Heredity (2021) 126:424–441
https://doi.org/10.1038/s41437-020-00382-w

the geneticssociety

ARTICLE

Evolution of chaperome gene expression and regulatory elements in the antarctic notothenioid fishes

Kevin T. Bilyk[1] · Xuan Zhuang[2] · Luis Vargas-Chacoff[3] · C-H Christina Cheng[4]

Received: 21 September 2019 / Revised: 15 October 2020 / Accepted: 19 October 2020 / Published online: 5 November 2020

Abstract

Confined within the cold-stable Southern Ocean, Antarctic notothenioid fishes have undergone an evolutionary loss of the inducible heat shock response (HSR), while facing perpetual low-temperature challenges to cellular proteostasis. This study examines how evolution in chronic cold has affected the shared cellular apparatus that mediates proteostasis under normal and heat stressed states. To deduce Antarctic-specific changes, we compared native expression levels across the full suite of chaperome genes and assessed the structural integrity of two crucial HSR regulators – Heat Shock Factor 1 (HSF1) that activates HSR, and heat shock elements (HSEs), the binding sites for HSF1 – between Antarctic fishes and the basal temperate notothenioid *Eleginops maclovinus*. Native expression levels of Antarctic fish chaperomes showed very modest changes overall, contrary to the common view of constitutive upregulation in the cold. Only a few cytosolic HSP70 genes showed greater transcription, with only the ancestrally-inducible *HSPA6* strongly upregulated across all Antarctic species. Additionally, the constant cold has apparently not relaxed the selective pressures on maintaining HSF1 and HSEs in Antarctic fish. Instead, we found HSF1 experienced intensified selective pressure, with conserved sequence changes in Antarctic species suggesting optimization for non-heat-stress functional roles. HSEs of the HSP70 gene family have largely remained conserved in canonical sequence motifs and copy numbers as in *E. maclovinus*, showing limited impact of relaxed selective pressure. This study shows that evolution in chronic cold has led to both subtle and distinctive changes in the cellular apparatus for proteostasis and HSR, with functional consequences amenable to experimental evaluation.

Introduction

The chronically icy, freezing waters of the Southern Ocean (SO) (Barnes et al. 2006; Cziko et al. 2014; Hunt et al. 2003) impose strong selective pressures on the endemic Antarctic notothenioid fishes, leading to remarkable trait changes. The evolution of the novel antifreeze glycoproteins that prevent death from freezing exemplifies one such adaptive trait gain (DeVries and Cheng 2005). At the other end of the spectrum, trait reduction under the constant cold manifests itself in a greatly diminished organismal heat tolerance (Bilyk and DeVries 2011), with high-latitude species unable to survive temperatures above 5–7 °C (Somero and DeVries 1967) or to mount a cellular heat shock response (HSR) (Hofmann et al. 2000). Evolution in perpetual extreme cold has clearly influenced notothenioid cellular and transcriptional processes (Beers and Jayasundara 2015), which expectedly would include the genes of the chaperome as well as the regulatory machinery that normally controls the response of these genes to heat stress.

Associate editor Bastiaan Star

Supplementary information The online version of this article (https://doi.org/10.1038/s41437-020-00382-w) contains supplementary material, which is available to authorized users.

✉ Kevin T. Bilyk
bilykk@montclair.edu

✉ C-H Christina Cheng
c-cheng@illinois.edu

[1] Department of Biology, Montclair State University, 1 Normal Ave., Montclair, NJ 07043, USA

[2] Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA

[3] Instituto de Ciencias Marinas y Limnológicas, Laboratorio de Fisiología de Peces, Centro Fondap de Investigación de Altas Latitudes (IDEAL), Universidad Austral de Chile, Valdivia, Chile

[4] Department of Evolution, Ecology, and Behavior, University of Illinois, Urbana-Champaign, IL 61801, USA

The array of cellular molecular chaperones and co-chaperones that collectively support protein homeostasis or proteostasis comprise the chaperome. The members of the chaperome are key effectors and critical links across the proteostasis network's functional branches that engage in nascent protein synthesis, protein folding, and protein degradation (Brehme et al. 2014; Klaips et al. 2018; Schmidt and Finley 2014). As effectors of cellular health, the activity of the chaperome contributes to delineating the functionally viable temperature range of organisms. The freezing Antarctic water temperatures are thought to present challenges to proper protein folding, thereby imposing selective pressure on the proteostasis network (Peck 2016). One outcome is the apparent co-option of an inducible heat shock protein HSP70 into constitutive expression in the Antarctic notothenioids, which was hypothesized to mitigate the challenges of maintaining a functioning cellular protein pool in the cold (Place and Hofmann 2004; Place et al. 2004).

