Proteomics reveals sex-specific heat shock response of Baikal amphipod *Eulimnogammarus cyaneus*

Daria Bedulina[1,2,*,#], Polina Drozdova[1,2,*], Anton Gurkov[1,2], Martin von Bergen[3,4], Peter F. Stadler[5-10], Till Luckenbach[11], Maxim Timofeyev[1,2], Stefan Kalkhof[3,12,13]

* contributed equally

corresponding author; daria.bedulina@gmail.com

1– Institute of Biology, Irkutsk State University, Lenin str. 3, Irkutsk, Russia
2– Baikal Research Centre, Lenin str. 21, Irkutsk, Russia
3– Department of Molecular Systems Biology, Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, Leipzig, Germany
4– Institute of Biochemistry, Faculty of Life Sciences, University of Leipzig, Germany
5– Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, Universität Leipzig, Härtelstraße 16-18, 04107 Leipzig, Germany.
6– Competence Center for Scalable Data Services and Solutions Dresden/Leipzig, Interdisciplinary Center for Bioinformatics, German Centre for Integrative Biodiversity Research (iDiv), and Leipzig Research Center for Civilization Diseases, Universität Leipzig, Augustusplatz 12, 04107 Leipzig, Germany.
7– Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany.
8– Department of Theoretical Chemistry, University of Vienna, Währinger Straße 17, 1090 Vienna, Austria.
9– Facultad de Ciencias Universidad National de Colombia, Sede Bogotá, Ciudad Universitaria, 111321 Bogotá, D.C. Colombia.
10– Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM87501 USA
11– Department of Bioanalytical Ecotoxicology, Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, 04318 Leipzig, Germany
12– Fraunhofer Institute for Cell Therapy and Immunology, Dept. Cell Engineering, Perlickstr. 1, 04103 Leipzig, Germany
13– Institute of Bioanalysis, University of Applied Sciences and Arts of Coburg, Friedrich-Streib-Str. 2, 96450 Coburg, Germany

Abstract

The ancient Lake Baikal is the largest source of liquid freshwater on Earth and home to a unique fauna. Several hundred mostly cold-adapted endemic amphipod species inhabit Baikal, an ecosystem that is already being influenced by global change. In this study, we characterized the core proteome and heat stress-induced changes in a temperature-tolerant endemic amphipod, *Eulimnogammarus cyaneus*, using a proteogenomic approach (PRIDE dataset PXD013237) to unravel the molecular mechanisms of the observed adverse effects. As males were previously found to be much more tolerant to thermal stress, we placed special emphasis on differences between the sexes. For both sexes, we observed adaption of energy metabolism, cytoskeleton, lipid, and carbohydrate metabolism upon heat stress. In contrast, significant differences were determined in the molecular chaperone response. Females from the control conditions possessed significantly higher levels of heat shock proteins (HSP70, HSPb1, Hsc70-3), which, in contrast to males, were not further increased in response to heat stress. The inability of females to further increase heat shock protein synthesis in response to temperature stress may be due to sex-specific processes, such as egg production, requiring a large proportion of the available energy. As ovigerous females synthesize generally higher amounts of protein, they also need higher levels of molecular chaperones for the folding of these new proteins. Thus, the higher sensitivity of females to heat shock may be due to the lack of molecular chaperone molecules to counteract the heat-induced protein denaturation.

Keywords

Proteogenomics, Amphipoda, sexual dimorphism, stress response, high-throughput proteomics

1. Introduction

Lake Baikal, the largest freshwater lake in the world and a UNESCO world heritage site, is inhabited by an exceptionally species-rich, largely endemic fauna, which is distinct from the faunas from other freshwaters in the Palearctic (Cristescu et al., 2010). With regard to species richness, extremely high abundance of individuals, and high total biomass, amphipods (Amphipoda, Crustacea) constitute a major, ecologically highly relevant animal taxon of Lake Baikal (Takhteev et al., 2015).

More than 350 endemic amphipod species and subspecies are described in Lake Baikal (Takhteev et al., 2015). They inhabit all kinds of substrates and all water depths of the lake (Timoshkin et al., 2001). Thus, Baikal amphipods are important as models in ecological and evolutionary studies (Takhteev et al., 2000). The endemic amphipods of Lake Baikal evolved under stable environmental conditions characterized by cold temperatures, low nutrient levels, high dissolved oxygen, and low mineralization (Mats et al., 2011).

Currently, some local benthic communities in the lake undergo rapid changes associated with growing human impact and with climate change, which increase the frequency and range of temperature fluctuation in the shallow littoral zone (Moore et al., 2009; Timoshkin et al., 2018). Our previous studies demonstrated high diversity in the tolerance and degrees of cellular responses to environmental stressors, such as fluctuating temperatures, in different amphipod species in the lake (Bedulina et al., 2013, Jakob et al., 2016, Axenov-Gribanov et al., 2016). Furthermore, sex-specific differences in temperature stress tolerance, as well as in the proteomic stress response, were shown for one of the most widespread upper-littoral species *Eulimnogammarus cyaneus* (Dybowsky, 1874). The responsiveness of the stress protein system was much higher in males, making them more resistant to temperature stress (Bedulina et al., 2017). However, the classical approach of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) followed by matrix-assisted laser desorption/ionization with tandem time-of-flight mass spectrometry (MALDI-ToF-ToF-MS/MS) only allowed for studying the high-abundant proteins.

Eulimnogammarus cyaneus, one of the most abundant species in the littoral zone of Lake Baikal (Timoshkin, 2001), has been extensively studied within the last decade (Bedulina et al., 2013; 2017; Jakob et al., 2016; Drozdova et al., 2019a; 2019b) and is considered a potential bioindicator for monitoring the ecological state of the benthic community (Timofeyev, 2010). *E. cyaneus* is a small (adult body size 11–15 mm) endemic amphipod, widespread along the shoreline of the entire lake. It is a representative of the upper littoral zone. This species inhabits the rocky shore of the lake with the highest population density at the narrow zone of the water

edge (Bazikalova, 1941). Reproduction takes place three times during the summer, starting in May when the first females with eggs appear (Gavrilov, 1949). This species has enhanced thermal tolerance compared to the majority of Baikal littoral species (Bedulina et al., 2013) and prefers the temperatures of approximately 11 to 12 °C (Timofeyev, Kirichenko, 2004). This species can be maintained in the laboratory for extended periods of time and is a good candidate to develop molecular markers and tools for environmental monitoring.

In this study, we aimed to investigate the proteome of this species for sexual dimorphism and response to acute thermal stress using LC-MS/MS for protein identification and quantification. Furthermore, protein identification was facilitated by the application of the species-specific transcriptomic database, which allowed us to identify and quantify 994 protein groups, to characterize the core male and females proteomes, and to deepen our understanding of the mechanism behind thermal stress response.

2. Methods

2.1. Experiment

Amphipods were collected in June 2014 with a hand net in the upper littoral zone in the shallow water right at the shore (0-0.1 m) near the coastal town of Listvyanka (51°84’85″ N, 104°88’37″ E). Animals were transferred to the laboratory in thermostatic boxes. Only individuals in amplexus (**Fig. 1A**) were selected for further study because in this state the sex of the amphipods could be easily determined without examination under the microscope (i.e., the upper individual in the amplexus is a male). The investigated females were on the second or third stage of their molting cycle (developed but hairless oostegites or with eggs in the marsupium) (Bazikalova et al., 1941). The males and females in amplexus were carefully separated, and males and females were separately allowed to acclimate in 2-L glass aquaria (at most 100 animals per aquarium) with constantly aerated lake water for 5 days at 6–7 °C (average annual temperature of Baikal littoral; Weiss et al. 1991; Falkner et al. 1991). During the acclimation period, the amphipods were fed with the commercial food TetraMin (Tetra, Germany) and a dried mix of algae and amphipods from their environment. The water was changed every other day. No mortality was observed during acclimation.

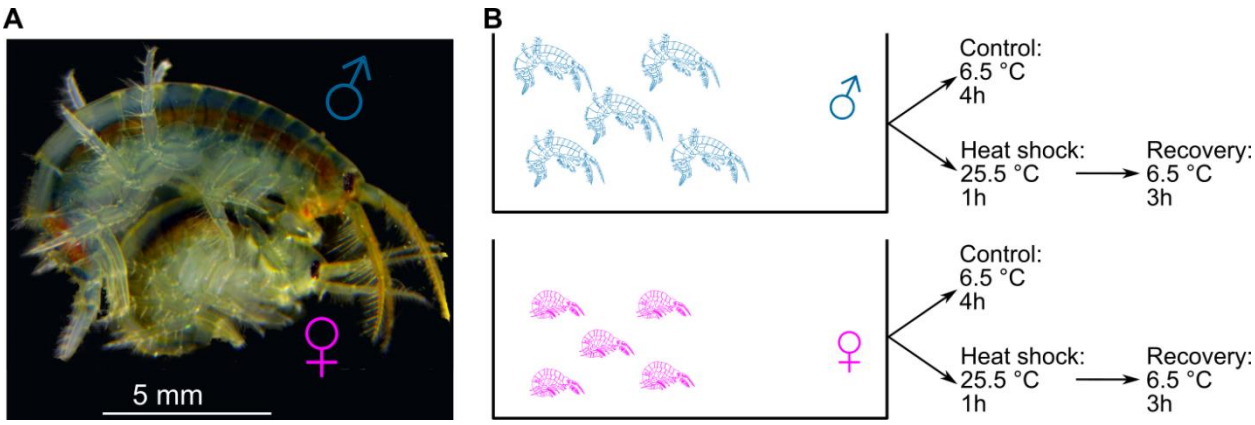


Fig. 1. A photograph of an *E. cyaneus* amplexus (A) and a schematic of the experimental heat shock treatment (B).

