp H, et al. 2002. The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Res.* 12:1611–1618.
Steenwyk JL, Goltz DC, Iii TJB, Li Y, Shen X-X, Rokas A. 2022. OrthoSNAP: A tree splitting and pruning algorithm for retrieving single-copy orthologs from gene family trees. *PLOS Biol.* 20:e3001827.
Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M. 2012. Drift-barrier hypothesis and mutation-rate evolution. *Proc. Natl. Acad. Sci.* 109:18488–18492.
Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms.Gibas C, editor. *PLoS ONE* 6:e21800.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34:W609–W612.
Tan G-M, Xu L, Bu D-B, Feng S-Z, Sun N-H. 2006. Improvement of Performance of MegaBlast Algorithm for DNA Sequence Alignment. *J. Comput. Sci. Technol.* 21:973–978.
Tarling G. 2010. Biology of Northern Krill. Academic Press
Teixeira JC, Huber CD. 2021. The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci.* [Internet] 118. Available from: https://www.pnas.org/content/118/10/e2015096118
Teschke M, Kawaguchi S, Meyer B. 2008. Effects of simulated light regimes on maturity and body

composition of Antarctic krill, Euphausia superba. *Mar. Biol.* 154:315–324.
Toullec J-Y, Cascella K, Ruault S, Geffroy A, Lorieux D, Montagné N, Ollivaux C, Lee C-Y. 2020. Antarctic krill (Euphausia superba) in a warming ocean: thermotolerance and deciphering Hsp70 responses. *Cell Stress Chaperones* 25:519–531.
Toullec J-Y, Corre E, Bernay B, Thorne MAS, Cascella K, Ollivaux C, Henry J, Clark MS. 2013. Transcriptome and Peptidome Characterisation of the Main Neuropeptides and Peptidic Hormones of a Euphausiid: The Ice Krill, Euphausia crystallorophias. *PLOS ONE* 8:e71609.
Tsagkogeorga G, Cahais V, Galtier N. 2012. The Population Genomics of a Fast Evolver: High Levels of Diversity, Functional Constraint, and Molecular Adaptation in the Tunicate Ciona intestinalis. *Genome Biol. Evol.* 4:852–861.
Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 43:W401-407.
Urso I, Biscontin A, Corso D, Bertolucci C, Romualdi C, De Pittà C, Meyer B, Sales G. 2022. A thorough annotation of the krill transcriptome offers new insights for the study of physiological processes. *Sci. Rep.* 12:11415.
Vereshchaka AL, Kulagin DN, Lunina AA. 2019. A phylogenetic study of krill (Crustacea: Euphausiacea) reveals new taxa and co-evolution of morphological characters. *Cladistics* 35:150–172.
Voolstra O, Oberhauser V, Sumser E, Meyer NE, Maguire ME, Huber A, von Lintig J. 2010. NinaB is essential for Drosophila vision but induces retinal degeneration in opsin-deficient photoreceptors. *J. Biol. Chem.* 285:2130–2139.
Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7:256–276.
Wiebe PH, Bucklin A, Kaartvedt S, Røstad A, Blanco-Bercial L. 2016. Vertical distribution and migration of euphausiid species in the Red Sea. *J. Plankton Res.* 38:888–903.
Xiao R, Xu XZS. 2021. Temperature Sensation: From Molecular Thermosensors to Neural Circuits and Coding Principles. *Annu. Rev. Physiol.* 83:205–230.
Yang Z. 2006. Computational Molecular Evolution. Illustrated edition. Oxford: Oxford University Press
Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* 24:1586–1591.
Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17:32–43.
Yang Z, Nielsen R. 2002. Codon-Substitution Models for Detecting Molecular Adaptation at Individual Sites Along Specific Lineages. *Mol. Biol. Evol.* 19:908–917.
Yang Z, Wong WSW, Nielsen R. 2005. Bayes Empirical Bayes Inference of Amino Acid Sites Under Positive Selection. *Mol. Biol. Evol.* 22:1107–1118.
York JM, Zakon HH. 2022. Evolution of Transient Receptor Potential (TRP) Ion Channels in Antarctic Fishes (Cryonotothenioidea) and Identification of Putative Thermosensors. *Genome Biol. Evol.* 14:evac009.
Yuan K, Yuan F-H, He H-H, Bi H-T, Weng S-P, He J-G, Chen Y-H. 2017. Heat shock 70 kDa protein cognate 5 involved in WSSV toleration of Litopenaeus vannamei. *Dev. Comp. Immunol.* 72:9–20.
Yuan X, Zhu Y, Ruiz-Carrillo D, Kadowaki T. 2022. Evolutionary history of metazoan TMEM16 family. *Mol. Phylogenet. Evol.* 177:107595.
Zane L, Patarnello T. 2000. Krill: a possible model for investigating the effects of ocean currents on the genetic structure of a pelagic invertebrate. *Can. J. Fish. Aquat. Sci.* 57:16–23.
Zhang H, Wang C, Zhang K, Kamau PM, Luo A, Tian L, Lai R. 2022. The role of TRPA1 channels in thermosensation. *Cell Insight* 1:100059.
Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. *Mol. Biol. Evol.* 22:2472–2479.
Zhang Z, Li J, Zhao X-Q, Wang J, Wong GK-S, Yu J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics* 4:259–263.

46

Acknowledgements

This research was supported by grants 2018-04444 from the Swedish Research Council Vetenskapsrådet and the Inez Johanssons scholarship foundation awarded to AW. The computational infrastructure was enabled by project SNIC 2022/5-472 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX and the PDC Center for High Performance Computing, KTH Royal Institute of Technology, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. We thank Dr N. Tremblay (Quebec university, Rimouski UQAR) for kindly providing Pacific Ocean krill samples, Camille Merland (Université Pierre et Marie Curie) and Marion Thellier (LOCEAN-IPSL, CNRS) for help with species identifications, the helpful research vessel and marine station staff that supported field sampling, Julien Yann Dutheil (Max Planck Institute for Evolutionary Biology) for helpful discussions and Mårten Larsson (Uppsala University) for folding the Peropsin protein. The REPCCOAI (Responses of the Pelagic Ecosystem to Climate Change in the Southern and South Indian Ocean) surveys were directed by Philippe Koubbi (Roscoff Biological Station) and J-Y. T. These surveys were supported by the French oceanographic fleet, the CNRS Antarctic Workshop Zone, the CNES KERTREND-SAT OSTST project led by Francesco d'Ovidio (LOCEAN), the European MESOPP H2020 programme and the TAAF National Nature Reserve programmes. We thank Philippe Koubbi for the opportunity to take part in these surveys.

Author contributions

A. W. and J-Y. T. conceived and designed the study. A. W., J-Y. T., A. C., G. T., M. C., E. C. performed collection of data in the field, lab work or data curation. F. L. contributed software. A. W., M. C., J-Y. T. analyzed the data and wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Data availability

The new sequence data underlying this article are available in the European Nucleotide Archive at https://www.ebi.ac.uk/ena/browser/home, and can be accessed with project accession PRJXXXXXXX. Annotated SNP datasets, transcripts and orthology datasets are available in the

47

SciLifeLab Data Repository at XXXXXXXXXX.

Code availability

Code for basefinder is available on Github: https://github.com/fellen31/basefinder.

48