Besides ensuring proteostasis under native conditions, the chaperome undergoes a critical reorganization during periods of proteotoxic stress such as from heat shock. The stress leads to the activation of the HSR, whereby expression of specific molecular chaperones and co-chaperones sharply increases, which acts to stabilize denaturing proteins and other macromolecules (Richter et al. 2010). The master control of the HSR is the transcription factor, Heat Shock Factor 1 (HSF1). HSF1 is normally held quiescent by association with molecular chaperones during non-stressed conditions, but is released by stress-induced accumulation of denaturing macromolecules in the cytosol (Åkerfelt et al. 2010). The activation of the released HSF1 and the magnitude of the ensuing HSR are further controlled by a variety of post-translational modifications to the transcription factor. The targets of the activated HSF1 are the heat shock elements (HSEs) in the *cis*-regulatory region of stress-responsive genes. The canonical HSE motif consists of three tandem inverted repeats of nGAAn, with the spacing of these repeats and the position of the guanine residue being critical for HSF1 recognition (Amin et al. 1988).

Absent high temperature excursions in the perpetually cold SO, selective pressures on maintaining cellular capability to respond to heat stress would expectedly become relaxed. Accordingly, various Antarctic notothenioids are found to lack a HSR upon experimental heating (Bilyk et al. 2018; Buckley and Somero 2009; Hofmann et al. 2000; Huth and Place 2016; Thorne et al. 2010). How evolution in chronic cold might have affected the native functional character of the full membership of the chaperome remains unexplored. Additionally, whether the relaxation of selective pressure has affected the sequence integrity of the HSR regulatory components, HSF1 and HSEs, potentially contributing to the muted HSR in Antarctic notothenioids, has also not been well studied. Only one prior study each has examined notothenioid HSF1 (Shin et al. 2014) and the *cis*-regulatory region of HSR genes (Bogan and Place 2019).

To gain deeper insight into the molecular basis of the cold-related trait alterations in Antarctic notothenioids, this study investigates the evolutionary status of three key components: the native expression across the full suite of chaperome genes; the integrity of HSF1 as the master control of the HSR; and the integrity of HSEs that are the binding targets of HSF1. These components underlie both the maintenance of cellular proteostasis in the constant ambient cold as well as the capacity to respond to heat challenges. We leverage the recent influx of RNAseq gene expression data that covers a wide sampling of Antarctic notothenioids to quantitatively evaluate the consequence of evolution in the chronic cold on the native expression of the broad suite of chaperome gene members. Furthermore, we utilize high-quality notothenioid genome assemblies that became available in the past year, with additional sequences we generated for select species, to assess whether evolutionary changes in sequence integrity of HSF1 and HSEs have occurred that may affect function.

Materials and methods

Native expression of notothenioid chaperome genes

Species sequence availability

Seven notothenioid species were used in a comparative analysis to assess whether the native expression levels of chaperome genes have been altered during Antarctic notothenioid evolution in chronic cold. These species comprise the largest sequenced taxon sampling that has publicly available RNAseq transcripts from a common tissue, gill, enabling direct cross-species comparison of gene expression. The S. American notothenioid *Eleginops maclovinus*, the monotypic species of the basal family Eleginopsidae and the closest sister taxon to the Antarctic clade, serves as an appropriate proxy for the ancestral state for trait comparison. The other six species include representatives from three of the five Antarctic families: *Dissostichus mawsoni*, *Pagothenia borchgrevinki*, and *Trematomus bernacchii* (Nototheniidae); *Parachaenichthys charcoti* (Bathydraconidae); and *Chionodraco hamatus* and *C. rastrospinosus* (Channichthyidae). Together, these seven species form a basal-to-derived evolutionary series of notothenioid lineages, allowing for detection of Antarctic-specific trait changes. Fish holding conditions were explicitly reported for five of the seven species, with specimens held in flow-through sea water aquaria at native ambient temperatures: ~15 °C for *E. maclovinus*; ≤−1.6 °C for *D. mawsoni*, *P. borchgrevinki* and

T. bernacchii; and ≤−1.0 °C for *C. rastrospinosus* (Bilyk et al. 2018; Chen et al. 2019; Huth and Place 2016). Holding conditions were not explicitly reported for either *P. charcoti* (Berthelot et al. 2018) or *C. hamatus* (Song et al. 2018), only that wild-caught specimens were dissected and tissues were flash frozen in liquid nitrogen.

Identification and isolation of chaperome genes

Membership of the chaperome has been variously defined in past studies (Brehme et al. 2014; Finka and Goloubinoff 2013; García-Prat et al. 2016; Joshi et al. 2018). To identify the notothenioid chaperome members, we built on the relatively restrictive definition of Finka and Goloubinoff (2013). We incorporated all the major HSP families of molecular chaperones and associated co-chaperones from their list, added known teleost-specific chaperone family members, expanded the representations of ER (endoplasmic reticulum) and mitochondrial chaperones, and introduced several key transcription factors (Supplementary Table 1). We used the gene names of this inclusive list to search the annotated genome of the bullhead notothen, *Notothenia coriiceps* (Shin et al. 2014), which yielded 133 putative chaperome genes to form the notothenioid reference set.