To determine the heat shock response, 5 replicates of each sex (4-5 specimens per replicate) were incubated in 1-L glass aquaria with well-aerated lake water, which was preheated to 25.5 °C, the temperature previously determined to cause 50% mortality of the mixed adult population (Bedulina et al., 2013). The individuals were exposed to heat shock for one hour and afterward transferred to a separate tank with water at 6–7 °C for a 3-hour recovery period, as we have previously shown that in some cases the recovery allows to more clearly discover the protein-level Hsp70 response (Bedulina et al., 2013). Following the recovery, the entire animals (4-5 specimens per sample) were flash-frozen in liquid nitrogen. The control animals kept at acclimation (6–7 °C) conditions throughout the experiment were shock-frozen in parallel at the end of the experiment (**Fig. 1B**).

2.2. Sample preparation

Protein isolation for the one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (1D-SDS-PAGE) was performed as described previously in Bedulina et al. (2013). Four to five specimens of *E. cyaneus* were pooled for each protein sample. SDS-PAGE was performed in a 70×80×0.75 mm^3 gel blocks (Laemmli, 1970) using a Mini-PROTEAN II Electrophoretic Cell (BIO-RAD CA, USA) apparatus for electrophoresis in a 10% polyacrylamide gel. Gels were loaded with 45 µg of total protein, stained with Coomassie Blue R (Invitrogen), and destained overnight in water.

The total protein from each gel lane was divided into six portions by cutting each gel lane with a scalpel into six slices. The proteins were reduced, alkylated, and subjected to tryptic in-gel digestion (100 ng bovine trypsin per slice, Roche, Germany) as described previously (Müller et

al., 2013). The resulting proteolytic peptide mixtures were eluted, dried, and resolubilized with 0.1% formic acid for nano-HPLC/nano-ESI-MS/MS analyses. The peptides were separated on a nano-HPLC system (nanoAcquity, Waters, Milford, MA, USA) connected to an LTQ Orbitrap XL ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA). MS analyses were performed in a data-dependent mode. In brief, 5 µL of each sample were loaded and first trapped and desalted online on a reverse-phase precolumn (nanoAcquity UPLC column, C_{18}, 180 µm×20 mm, 5 µm, Waters). Then, the peptides were separated using a 90-min gradient on a nanoscale reverse-phase precolumn (nanoAcquity UPLC column, C_{18}, 75 µm×150 mm, 1.7 µm, Waters) at a flow rate of 300 nL/min with a nonlinear gradient of CH_3CN, 0.1% formic acid (2 min, 2%; 7 min, 6%; 55 min 20%; 91 min, 40%; 94 min, 80% content of CH_3CN). The mass spectrometer automatically switched between full scan MS mode (m/z 300-1600, R=60,000) and tandem MS acquisition of 6 most intense ions of charge state 2+, 3+ or 4+. Peptide ions exceeding an intensity of 2,000 counts were fragmented within the linear ion trap by CID (isolation width 3 amu, normalized collision energy 35%, activation time 10 ms, activation Q 0.25). A dynamic precursor exclusion time of 2 min was applied for MS/MS measurements.

2.3. Proteogenomic database development

The database for protein identification was constructed from a *de novo* RNA sequencing data assembly described earlier (Drozdova et al., 2019b). Briefly, around 1 billion paired-end 100-base long Illumina reads of amphipods exposed to such stressors as increased temperature, $CdCl_2$, phenanthrene and acetone, as well as control animals, were quality trimmed with trim_galore (Krueger, 2017) v0.4.1, filtered with bowtie2 (Langmead and Salzberg, 2012) v2.1.0, corrected with rcorrector (Song and Florea, 2015), and assembled with Trinity (Grabherr et al., 2011) v2.5.1. The resulting assembly (EcyBCdTP1_all) contained 1,070,909 contigs. Assembly completeness was checked with BUSCO (Simão et al. 2015) v3.0.2 and was high with over 92% of complete conserved Arthropoda genes found in this assembly (Drozdova et al., 2019b). The proteome was predicted with the EvidentialGene pipeline (Gilbert, 2019) version 2013.07.27. We used the tr2aacds script and merged the 'okay' and 'alternative' sequence sets. The resulting protein database contained 167,948 sequences. These sequences were used as a reference database for peptide search.

2.4. Protein identification and quantitation

Protein identification was carried out with MaxQuant (Cox and Mann, 2008) v1.6 with the following parameters: match between runs checked; protein FDR, site FDR, and site decoy fraction set to 0.05; and all remaining parameters set to their default values. In total, 994 protein

groups (sets of proteins united by shared peptides) could be identified after filtering out those identified by less than two peptides or being potential contaminants and were subjected for subsequent analyses. Annotation of all identified protein sequences was performed with Blast2GO (Ashburner et al., 2000; Conesa et al., 2005) v5.0 and FunctionAnnotator (Chen et al., 2017), as well as with pantherScore v2.1.0 provided by Panther (Mi et al., 2019).

Statistical analysis of the normalized label-free quantification (LFQ) intensity values, (hereafter called protein intensities for brevity), was performed with custom R (R Core Team, 2017) scripts. Briefly, two analyses were performed. First, protein groups with qualitative differences in protein abundance were defined as those which were not identified in any sample of one group but detected in at least two samples of the second group. Second, protein groups present in all samples were subjected to quantitative analysis (experimental Bayes t-test) using the limma (Ritchie et al., 2015) and q-value (Storey et al., 2019) packages. The scripts are deposited at https://github.com/drozdovapb/code_chunks/tree/master/proteome.

The mass spectrometry proteomics data were deposited at the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al., 2012; Perez-Riverol et al. 2016) partner repository with the dataset identifier PXD013237.

3. Results

3.1. The core proteome of E. cyaneus

A proteogenomic workflow was used to define a core proteome *of E. cyaneus* (*i.e.* proteins that can be reliably and reproducibly detected in all individuals of the same sex of this particular species, according to Trapp et al. (2016) and Gouveia et al. (2019)) and to investigate the sex-specific adaptation to thermal stress. Both females and males were either subjected to thermal stress (treated group) or exposed to the acclimation temperature (control group). The lysis protocol, which has been optimized and checked for complete protein extraction, resulted in an amount of about 1 µg total protein per *E. cyaneus* individual. All individual lysates were separated by 1D-SDS-PAGE and analyzed by LC-MS/MS. Fourteen of them had sufficient quality and were used in the subsequent analysis resulting in 6947 identified unique peptides in total (minimum 1526, maximum 4530, and mean 3295 peptides per sample). Using an ORF database generated based on the previously obtained *E. cyaneus* transcriptome data (Drozdova et al., 2019b), the MS/MS spectra were assigned to a total of 2563 contigs. These could be grouped to a total of 994 distinct protein groups (*i.e.* sets of proteins united by shared identified peptides) with each protein group including from 1 to 101 contigs of the transcriptome assembly. The

numbers of identified proteins were comparable to the results obtained for related amphipod species (Trapp et al., 2014; Trapp et al., 2015; Gouveia et al., 2019).

Functional annotation with Gene Ontology (GO) terms revealed a nearly comprehensive representation of the predicted proteome by the identified proteins: only 23 out of 10782 Biological Process (BP) terms, 19 out of 2513 Molecular Function (MF) terms, and eight out of 1145 Cellular Component (CC) terms were present in the predicted proteome but absent from the experimentally obtained proteome. Furthermore, only 41, 43, and 9 of the BP, MF, and CC terms, respectively, were present but significantly under-represented. The largest (by the number of annotated transcripts) GO terms being completely absent from the identified proteins were connected to post-translational protein modifications (including histone modifications), transcription, and membrane-related processes (**Fig. 2** and **Table S1**).

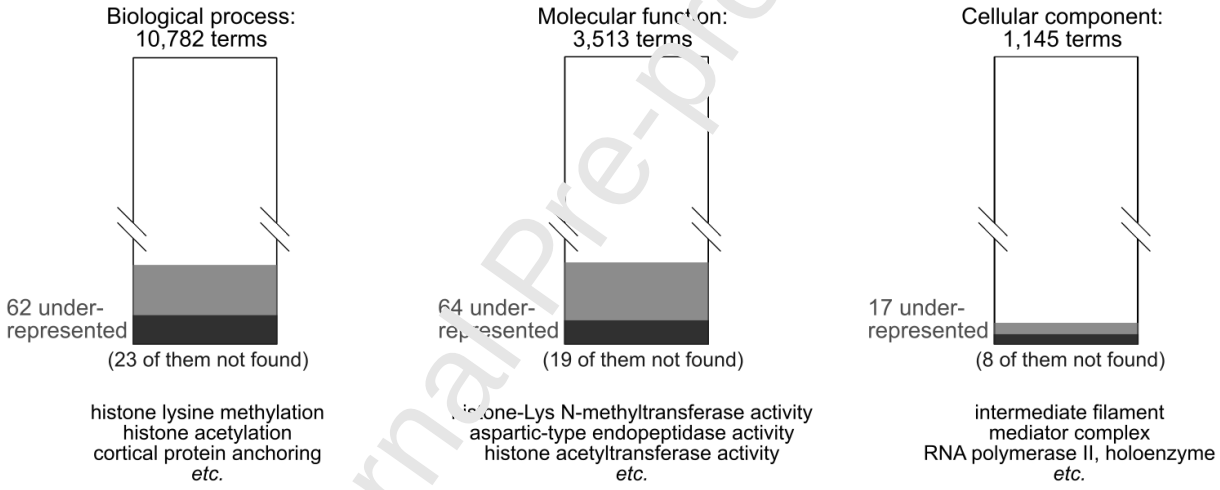


Fig. 2. Relative comparison of the composition of the predicted and obtained *E. cyaneus* core proteomes considering biological process, molecular function, and subcellular localization. The three largest (by the number of annotated transcripts) GO terms absent from the list of identified proteins are listed under each bar. The complete lists of under-represented GO terms are presented in **Table S1**.

3.2. Sex-specific features of the core proteome

A total of 585 protein groups were quantified in male samples from the control conditions. Functional clustering (according to the Panther database) resulted in 14 major protein families, each of which represented at least 1% of the total protein intensity, and they summed up to *ca.* 71% of the overall protein intensity of the sample (**Fig. 3**, upper panel). The three most

abundant protein families were arginine kinase, filamins, and hemocyanins, summing up to 42% of the total protein intensity. Other major proteins were distributed among several functional groups. Cytoskeleton-related proteins were represented by actin, tubulin and a number of actin-binding protein groups. Several of the major protein families identified are involved in carbohydrate metabolism, in particular glycolysis (glyceraldehyde-3-phosphate dehydrogenase and triosephosphate isomerase), gluconeogenesis (fructose-1,6-bisphosphatase), glycogenolysis (glycogen phosphorylase), and cellulose digestion (glycoside hydrolyase) (**Table S2**).

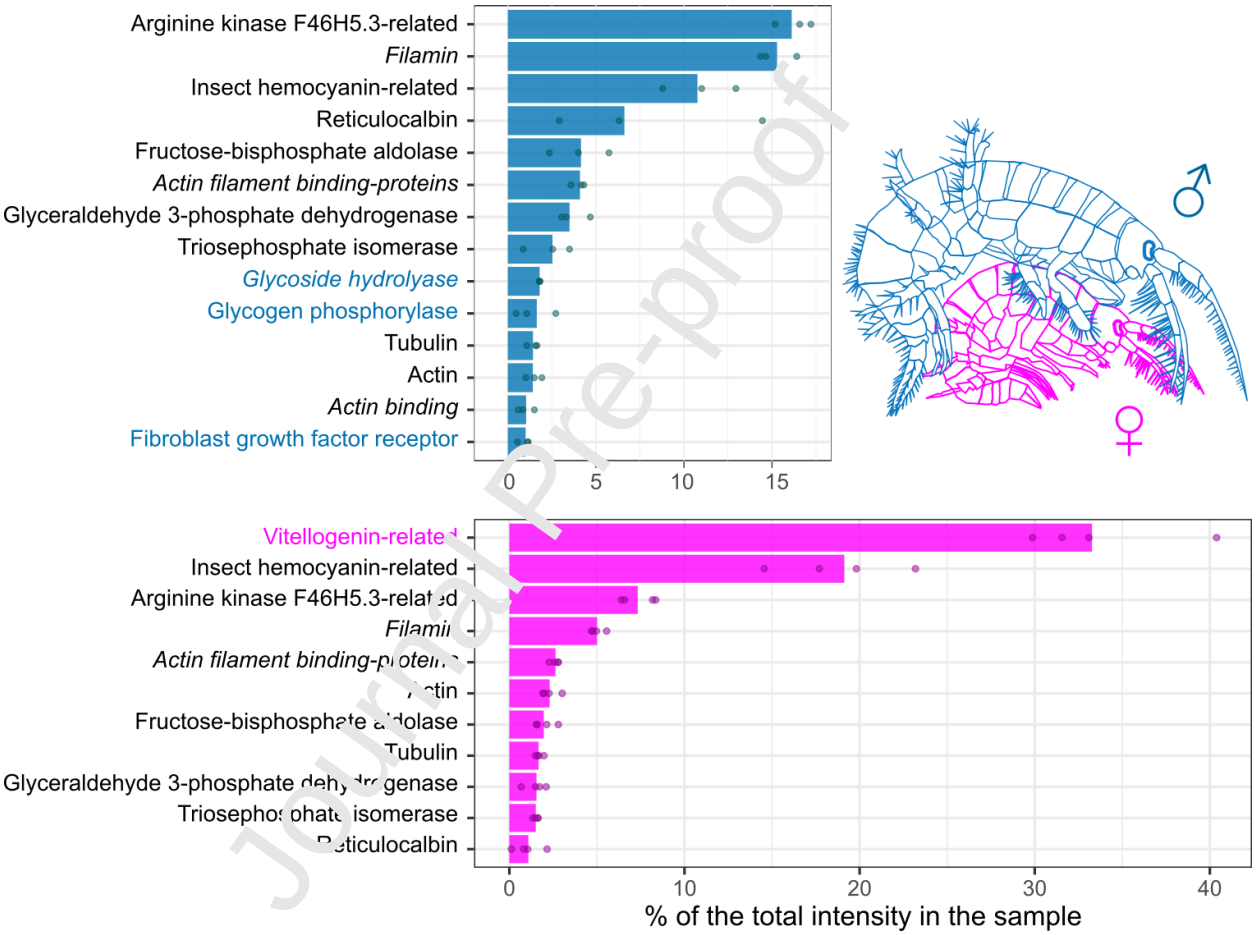


Fig. 3. The most abundant protein families (each group >1% of the total intensity) in *E. cyaneus* male (top) and female (bottom) proteomes from the control conditions. Protein groups with similar annotations mapped to the same Panther family were combined. Names in upright font correspond to the Panther database family or subfamily names; if the family was not named in Panther, we used the name of the most abundant subfamily or the common annotation of the subfamilies (in italics). Proteins with the letters of the names in black were present in high amounts in animals of both sexes, while proteins with the names in blue and pink were abundant in only male or female proteomes, respectively. The complete data on protein abundances are listed in **Table S2**.

In ovigerous female specimens from control conditions, 760 protein groups, about 30% more than in males, were identified. These proteins were assigned to protein families, from which ten comprised at least 1% of the total protein abundance and summed up to *ca.* 72% of the total intensity. As expected, about one-third of total protein were vitellogenins. Excluding vitellogenins, the three most abundant protein families in the female proteome were, like in the male proteome, hemocyanins, arginine kinase, and filamin, summing up to 29% of the total proteome (**Fig. 3**, bottom).

Comparative analysis of the proteomes of males and females from control conditions revealed 160 protein groups exclusively found in at least two female control samples but in no male control sample, while 18 protein groups were found in at least two male control samples but not in any of the female control samples (**Table S2**). At the same time, 176 protein groups were present in each of the seven control samples (both males and females), of which 11 were significantly differentially expressed (**Fig. 4A**; **Table S3**). Overall, this result leaves us with 23 male-specific and 166 female-specific protein groups (**Fig. 4B**).

These exclusively or highly overrepresented protein groups for both sexes were manually assigned to functional categories (**Table S3**). The most diverse functional categories (over ten proteins) among the female-specific proteins being more abundant in female samples were cytoskeletal and muscular, energy, and protein metabolism-related. In addition, such functions as carbohydrate metabolism, xenobiotic and antioxidant response, protein biosynthesis and folding, immune response, molting, and reproduction, were represented by at least five protein groups in female samples (**Fig. 4C**). Twenty-four female-specific protein groups remained without functional annotation (Unclassified).