All available RNAseq reads from tissues of the seven notothenioids were downloaded from NCBI SRA (Sequence Read Archive) (accession numbers in Supplementary Table 2). The transcript reads were quality trimmed with Fastp (Chen et al. 2018) using default parameters and then mapped to the *N. coriiceps* reference chaperome gene set using Bowtie2 with default parameters (Langmead and Salzberg 2012), thereby allowing us to isolate reads that belonged to chaperome genes for each species. The isolated chaperome reads were then assembled with Trinity v2.6.5 (Grabherr et al. 2011) to produce species-specific chaperome assemblies. To determine the orthologs among the species chaperome assemblies, we performed reciprocal best blast searches against a filtered set of *N. coriiceps* coding domain sequences (CDS). The filtering of *N. coriiceps* CDS was carried out with CD-HIT (Fu et al. 2012) to remove highly similar sequences – in particular, splice variants that confounded cross-species gene identifications. Chaperome contigs from separate species that shared a reciprocal best blast relationship with their matches in the *N. coriiceps* CDS were then considered as orthologs. Of the initial 133 chaperome genes, 119 orthologs could be identified across all seven notothenioid fishes. These form the species-specific chaperome gene sets used in quantifying gene expression.

Differential gene expression analysis

The expression levels of the chaperome genes for each species were quantified by mapping gill RNAseq reads to the species-specific chaperome gene set using Bowtie2 with default settings (Langmead and Salzberg 2012). To enable direct comparisons across species, FPKM estimates were made for each gene using the read counts and effective contig length estimated by RSEM (Li and Dewey 2011) from the Bowtie2 mappings, but using the total read counts rather than only counting the mapped reads. This latter substitution was necessary to avoid double counting changes in expression as the number of mapped reads directly reflected the changing expression of the chaperome genes measured in this study. To identify changes in gene expression, FPKM estimates were $\log_2$ transformed and then analyzed in R using limma (Ritchie et al. 2015) through a generalized linear model (GLM), with species as the only independent variable. Contrasts were then used to identify genes that were differentially expressed (DE) between *E. maclovinus* and each of the Antarctic species using an FDR adjusted *P* value threshold of 0.01. Finally, to validate that the results did not include anomalous mapping to genes that were not included in the chaperome, we re-ran the analysis using stringent mapping settings that allowed only exact sequence matches.

Evolution of the HSF1 coding sequence

Species and HSF1 sequences

HSF1 coding sequences were obtained for 10 notothenioid species. These include the aforementioned seven species with available gill transcriptomes, plus three additional Antarctic species with available transcriptomes of other tissues, namely *N. coriiceps* (Nototheniidae), *Gymnodraco acuticeps* (Bathydraconidae) and *Chaenocephalus aceratus* (Channichthyidae). *HSF1* sequences from the model fish species *Danio rerio* and *Gasterosteus aculeatus* were first used to query the existing notothenioid transcriptomes via BLAST. This produced complete HSF1 CDS for *E. maclovinus* and *P. borchgrevinki*, but only partial sequences from the remaining eight Antarctic fish. To obtain complete HSF1 coding sequences from these eight species, we used the two complete HSF1 CDS as bait to isolate their HSF1 reads using Bowtie2 as described above for the chaperome genes. The isolated HSF1 reads were then assembled using Trinity v2.6.5 (Grabherr et al. 2011), and the gene identity and completeness of the resulting contigs were verified by BLASTX against Uniprot and NCBI's NR database. To validate the accuracy of the bioinformatically-obtained HSF1 CDS, we RT-PCR amplified and sequenced the full length HSF1 cDNA from liver or gill RNA of three species – *E. maclovinus*, *P. borchgrevinki*, and *C. rastrospinosus*. The PCR primers (Supplementary Table 3) were designed from the 5' and 3' UTR (untranslated region) sequences present in the bioinformatically-assembled HSF1 transcripts

of *E. maclovinus* and the Antarctic species *D. mawsoni, N. coriiceps*, and *C. aceratus*. The PCR amplicons were treated with SAP/ExoI (shrimp alkaline phosphatase/ Exonuclease I) and cloned into pGemTeasy vector (Promega); subsequently, one clone from each species was sequenced with BigDye Terminator v.3.1 chemistry (Applied BioSystems) (Genbank accession numbers MT813024- MT813026).

To evaluate sequence changes in HSF1 with Antarctic evolution, the translated HSF1 amino acid sequences from select notothenioid fishes were aligned using Clustal Omega with default parameters. The alignment included both the HSF1 from the model fish *G. aculeatus* (Uniprot Accession G3P880) as the teleost representative and the human HSF1 (Uniprot Accession Q00613) to provide structural context for mapping known functional sites and domains of HSF1 (Anckar and Sistonen 2011; Gomez-Pastor et al. 2018). The aligned HSF1 sequences were then manually inspected for any changes in putative functional domains and post-translational modification sites that might affect HSF1 function in the Antarctic fish.