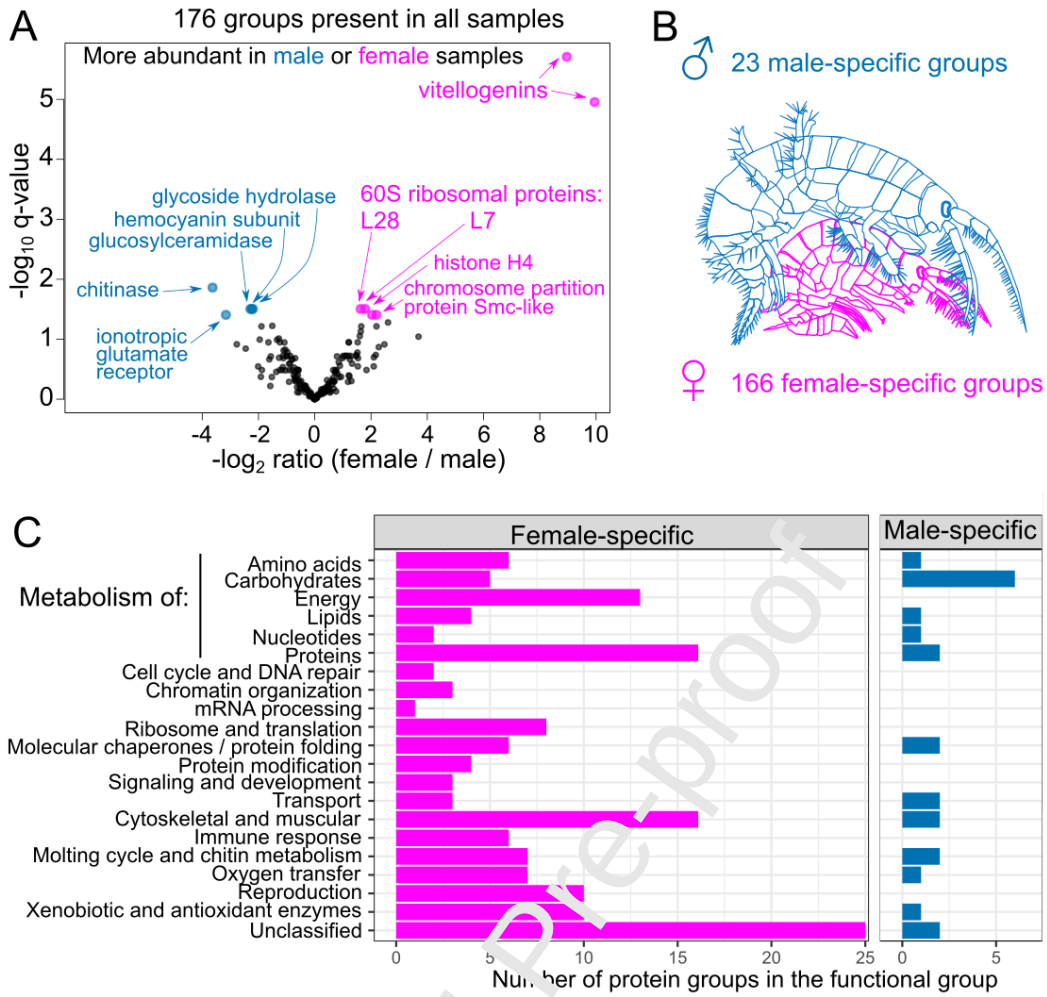


Fig. 4. The diversity of protein groups prevailing in *E. cyaneus* female and male proteomes.

A, differential abundance analysis of proteins present in all samples. Those significantly more up- (red) or down-regulated (blue) in females compared to males are highlighted. B, an overview of the sex-specific protein groups. C, functional diversity of sex-specific proteins. Protein groups with similar annotations were combined.

The 23 male-specific protein groups (**Fig. 4B**) could be associated with fewer functions than the female-specific protein groups (**Fig. 4C**). The only diverse (6 protein groups) male-specific functional group in males included proteins involved in carbohydrate metabolism. Interestingly, we found female- and male-specific hemocyanins (annotated as oxygen transfer proteins) and protein-folding related proteins (**Fig. 4C**).

3.3. Proteomic response to heat shock: females

In total, 319 protein groups were present in all samples and thus could be quantified. Of these proteins, 11 proteins were found to be either downregulated (5 proteins) or upregulated (6

proteins) with unadjusted p-value ≤ 0.05 and absolute $\log_2$ fold change (logFC) ≥ 1. Proteins found to be downregulated were carbonyl reductase 3, staphylococcal nuclease domain-containing protein 1, a notch homolog, calcium/calmodulin-dependent protein kinase type II, and ribosomal protein S2. Among the upregulated proteins were four ribosomal proteins (S18, L7, L13, and L27), a glutathione S transferase 2-like protein, and an unknown protein.

Furthermore, 21 protein groups could be detected exclusively in female samples from the heat shock/recovery treatment group, whereas 23 groups were present only in the control samples (Fig. 5A). The largest functional groups with more than two up-regulated proteins were cytoskeletal and muscular proteins (plastin, titin, and tubulin) and xenobiotic and antioxidant enzymes (two different glutathione S-transferases and a poly(ADP-ribose) glycohydrolase); however, both functional groups also included downregulated members (see below). Furthermore, proteins involved in lipid metabolism (long-chain specific acyl-CoA dehydrogenase, acetyl-CoA acetyltransferase), carbohydrate metabolism (endoglucanase E-4, phosphomannomutase 2), and signaling/development (1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase, cAMP-dependent protein kinase type II regulatory subunit isoform X1) could only be quantified in samples from heat shock/recovery treatment.

Twenty-three protein groups reproducibly disappeared after heat shock and recovery in females. Manual functional annotation revealed a wide diversity of cellular and metabolic functions, some of which also included up-regulated protein groups, such as cytoskeletal and muscular proteins (two protein groups annotated as actins) or xenobiotic and antioxidant enzymes (another glutathione S-transferase). The most frequent function was protein catabolism (chymotrypsin, ovochymase, and an isochorismatase domain-containing protein).

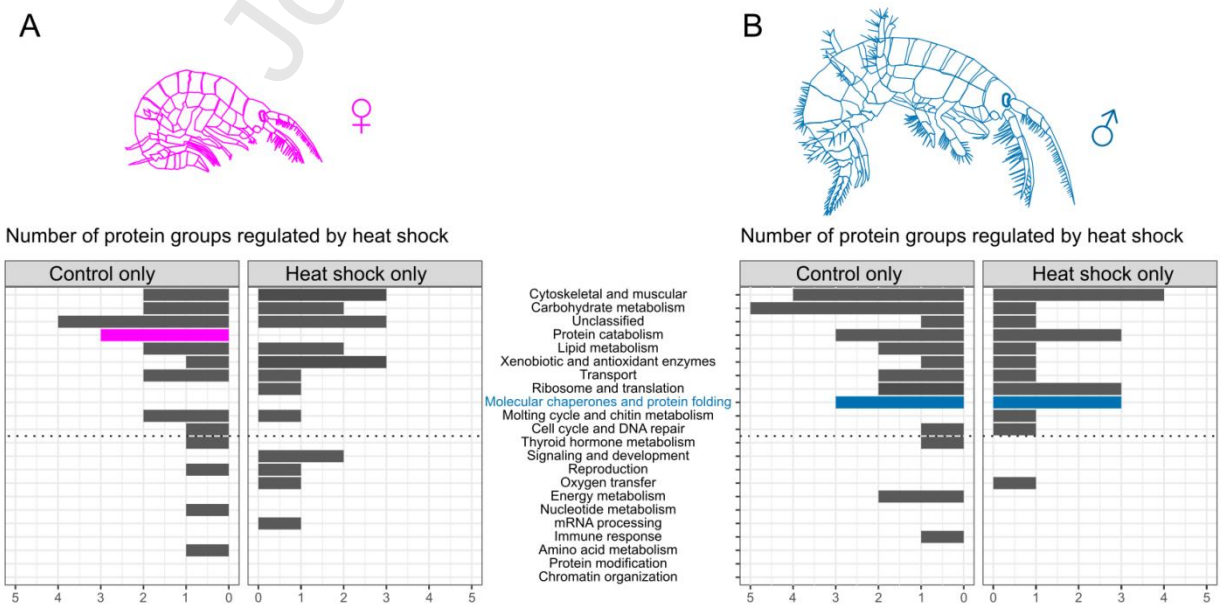


Fig. 5. Overview of the protein groups found in females and males of *E. cyaneus* only from control conditions or only upon heat stress and recovery. The x-axes show the numbers of proteins assigned to each functional group. The functional categories above the dotted line contain at least three protein groups.

3.4. Proteomic response to heat shock: males

For males, 147 proteins could be quantified in all samples, of which 14 were determined to be regulated with unadjusted p-value ≤ 0.05 and absolute logFC ≥1 **(Table S3**). The down-regulated proteins were represented by a myosin regulatory light chain 2, sarcoplasmic calcium-binding protein, catalase, vitellogenin protein precursor or hemolymph clottable-like protein, glycoside hydrolase family 7 protein, and V-type proton ATPase catalytic subunit A. A hemocyanin subunit, nucleoside diphosphate kinase, beta-1,3-glucan-binding protein, peptidyl-prolyl cis-trans isomerase, heat shock 70 kDa protein cognate 5, tubulin alpha, myosin heavy chain, and an unknown protein were more abundant after heat stress and recovery.