Testing for changes in selective pressure on Antarctic notothenioid HSF1

HSF1 from the Antarctic species were tested against a reference set of fishes that consisted of the basal notothenioid *E. maclovinus* and a phylogenetically-broad section of teleosts (10 species in 8 orders) that have genome data in Ensembl (accession numbers in Table S2). The HSF1 CDS were codon-aligned with MUSCLE (Edgar 2004) using default settings and deletion of gap regions. This generated a final multiple sequence alignment of 1299 nt (433 amino acids) in length.

The requisite phylogenetic framework for testing selective pressure on HSF1 was a Bayesian inferred phylogenetic tree constructed from mitochondrial *16S*, *ND2*, *COI*, and *Cytb* gene sequences that were downloaded for all species from databases (accession numbers in Supplementary Table 2). The four gene sequences were concatenated and then aligned using MUSCLE. Evaluation of models of nucleotide substitution for the alignment utilized jModelTest (Posada 2008), which identified GTR+I+G as the best fit model for *16S*, *ND2* and *Cytb*, and HKY+I+G for *COI*. The respective substitution model was specified for each gene partition and implemented in Bayesian phylogenetic analyses using MrBayes v.3.2.6 (Huelsenbeck and Ronquist 2001). The Markov chain Monte Carlo simulation was run for 10 million generations with four chains and sampled every 100 generations. Tracer 1.4.1 (Drummond and Rambaut 2007) was used to examine the resulting trace files to verify the chains reached convergence, and the first 25% of sampled trees were discarded as burn-in.

Notothenioid HSF1 was then tested for changes in selective pressure using a suite of tools available in HyPhy (Pond et al. 2005) through the Datamonkey Adaptive Evolution Server (http://datamonkey.org/; Weaver et al. 2018). The codon-aligned nucleotide sequences were first analyzed with RELAX to test for general shifts in selective pressure (Wertheim et al. 2015). As the RELAX analyses detected an intensification of selective pressure on the root branch to the Antarctic notothenioids, we further analyzed the aligned sequences with aBSREL (adaptive Branch-Site Random Effects Likelihood) (Smith et al. 2015) within the HyPhy software package to evaluate the occurrence of positive selective pressure (the other possibility being purifying selection). Finally, as a complement to aBSREL, a MEME (Mixed Effects Model of Evolution) (Murrell et al. 2012) analysis was conducted to assess whether signatures of positive selection could be isolated to specific sites within the HSF1 protein.

Evolution of the HSP70 *cis*-regulatory region

We used the members of the *HSP70* gene family as candidate heat-responsive genes for examining the structural integrity of the HSEs in Antarctic notothenioids. The presence and positions of HSEs were determined for eight of the nine *HSP70* gene family members that we identified in Antarctic notothenioid chaperomes (Supplementary Table 1). The sequences of the *HSP70* paralogs were curated for seven notothenioids that have published genome assemblies. These include the basal *E. maclovinus* (Chen et al. 2019) and six Antarctic species – *N. coriiceps* (Shin et al. 2014), *D. mawsoni* (Chen et al. 2019), *P. charcoti* (Ahn et al. 2017), *C. aceratus* (Kim et al. 2019), *Chionodraco myersi* and *C. hamatus* (Bargelloni et al. 2019). For each *HSP70* paralog, 3 kbp of genomic sequence upstream of the start codon were isolated and screened for the HSE sequence motif nGAAn/nTTCn using JASPAR (Khan et al. 2017) with the default relative profile score threshold of 80%.

Our initial analysis of the copy number and locations of putative HSEs in the *cis*-regulatory region identified a positional shift of the HSE cluster of the *HSPA6* gene – the key inducible member of the HSP70 family – in the Antarctic species relative to *E. maclovinus*. To verify this positional shift, we PCR-amplified the upstream region of *HSPA6* as well as *HSP70-like* (the second inducible *HSP70* member) from *E. maclovinus* and two additional Antarctic species, *P. borchgrevinki* and *C. rastrospinosus*. Amplification was carried out using primers developed from the genomic sequences of the two paralogs available from *E. maclovinus*, *D. mawsoni*, and *N. coriiceps* (Supplementary Table 3). The PCR amplicons were treated with SAP/ExoI,

cloned into pGemTeasy vector (Promega), and sequenced with BigDye Terminator v.3.1 chemistry (Applied Bio-Systems) (Genbank accession numbers MN275890-MN275895).

We used STAR aligner v 2.6 (Dobin et al. 2013) with available single ended RNAseq gill reads to delineate the transcription start site (TSS) and identify intron sequences in the 5' untranslated region (UTR) of the cloned upstream sequences of *HSPA6* and *HSP70-like*. In addition, the *cis*-regulatory regions of all members of the HSP70 gene family were screened for potential transposable element (TE) insertions using the Dfam database.