Twenty-nine protein groups could be detected in males only after heat shock and recovery but not in the control samples (**Fig. S3**). Manual annotation revealed that the largest functional group comprised cytoskeletal or muscular proteins (tubulin beta, echinoderm microtubule-associated protein, nesprin, and skeletal muscle actin). Other affected groups were ribosome- and translation-related proteins (two ribosomal proteins of the large subunit and a tRNA ligase), molecular chaperones (70 kDa heat shock proteins), and protein metabolism-related proteins (prolyl endopeptidase, N-acetylated-alpha-linked acidic dipeptidase, and 26S proteasome non-ATPase regulatory subunit).

Interestingly, three other protein groups related to protein degradation, two aminopeptidases N and the ubiquitin fusion protein, were found only in the controls but not upon heat shock/recovery. In total, 28 protein groups were absent in the heat shock/recovery treatment samples. The largest functional group was carbohydrate metabolism (sucrase-isomaltase, glucosidase, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, phosphoglucomutase, and protein obstructor), followed by cytoskeletal and muscular proteins (actin, myosin heavy chain, nesprin, and sarcalumenin), protein groups participating in protein catabolic processes (two aminopeptidases and a ubiquitin-fusion protein) and in protein folding (namely, peptidyl-prolyl cis-trans isomerase, T-complex protein, and serine/threonine-protein phosphatase).

4. Discussion

In this study we aimed to investigate (1) the core proteome, (2) the sex-specific differences in the proteomes, and (3) the sex-specific differences in the proteomic responses to heat shock/recovery in the Baikal endemic amphipod species *E. cyaneus*.

4.1. The E. cyaneus *core proteome: LC-MS vs. 2D-PAGE and MALDI-ToF-ToF-MS/MS*

In total, the LC-MS approach used in this work allowed us to identify 994 protein groups, while 2D-PAGE and MALDI-ToF-ToF-MS/MS (2DE-MALDI-MS) previously resulted only in 317 spots, and only half of the 36 spots that were determined to be regulated could be identified (Bedulina et al., 2017). The state of the art LC-MS technique enables separation and spectrum acquisition of thousands of peptides matching to hundreds or even thousands of proteins per sample and thus enables deep proteome analyses, even if such complex protein mixtures as whole animal lysates are used.

In line with our previous data, the most abundant proteins, which formed the largest spots on Coomassie Brilliant Blue (CBB) stained 2D gels in *E. cyaneus*, were hemocyanins and arginine kinase (Bedulina et al., 2017). This has been confirmed with the LC-MS approach, as the highest intensities were observed for these protein groups.

Hemocyanins play an important role in gas exchange in crustaceans. These proteins can also be involved in stress and immune responses (deFur et al., 1990; Mangum et al., 1991; Mangum, 1997; Decker and Foll, 2000; Spicer and Hodgson, 2003). Thirty-four protein groups in total were annotated as hemocyanin alpha subunit, hemocyanin alpha subunit 2, hemocyanin B chain-like, and hemocyanin subunit 1, one of which represented 6.4% of the identified proteome based on the spectra intensities. According to our previous results, gammarids from the genus *Eulimnogammarus* have up to twenty-four different spots on 2D gels which can correspond to different protein species of hemocyanin (Bedulina et al., 2016). Hemocyanin was found to be the most abundant protein in the core proteome of another gammarid, *Gammarus pulex* (Linnaeus, 1758) (Cogne et al., 2019).

Arginine kinase (AK) is responsible for the reversible transfer of the high-energy phosphate from arginine phosphate to ADP to form ATP. This enzyme is a key metabolic enzyme in cellular respiration in crustaceans (Ellington, 2001; Uda et al., 2006). High abundance of this protein was also found in the proteome of the gammarid *G. fossarum* (Trapp et al., 2014). Interestingly, at least two large spots on Coomassie Brilliant Blue stained 2D gels with pI/MW values of 7-8/37-38 kDa were identified as AK (Bedulina et al., 2017); however, with the LC-MS method we found only one transcriptome sequence identified as AK. This sequence contained an open reading frame, deduced amino acid sequence of which included 356 aa residues. Theoretical pI/MW values calculated with IPC (Kozlowski, 2016) were 5.95/40 kDa,

suggesting a slightly more acid pI than the identified spots on CBB stained gels 2D gels. The identified putative protein species were also found to be very abundant with LC-MS analysis (7% of the total intensity in females and 16% in males). This protein probably has various post-translational modifications, altering its molecular weight and isoelectric point. In particular, the MW could be reduced due to cleavage of the initiator methionine, which was suggested for the AK of *Litopenaeus vannamei* (Garcia-Orozco et al., 2007), while the pI value may have changed due to differential phosphorylation especially of the C-terminal arginine/serine repeat sequence, a post-translational modification shown for a crayfish *Orconectes virilis* (Hagen, 1870) under hypoxic conditions (Dawson and Storey, 2011).

Another major group based on their intensities were cytoskeletal and muscle-related proteins. The most abundant protein, identified as "filamin-A isoform X3", represented 3.36% of the intensity of the identified proteome and 34% of all protein with cytoskeletal and muscular function. Filamin is a large (280 kDa) cytoplasmic protein, which has both structural and regulatory functions. It binds actin and is involved in cell motility and signaling. It was shown as a major protein in many crustacean species: i.e., the porcelain crab *Petrolisthes cinctipes* (Garland et al., 2015), red swamp crawfish *Procambarus clarki* (Yang et al., 2017), and spiny lobster *Sagmariasus verreauxi* (Ventura et al., 2015).

Among other cytoskeletal and muscular-related proteins, the most abundant in *E. cyaneus* proteomes were various protein species of actin (alpha-, beta-, skeletal muscle actin *etc.*), tubulin (alpha- and beta-chains), muscle LIM protein, nesprin (nuclear envelope spectrin-repeat protein), and actinin. These proteins are essential structural and regulatory proteins in the majority of animals. A high intensity (0.55%) within the characterized proteome was found for the contig identified as "sarcoplasmic calcium-binding protein". This protein was found as a major protein only in invertebrates, predominantly in muscle and neurons; it plays an essential role in calcium homeostasis and muscle contraction (Hermann and Cox, 1995).

Several proteins involved in energy metabolism are abundant in the proteomes of the studied species, such as three enzymes involved in glycolysis, fructose-bisphosphate aldolase (1.3%), triosephosphate isomerase (1.08%), and glyceraldehyde-3-phosphate dehydrogenase (0.8%).

According to the obtained intensities, an essential part of the core proteomes was heat shock 70 kDa protein. Fifteen protein groups annotated as heat shock protein 70 were found in the *E. cyaneus* proteome, with one of them comprising 0.44% intensities of the whole proteome. Heat shock proteins 70 (Hsp70) are important cellular proteins with molecular chaperone function, involved in protein folding, transport, degradations, and many signal transduction

pathways controlling cell homeostasis, proliferation, differentiation, and cell death (Rosenzweig et al., 2019).

Besides proteins with predicted functions, we found a highly abundant protein with the best hit to the *H. azteca* sequence "PREDICTED: uncharacterized protein LOC108673699" (1.06% of the total proteome). In a parallel study, we found that this protein is a novel putative carotenoid-binding protein defining the characteristic blue color of *E. cyaneus* (Drozdova et al., 2020).

As LC-MS data also provided us with a global view of the proteome, we compared the obtained proteome with the predicted transcriptome-based sequences (**Fig. 2**). We found a good representation of the functional groups of the predicted sequences, except for chromatin modification-, transcription- and membrane-related proteins (**Table S1, Fig. 2**). The lack of these proteins could be explained by their generally low abundance, which is probably true for many transcription-related proteins, or low extraction efficiency, which could be the case for membrane-related proteins and should be taken into account when adjusting protein extraction methods.

4.2. Sexual dimorphism of E. cyaneus proteomes

In the proteomes of the laboratory controls, we found more female-specific than male-specific protein groups. This result is congruent with the previous finding of more female-specific spots on 2D gels for *E. cyaneus* females (Bedulina et al., 2017). Earlier, we hypothesized that females have higher metabolic activity than males (Bedulina et al., 2017). Indeed, the functional groups to which female-specific proteins were assigned included energy metabolism and protein synthesis, folding, and metabolism. A higher amount of hemocyanins in the female proteomes also may indicate an enhanced metabolic rate in these animals. The higher metabolic activity in ovigerous females is probably related to egg development. The largest functional groups in the proteomes of *E. cyaneus* females included, besides the basic cellular metabolic functions, molecular chaperones (those annotated as 78 kDa glucose-regulated protein, several proteins identified as heat shock protein 70 kDa, hsc70-interacting protein, heat shock protein 90, and protein disulfide-isomerase A6 homolog) and antioxidant enzymes (identified as catalase, glutathione S-transferase, mitochondrial manganese superoxide dismutase and thioredoxin reductase).