Results

Native expression of notothenioid chaperome genes

Of the original 133 chaperome genes identified from the *N. coriiceps* genome, 119 have orthologs in the seven notothenioid species we studied, and 107 of these have detectable gill expression in all species (Supplementary Table 1). When compared to the ancestral proxy *E. maclovinus*, the native chaperomes of the six Antarctic species showed a common modest transcriptional response by the measure of $\log_2$FC, with nearly all differentially expressed (DE) genes falling within the 2 and −2 $\log_2$FC interval (Fig. 1). Three species, *P. borchgrevinki, C. rastrospinosus*, and *T. bernacchii*, for which we have replicated gill RNAseq data (Bilyk et al. 2018) allowing robust quantification, showed an even more restricted response, with most DE genes exhibiting changes <1 $\log_2$FC (Fig. 1a). The number of DE chaperome genes ranged from 56 to 64 in the six Antarctic species, representing a slight majority (52%–59%) of the identified chaperome members. However, only 11 of these chaperome genes were consistently differentially expressed in all six Antarctic notothenioids compared to *E. maclovinus* (Fig. 1b).

The scatterplots in Fig. 2 show chaperome gene expression for the six Antarctic notothenioids, with the nine members of the *HSP70* family indicated. In accordance with Fig. 1a, most changes in expression were modest relative to *E. maclovinus*. Three *HSP70* gene family members – *HSPA6, HSP70-like*, and *HSC71-like* – were noticeable for larger $\log_2$FC expression increases. Of these, the key inducible *HSPA6* uniquely exhibited a large increase, exceeding 4 $\log_2$FC across all species. In contrast, *HSP70-like* and *HSC71-like* showed a more lineage-specific pattern of change. *HSPA8*, the normally highly expressed constitutive *HSP70* member, showed a modest differential downregulation in five of the six Antarctic notothenioids.

The side-by-side comparison of the differential expression level of each of the 107 chaperome genes in the six Antarctic species (Fig. 3a) highlights the generally modest changes. Most of the chaperome genes exhibited expression change in one or more species, but only a small subset of 11 genes were differentially expressed in all six fishes (Figs. 1b and 3b). *HSP70* gene members and the HSP70 co-chaperone *HSPBP1* comprise the largest single component in this subset (5 of 11 genes). The remainder included one member of the *CCT* (Chaperonin Containing TCP1-complex), three ER *cis-trans* peptidylprolyl and prolyl isomerases (*PPWD1* and *FKBP* paralogs), and the mitochondrial chaperone *BCS1*, which were all mildly upregulated, and the proteasome assembly related chaperone *psmg2* that exhibited mild downregulation (Fig. 3b). Again, most distinctive among cytosolic chaperones and indeed across the entire chaperome is the consistently >4 $\log_2$FC increase of the ancestrally inducible *HSPA6* (Fig. 3a, b). Remapping gill RNAseq reads using highly stringent criteria (S Material 1) corroborated these expression patterns obtained with default mapping criteria. In sum, evolutionary changes in native gene expression of the chaperome appear to be mainly associated with the cytosolic *HSP70* family members, with *HSPA6* being the single paralog that is consistently and prominently upregulated across species.

Evolution of the HSF1 coding sequence

To evaluate sequence integrity, HSF1 coding sequences (CDS) were curated for 10 notothenioids with available transcriptomes. The reliability of these bioinformatically-derived HSF1 CDS was verified by the cloned full length HSF1 cDNA from three of the species: *E. maclovinus, P. borchgrevinki*, and *C. rastrospinosus* (Supplementary Fig. 1). Both HSF1 nucleotide and translated amino acid alignments (Supplementary Fig. 2) showed high sequence conservation among Antarctic species (>98% to 100% identities), whereas HSF1 of *E. maclovinus* was more dissimilar at about 91% identity with the Antarctic fish HSF1. The comparison also showed two major Antarctic-specific changes: a 17-residue deletion, and a 3-residue indel (Supplementary Fig. 2).

The HSF1 protein sequences of *E. maclovinus P. borchgrevinki*, and *C. rastrospinosus* were aligned with the human and stickleback references (Fig. 4), with known post-translational modification (PTM) sites in human HSF1 indicated (Gomez-Pastor et al. 2018). Most of the PTM sites were found to be conserved between the human reference and the notothenioids. One notable difference is the substitution of S303 (in human; human HSF1 numbering) to N303, which had previously been

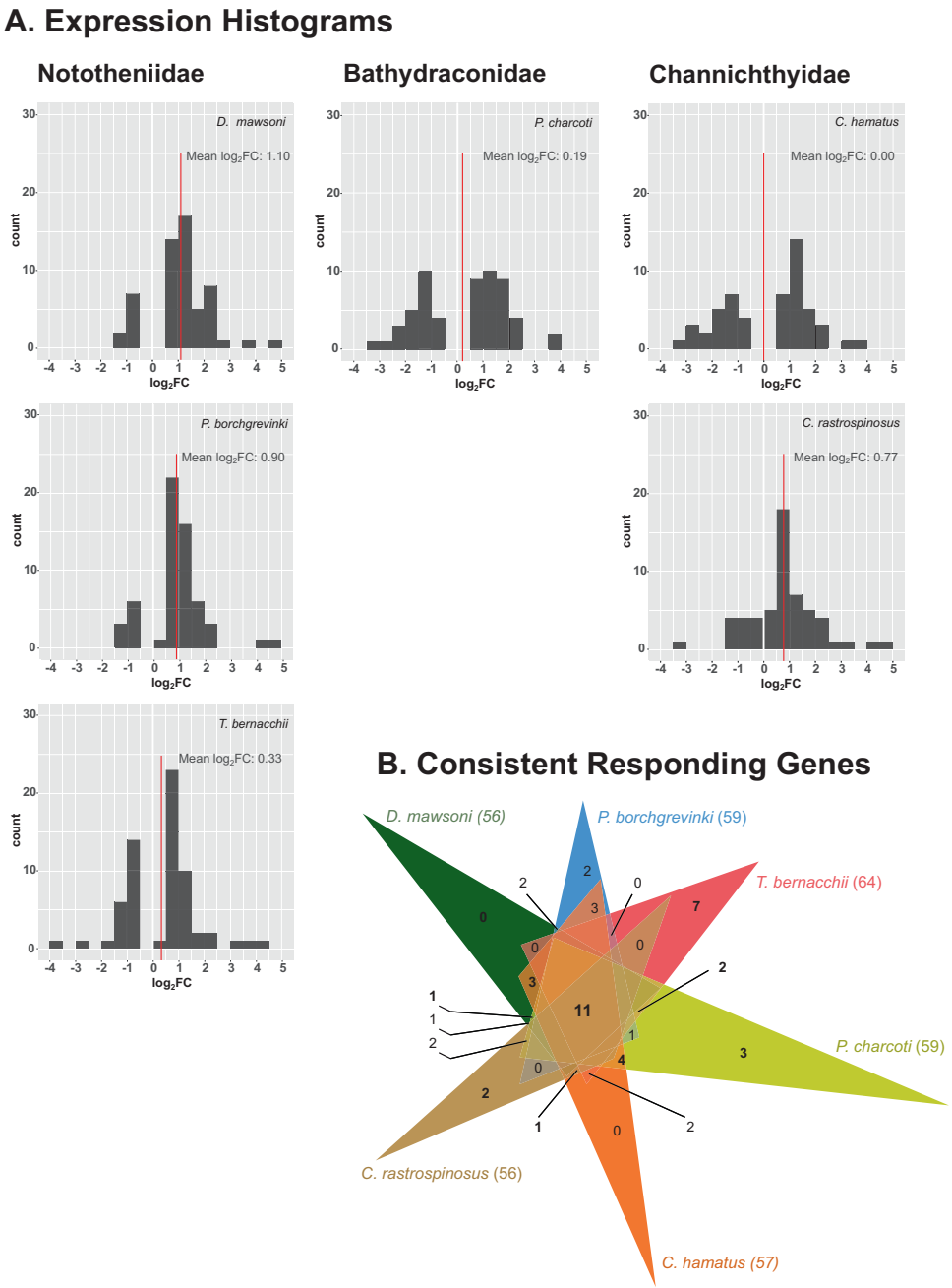


Fig. 1 Profiles of differentially expressed (DE) chaperome genes in gill tissue of Antarctic notothenioids compared to *E. maclovinus*. **a** Histograms show small log_2FC in native expression of chaperome genes in gill tissue of six Antarctic notothenioid species (organized by family) compared to *E. maclovinus*. Vertical red lines indicate mean log_2FC for each species, with the value given. **b** The total number of DE chaperome genes of each species (given in parenthesis next to species name) ranged from 56 to 64. The Venn diagram depicts the 11 consistently DE genes shared by all six Antarctic species. The number of DE genes specific to each species and the number of DE genes shared between species pairs are also indicated. For clarity, the number of DE genes shared by 3, 4, or 5 species are not included in the diagram.

identified in *N. coriiceps* by Shin et al. (2014) as a functionally important Antarctic-specific change. Here, we found S303/N303 is in fact plesiomorphic, with N303 already present in *E. maclovinus* (Fig. 4). Between *E. maclovinus* and the Antarctic species, more amino acid substitutions are observed in the regulatory and transactivation domains that comprise the C-terminal half of HSF1. However, none of the substitutions occur at putative PTM sites based on the human HSF1 reference, except for one apparent reversal in a potential glycosylation site (Thr367 in human HSF1) – from a Pro in *E. maclovinus* to a Ser in Antarctic species (Fig. 4).