At the same time, most of the protein groups in proteomes of *E. cyaneus* males were assigned to carbohydrate metabolism. Among molecular chaperones found in the male proteomes from the controls, there were no inducible heat shock proteins but chaperones with housekeeping functions (T-complex protein and peptidyl-prolyl cis-trans isomerase). These

differences between the sexes may reflect a higher level of cellular stress in females, which is supported by the previous finding of the higher sensitivity of females to increased temperature. Borges et al. (2018) demonstrated a higher sensitivity of reproductive function to hypercapnia in females of the amphipod *Gammarus locusta*. Cellular stress in females is underlined by higher tissue levels of Hsp70 determined with western blotting (Bedulina et al., 2017). In turn, it might be connected to the overall higher mRNA and protein synthesis and modification in females. This supports the data about total protein concentration, which indicated higher levels of total protein in females (**Table S4**).

4.3. Proteomic response to heat shock

Our previous results demonstrated clear differences in tolerance to acute thermal stress between males and ovigerous females of two Baikal endemic amphipods, *E. cyaneus* and *E. verrucosus* (Bedulina et al., 2017). Females of *E. cyaneus* were significantly more sensitive to acute thermal stress than males with the estimated median lethal time being 13.6 hours for females compared to 35 hours for males. The first attempt to investigate the reason for these differences and to characterize the proteomic stress-response using 2D-PAGE and MALDI-MS/MS revealed a significant difference between sexes in gel patterns for samples from before and after heat shock and recovery. However, the proteome coverage using this approach was quite limited. Complementarily, we estimated the level of Hsp70 with western blotting and found slightly higher levels of Hsp70 in females of this species but the amount of Hsp70 remained stable in both males and females after 1 hour of heat shock and recovery.

To receive more detailed insights into the proteomic stress response in this important species, we here used the LC-MS based approach.

In both sexes of *E. cyaneus*, heat shock and recovery led to the accumulation of cytoskeletal and muscular, ribosomal-related proteins, enzymes of lipid metabolism, hemocyanins, and proteins with the antioxidant function.

Previous results, obtained with 2D-PAGE and MALDI-ToF-MS/MS, revealed significant differences in gel patterns in males and females of *E. cyaneus* after heat shock (25.5 °C, 1 hour) and recovery (6 °C, 3 hours) (Bedulina et al., 2017). Our LC-MS-based investigation confirmed these results, demonstrating different sets of up- and down-regulated proteins in males and females of this species. The common functional groups up-regulated in both sexes were cytoskeletal and muscular proteins, ribosomal proteins, lipid metabolism-related proteins, and hemocyanins. Among the down-regulated functional groups, antioxidant, lipid metabolism, carbohydrate metabolism, cell cycle, protein catabolism-related, cytoskeletal and muscular, ribosomal and transport proteins were common for both sexes. Up- and down-regulation of

proteins from the same functional group can indicate a switch in these processes in cells triggered by the presented stimulus. This was observed for cytoskeletal and muscular, ribosomal and lipid metabolism-related proteins.

The involvement of cytoskeletal and muscular proteins in heat shock response in crustaceans (a crab *P. cinctipes*) was shown earlier by Garland et al. (2015). It can be a sign of heat shock-induced actin filament restructuring and stabilization (Lee and Dominguez, 2010).

Interestingly, a slight (2.7-fold) upregulation of molecular chaperones was found only in the males of *E. cyaneus*. Additionally, three of the contigs identified as Hsp70 or heat shock protein cognate were found only after heat shock and recovery. Besides, another protein involved in protein folding, peptidyl-prolyl cis-trans isomerase, was slightly (2.8-fold) upregulated in males after heat shock and recovery.

Previously published results indicated the absence of significant Hsp70 upregulation with western blotting in either sex after the same experimental treatment (Bedulina et al., 2017). Given the higher sensitivity of LC-MS, however, Hsp70 upregulation may have been below the detection limit of western blotting. Females might have constitutive Hsp70 levels in their cells that are sufficiently high for chaperone function so that further Hsp70 protein synthesis is either not required or it not possible due to the lack of the necessary energy. Instead, larger numbers of antioxidant proteins are involved in the heat shock response in females, such as glutathione S-transferase and dual oxidase. Similarly, Grilo et al. (2018) demonstrated stimulation of Hsp70 production after heat shock in males of topshells (*Trochus histrio*) that was more pronounced compared to females that reacted by activation of the antioxidative enzyme catalase.

5. Conclusion

The applied proteogenomic approach significantly deepened our understanding of the proteome of the endemic amphipod *Eulimnogammarus cyaneus* from Lake Baikal and allowed a comprehensive investigation of both qualitative and quantitative features of this non-model organism. The proteome responses of both sexes of *E. cyaneus* comprised differential expression of hemocyanins, cytoskeletal and muscular proteins, antioxidant proteins, ribosomal proteins and transport proteins as well as proteins being involved in lipid metabolic processes, carbohydrate metabolism, cell cycle, protein catabolism. Differences in the proteomic responses to acute thermal stress impact and recovery suggest an explanation for the differences between the sexes in the whole organism responses on a molecular level: the higher sensitivity of females to stressor impact may be associated with an increased general level of cellular stress as a major proportion of the organismal resources is diverted to the processes related to egg development.

6. Acknowledgments

We are grateful to the team of the "Biosystems' adaptation" laboratory for field and laboratory assistance; in addition, we would like to specifically thank Ekaterina Shchapova for the drawing of *E. cyaneus* in amplexus and Polina Lipaeva for fruitful discussions.

The study was supported by the Russian Science Foundation/Helmholtz Association of German Research Centres (RSF grant number 18-44-06201).

7. References

1. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H., Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; Harris, M.A. Gene ontology: tool for the unification of biology. *Nat. Genet.* **2000**, *25*(1), 25.
2. Bazikalova, A.Y. [Materials on reproduction of Baikal amphipods] (In Russian). *Doklady Akad Nauk SSSR* **1941**, *3*, 407-426.
3. Bedulina, D.; Meyer, M.F.; Gurkov, A.; Kondratjeva, E.; Baduev, B.; Gusdorf, R.; Timofeyev, M.A. Intersexual differences of heat shock response between two amphipods (*Eulimnogammarus verrucosus* and *Eulimnogammarus cyaneus*) in Lake Baikal. *PeerJ* **2017**, *5*, p.e2864.
4. Bedulina, D.S.; Evgen'ev, M.B.; Timofeyev, M.A.; Protopopova, M.V.; Garbuz, D.G.; Pavlichenko, V.V.; Luckenbach, T.; Shatilina, Z.M.; Axenov- Gribanov, D.V.; Gurkov, A.N. and Sokolova, I.M. Expression patterns and organization of the *hsp70* genes correlate with thermotolerance in two congener endemic amphipod species (*Eulimnogammarus cyaneus* and *E. verrucosus*) from Lake Baikal. Mol. Ecol. **2013**, *22*(5), 1416-1430.
5. Bedulina, D.S.; Gurkov, A.N.; Baduev, B.K.; Borvinskaya, E.V.; Dimova, M.D.; Timofeyev, M.A. Preliminary analysis of hemocyanins in hemolymph plasma of Baikal endemic amphipods. *J. Stress Physiol. Biochem.* **2016**, *12*(1), 75-86.
6. Borges, F. O.; Figueiredo, C.; Sampaio, E.; Rosa, R.; Grilo, T. F. Transgenerational deleterious effects of ocean acidification on the reproductive success of a keystone crustacean (Gammarus locusta). *Mar. Environ. Res.* **2018**, 138, 55-64.

7. Chen, T.W.; Gan, R.C.; Fang, Y.K.; Chien, K.Y.; Liao, W.C.; Chen, C.C.; Wu, T.H.; Chang, I.Y.F.; Yang, C.; Huang, P.J.; Yeh, Y.M. FunctionAnnotator, a versatile and efficient web tool for non-model organism annotation. *Sci. Rep.* **2017**, *7*(1), 10430.

8. Cogne, Y.; Almunia, C.; Gouveia, D.; Pible, O.; François, A.; Degli-Esposti, D.; Geffard, O.; Armengaud, J.; Chaumot, A. Comparative proteomics in the wild: Accounting for intrapopulation variability improves describing proteome response in a *Gammarus pulex* field population exposed to cadmium. *Aquat. Toxicol.* **2019**, *214*, 105244.

9. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, 21(18), 3674-3676.

10. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*(12), 1367.

11. Cristescu, M. E.; Adamowicz, S. J.; Vaillant, J. J.; Haffner, D. G. Ancient lakes revisited: from the ecology to the genetics of speciation. *Molecular Ecology* **2010**, *19*(22), 4837-4851.

12. Dawson, N. J. and Storey, K. B. Regulation of tail muscle arginine kinase by reversible phosphorylation in an anoxia-tolerant crayfish. *J. Comp. Physiol. B* **2011**, *181*(7), 851-859.