Using the consensus teleost tree (Supplementary Fig. 3), RELAX analyses of the codon-aligned HSF1 CDS from a wide range of temperate teleosts and Antarctic notothenioids (Supplementary Fig. 4) identified a signature of intensified selective pressure strength on HSF1 (Fig. 5a, Supplementary Table 4). Serial testing of different permutations of background (reference) and foreground (test) branches localized the selective intensification to the root branch of the Antarctic notothenioid clade, where the value of the intensity parameter (K) was greatest (Fig. 5a.iii, Supplementary Table 4). From the aBSREL analyses, the signature of intensification was found to include significant

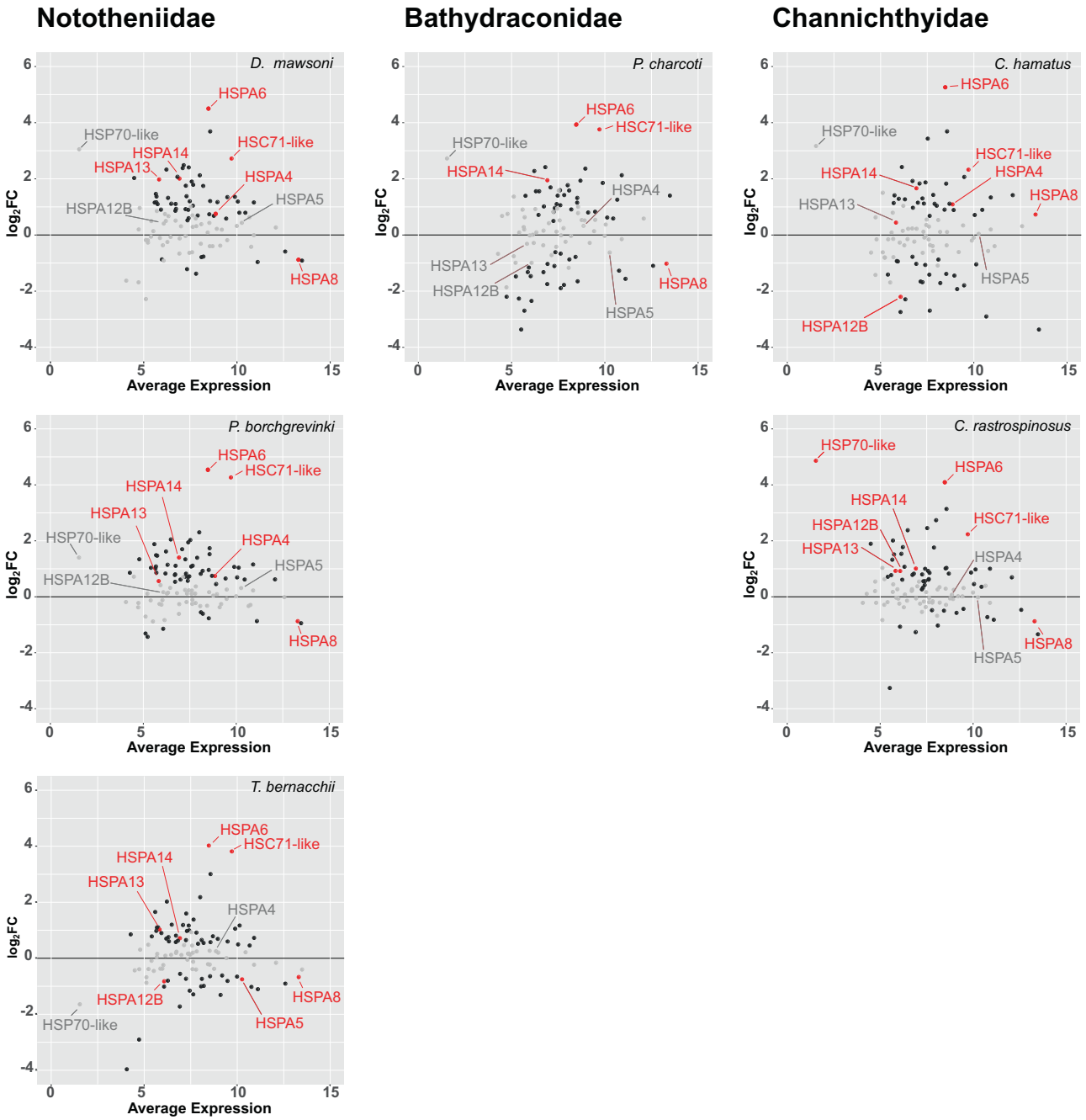


Fig. 2 Gene expression profile of the HSP70 gene family members in Antarctic notothenioids. Scatterplot of log_2FC in expression of chaperome genes in the six Antarctic notothenioids relative to *E. maclovinus* in Fig. 1, highlighting the nine members of the HSP70 gene family. Each dot represents one chaperome gene, with black and gray color respectively indicating significant and insignificant expression difference relative to *E. maclovinus*. The HSP70 gene family members are labeled, and those indicated in red significantly differed in expression from *E. maclovinus*. Most DE chaperome genes showed only modest increases in expression, at <2 log_2FC relative to *E. maclovinus*. *HSPA6* of the HSP70 family uniquely exhibited a consistent large increase (>4 log_2FC) in all examined Antarctic species.