13. Decker, H.; Föll, R. Temperature adaptation influences the aggregation state of hemocyanin from *Astacus leptodactylus*. *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **2000**, *127*(2), 147-154.

14. DeFur, P.L.; Mangum, C.P.; Reese, J.E. Respiratory responses of the blue crab *Callinectes sapidus* to long-term hypoxia. The Biological Bulletin **1990**, *178*(1), 46-54.

15. Drozdova, P.; Bedulina, D.; Madyarova, E.; Rivarola-Duarte, L.; Schreiber, S.; Stadler, P.F.; Luckenbach, T. and Timofeyev, M. Description of strongly heat-inducible *heat shock protein 70 transcripts* from Baikal endemic amphipods. *Sci. Rep.* **2019a**, *9*(1), 8907.

16. Drozdova, P.; Rivarola-Duarte, L.; Bedulina, D.; Axenov-Gribanov, D.; Schreiber, S.; Gurkov, A.; Shatilina, Z.; Vereshchagina, K.; Lubyaga, Y.; Madyarova, E.; Otto, C.; Jühling, F.; Busch, W.; Jakob, L.; Lucassen, M.; Sartoris, F.J.; Hackermüller, J.; Hoffmann, S.; Pörtner, H-O.; Luckenbach, T.; Timofeyev, M.; Stadler, P.F. Comparison between transcriptomic responses to short-term stress

exposures of a common Holarctic and endemic Lake Baikal amphipods. *BMC Genomics* **2019b**, *20*(1), 712.

17. Drozdova, P.; Saranchina, A.; Morgunova, M.; Kizenko, A.; Lubyaga, Y.; Baduev, B.; Timofeyev M. The level of putative carotenoid-binding proteins determines the body color in two species of endemic Lake Baikal amphipods. *PeerJ.* **2020,** 8, e9387.
18. Ellington, W.R. Evolution and physiological roles of phosphagen systems. *Annu. Rev. Physiol.* **2001**, *63*(1), 289-325.
19. Falkner, K.K.; Measures, C.I.; Herbelin, S.E.; Edmond, J.M.; Weiss, R.F. The major and minor element geochemistry of Lake Baikal. *Limnol. and Oceanogr.* **1991**, *36*(3), 413-423.
20. Garcia-Orozco, K. D.; Aispuro-Hernandez, E.; Yepiz-Plascencia, G.; Calderon-de-la-Barca, A. M.; Sotelo-Mundo, R. R., 2007. Molecular characterization of arginine kinase, an allergen from the shrimp *Litopenaeus vannamei*. *International archives of allergy and immunology* **2007**, *144*(1), 23-28.
21. Garland, M.A.; Stillman, J.H.; Tomanek, L. The proteomic response of cheliped myofibril tissue in the eurythermal porcelain crab *Petrolisthes cinctipes* to heat shock following acclimation to daily temperature fluctuations. J. Exp. Biol. **2015**, *218*(3), 388-403.
22. Gavrilov, G. B. [On the problem of the time of reproduction in amphipods and isopods in Lake Baikal] (In Russian). *Doklady Akad Nauk SSSR* **1949**, *64*(5), 739-742.
23. Gilbert, D.G. Gene of the pig, *Sus scrofa*, reconstructed with EvidentialGene. *PeerJ* **2019**, *7*, e6374.
24. Gouveia, D.; Cogne, Y.; Gaillard, J.C.; Almunia, C.; Pible, O.; François, A.; Degli-Esposti, D.; Geffard, O.; Chaumot, A.; Armengaud, J., 2019. Shotgun proteomics datasets acquired on *Gammarus pulex* animals sampled from the wild. *Data in brief* **2019,** 27, 104650.
25. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; Chen, Z. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **2011**, *29*(7), 644.

26. Grilo, T.F; Lopes, A.R.; Sampaio, E.; Rosa, R.; Cardoso, P.G. Sex differences in oxidative stress responses of tropical topshells (*Trochus histrio*) to increased temperature and high pCO2. Mar. Pollut. Bull. **2018** 131: 252-259. DOI: 10.1016/j.marpolbul.2018.04.031

27. Hermann, A.; Cox, J.A. Sarcoplasmic calcium-binding protein. *Comp. Biochem. Physiol., Part B*: *Biochem. Mol. Biol.* **1995**, *111*(3), 337-345.

28. Jakob, L.; Axenov- Gribanov, D.V.; Gurkov, A.N.; Ginzburg, M.; Bedulina, D.S.; Timofeyev, M.A.; Luckenbach, T.; Lucassen, M.; Sartoris, F.J.; Pörtner, H.O. Lake Baikal amphipods under climate change: thermal constraints and ecological consequences. *Ecosphere* **2016**, *7*(3), e01308.

29. Kozlowski, L.P. IPC–isoelectric point calculator. *Biology direct* **2016**, *11*(1), 55.

30. Krueger, F. Babraham Bioinformatics - Trim Galore! https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 2017-08-31

31. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*(5259), 680.

32. Langmead, B.; Salzberg, S. L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. methods* **2012**, *9*(4), 357.

33. Lee, S. H.; Dominguez, R. Regulation of actin cytoskeleton dynamics in cells. *Molecules and cells* **2010**, *29*(4), 311-325.

34. Mangum, C.P. Adaptation of the oxygen transport system to hypoxia in the blue crab, *Callinectes sapidus*. *Am. Zool.* **1997**, *37*(6), 604-611.

35. Mangum, C.P.; Graves, J.; Rainer, J.S. Oligomer composition and oxygen binding of the hemocyanin of the blue crab *Callinectes sapidus*. The Biological bulletin **1991**, *181*(3), 453-458.

36. Mats, V.D.; Shcherbakov, D.Y.; Efimova, I.M. Late Cretaceous-Cenozoic history of the Lake Baikal depression and formation of its unique biodiversity. *Stratigr. Geol. Correl.***2011**, *19*(4), 404-423.

37. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **2018**, *47*(D1), D419-D426.

38. Moore, M.V.; Hampton, S.E.; Izmest'Eva, L.R.; Silow, E.A.; Peshkova, E.V. and Pavlov, B.K. Climate change and the world's "sacred sea"—Lake Baikal, Siberia. *BioScience* **2009**, *59*(5), 405-417.

39. Müller, S.A.; Findeiß, S.; Pernitzsch, S.R.; Wissenbach, D.K.; Stadler, P.F.; Hofacker, I.L.; von Bergen, M.; Kalkhof, S. Identification of new protein coding sequences and signal peptidase cleavage sites of *Helicobacter pylori* strain 26695 by proteogenomics. *J. proteomics* **2013**, *86*, 27-42.

40. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D.J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M. and Pérez, E.; 2018. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.*, *47*(D1), D442-D450.

41. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: **2017**.

42. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*(7), e47.

43. Rosenzweig, R.; Nillegoda, N.B., Mayer, M.P.; Bukau, B. The Hsp70 chaperone network. *Nat. Rev. Mol. Cell. Biol.* 2019, 665-680.

44. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, 31(19), 3210-3212.

45. Song, L.; Florea, L. Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience* **2015**, *4*(1), p.48.

46. Spicer, J.I.; Hodgson, E. Structural basis for salinity-induced alteration in oxygen binding by haemocyanin from the estuarine amphipod *Chaetogammarus marinus (L.)*. *Physiol. Biochem. Zool.* **2003**, 76(6), 843-849.

47. Storey, J.D.; Bass, A.J.; Dabney, A.; Robinson, D. qvalue: Q-value estimation for false discovery rate control. R package version 2.16.0, **2019**, http://github.com/jdstorey/qvalue.

48. Takhteev, V. V. Trends in the evolution of Baikal amphipods and evolutionary parallels with some marine malacostracan faunas. *Adv. Ecol. Res.* **2000**, 31, 197-220.

49. Takhteev, V. V.; Berezina N. A.; Sidorov D. A. Checklist of the Amphipoda (Crustacea) from continental waters of Russia, with data on alien species. *Arthropoda Selecta* **2015**, *24*(3), 335-370.

50. Timofeyev, M.A. [*Ecological and physiological aspects of adaptation to abiotic environmental factors in endemic Baikal and Palearctic amphipods*] (In Russian); Tomsk: Tomsk State University, 2010.

51. Timofeyev, M.A.; Kirichenko, K.A. [Experimental evaluation of the role of abiotic factors in the restriction of distribution of endemics beyond Lake Baikal by the example of amphipods] (In Russian). *Sibirskii Ekologicheskii Zhurnal* **2004**, *11*(1), 41.

52. Timoshkin, O. A. *Lake Baikal: diversity of fauna, problems of its immiscibility and origin, ecology and "exotic" communities. Index of animal species inhabiting Lake Baikal and its catchment area*; Nauka: Novosibirsk, 2001.

53. Timoshkin, O. A.; Moore, M. V.; Kulikova, N. N.; Tomberg, I. V.; Malnik, V. V.; Shimaraev, M. N.; Troitskaya, E. S.; Shirokaya, A. A.; Sinyukovich, V. N.; Zaitseva, E. P.; Domysheva, V. M. Groundwater contamination by sewage causes benthic algal outbreaks in the littoral zone of Lake Baikal (East Siberia). *J. Great Lakes Res.* **2018**, *44*(2), 230-244.

54. Trapp, J.; Almunia, C.; Gaillard, J.C.; Pible, O.; Chaumot, A.; Geffard, O. and Armengaud, J. Data for comparative proteomics of ovaries from five non-model, crustacean amphipods. *Data in brief* **2015** *5*, 1-6.

55. Trapp, J.; Almunia, C.; Gaillard, J.C.; Pible, O.; Chaumot, A.; Geffard, O. and Armengaud, J. Proteogenomic insights into the core-proteome of female reproductive tissues from crustacean amphipods. *J. Proteomics* **2016**, *135*, 51-61.

56. Trapp, J.; Geffard, O.; Imbert, G.; Gaillard, J.C.; Davin, A.H.; Chaumot, A. and Armengaud, J.; **2014**. Proteogenomics of *Gammarus fossarum* to document the reproductive system of amphipods. *Molecular & Cellular Proteomics*, *13*(12), 3612-3625.

57. Uda, K.; Fujimoto, N.; Akiyama, Y.; Mizuta, K.; Tanaka, K.; Ellington, W.R.; Suzuki, T. Evolution of the arginine kinase gene family. *Comp. Biochem. Physiol., Part D: Genomics Proteomics* **2006**, *1*(2), 209-218.

58. Ventura, T.; Fitzgibbon, Q.P.; Battaglene, S.C.; Elizur, A. Redefining metamorphosis in spiny lobsters: molecular analysis of the phyllosoma to puerulus transition in *Sagmariasus verreauxi*. *Sci. Rep.* **2015**, *5*, 13537.

59. Vizcaíno, J.A.; Côté, R.G.; Csordas, A.; Dianes, J.A.; Fabregat, A.; Foster, J.M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; O'Kelly, G., 2012. The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **2012**, *41*(D1), D1063-D1069.

60. Weiss, R.F.; Carmack, E.C.; Koropalov, V.M. Deep-water renewal and biological production in Lake Baikal. *Nature* **1991**, *349*(6311), 665.

61. Yang, Y.; Zhang, Y.X.; Liu, M.; Maleki, S.J.; Zhang, M.L.; Liu, Q.M.; Cao, M.J.; Su, W.J.; Liu, G.M. Triosephosphate isomerase and filamin C share common epitopes as novel allergens of *Procambarus clarkii*. *J. Agric. Food Chem.* **2017**, *65*(4), 950-963.

Supplementary Information

Supplementary Table S1: Absent or significantly ($p < 0.01$) underrepresented gene ontology terms based on the identified proteins in respect to (a) biological processes (b), molecular functions and (c) cellular compartments

Supplementary Table S2: Full list of the label-free protein quantification results for control and heat stress conditions of male and female individuals

Supplementary Table S3: Overview of the proteins that were either solely present or significantly enriched in a sex or heat stress specific manner.

subtable 1: "Sex specific absence"
List of proteins being quantified at least in two replicates of one sex but not detected for the other sex

subtable 2: "Female/male significant"
List of proteins being differentially expressed (log2 FC > 1, p-value < 0.05) comparing male vs. female at control conditions

subtable 3: "Male heat shock absence"
List of proteins being quantified at least in two replicate of males at a given treatment (heat stress or control) but are not detected at the other treatment

subtable 4: "Male heat shock regulated"
List of proteins being differentially expressed (log2 FC > 1, p-value (unadjusted) < 0.05) comparing male treated with heat compared to control condition

subtable 5: "female heat shock absence"
List of proteins being quantified at least in two replicate of females

	at a given treatment (heat stress or control) but are not detected at the other treatment
subtable 6:	"female heat shock regulated" List of proteins being differentially expressed (log2 FC > 1, p-value (unadjusted) < 0.01) comparing females treated with heat compared to control condition
Supplementary Table S4:	Total protein concentrations in the extracts of female and male samples

Declaration of interests

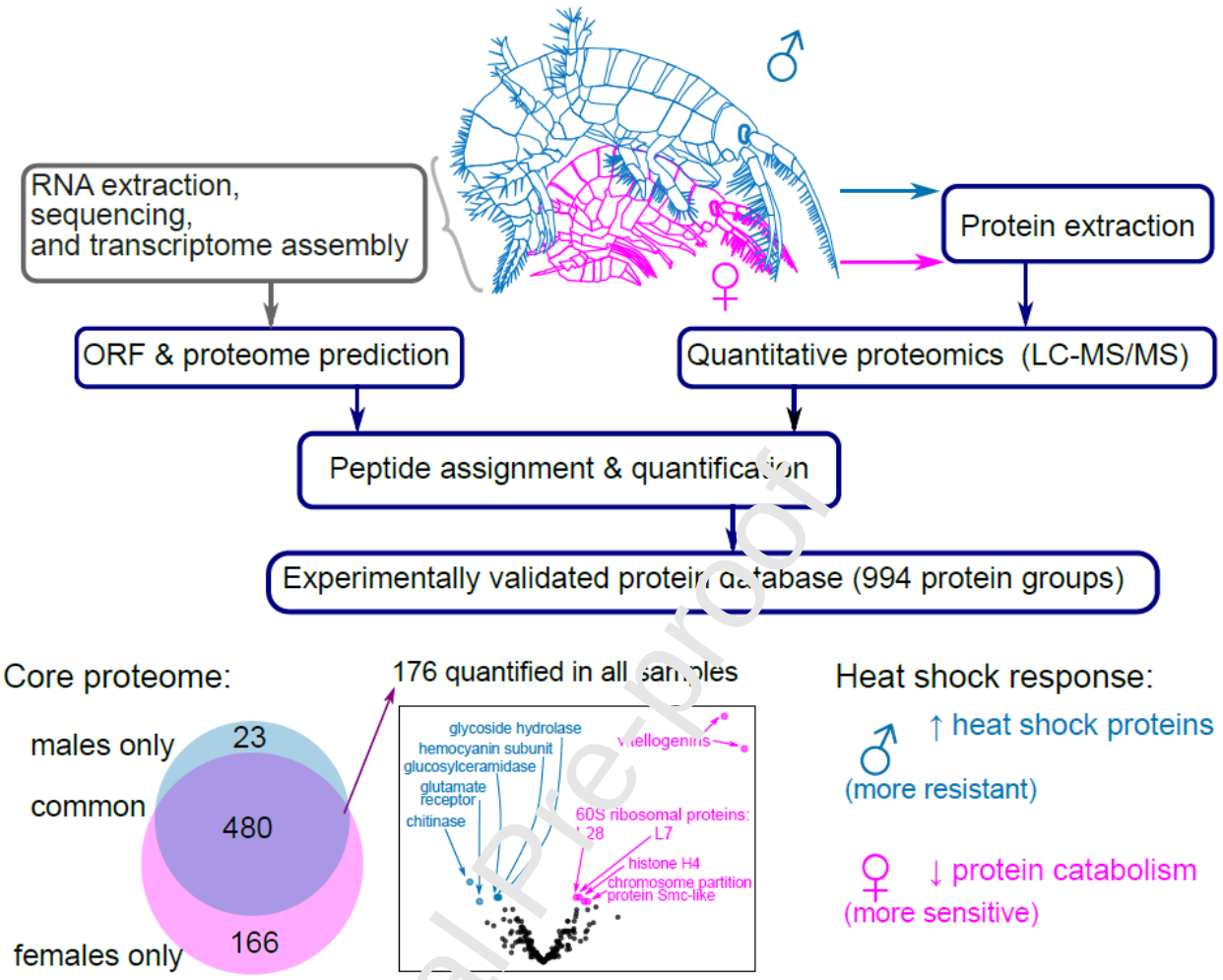
☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

CRediT author statement

Daria Bedulina: Conceptualization, Investigation, Formal analysis, Data curation, Original draft preparation; **Polina Drozdova**: Methodology, Software, Data Curation, Validation, Writing - Review & Editing, Visualization; **Anton Gurkov** Formal analysis, Investigation, Validation; **Martin von Bergen** Resources, Writing - Review & Editing; **Peter F. Stadler:** Data Curation, Writing - Review & Editing; **Till Luckenbach:** Conceptualization, Writing - Review & Editing; **Maxim Timofeyev:** Resources, Project administration; **Stefan Kalkhof**: Resources, Methodology, Writing - Review & Editing, Formal analysis, Data Curation, Validation

Graphical abstract



Highlights

- Endemic amphipods are extremely important in benthic communities of Lake Baikal.
- We studied the proteomes of males and females of *Eulimnogammarus cyaneus*.
- High-throughput mass spectrometry allowed us to find almost 1000 protein groups.
- Females possess higher levels of heat shock proteins in control conditions.
- Males activate heat shock proteins in heat stress conditions, while females do not.