positive selection on HSF1 at the root branch (Fig. 5b. i–iii, Supplementary Table 5). Together, RELAX and aBSREL results suggest that intensified positive selection on HSF1 occurred in the common ancestor of Antarctic notothenioids (root branch). Subsequent to that, HSF1 remained highly conserved within the Antarctic clade (Supplementary Fig. 2), indicating that purifying selection acted to maintain the Antarctic-specific mutations. Analyses using MEME identified eight positively selected sites (at 0.1 level of significance) in HSF1, three of which are substitutions specific

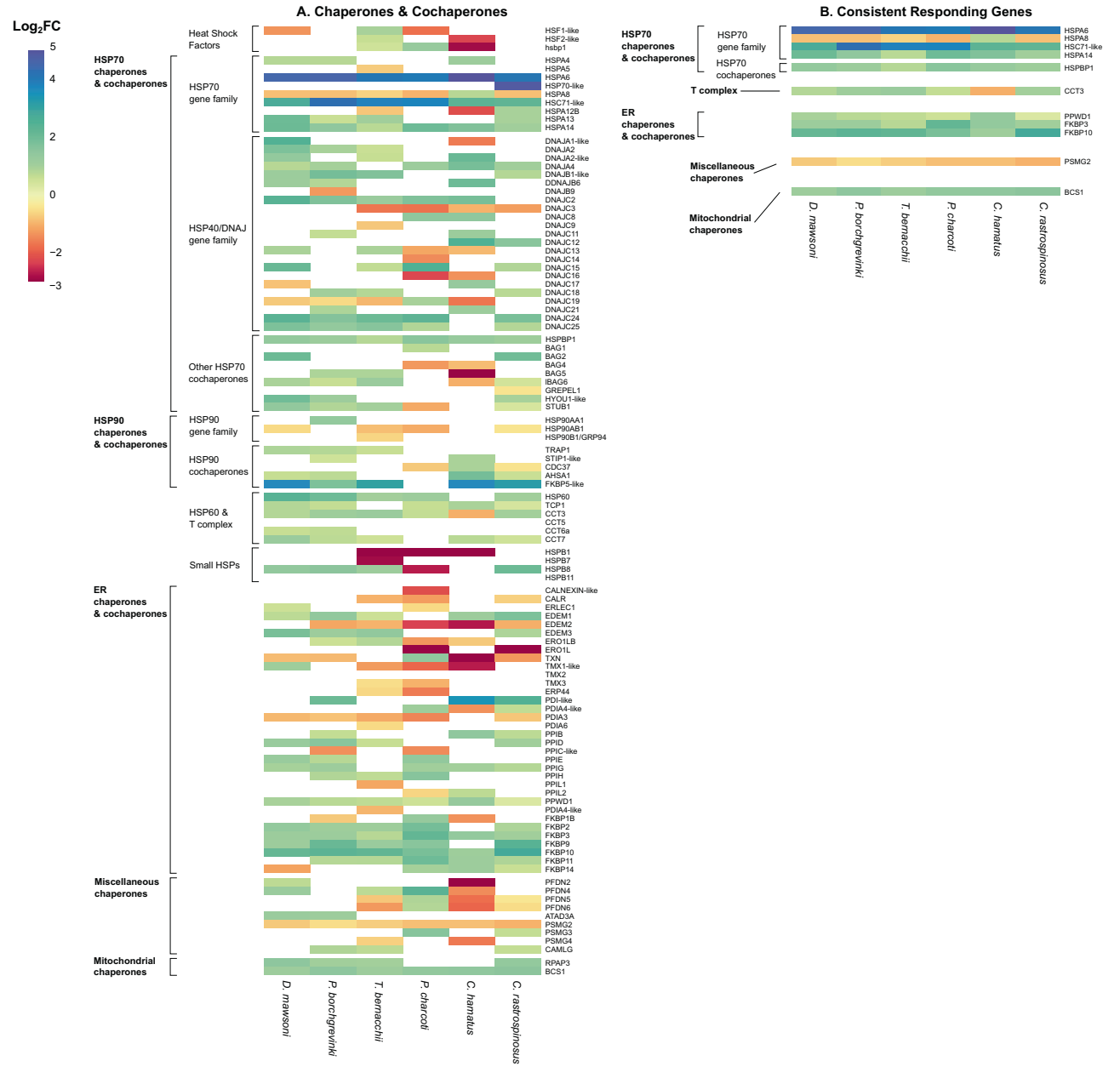


Fig. 3 Heat map of chaperome gene expression. Changes in native expression levels of the 107 chaperome gene members from six Antarctic notothenioid species compared to *E. maclovinus*. Gene names are indicated on the right of the heat map, and the gene family they belong to are labeled on the left. Different colors depict significant Log_2FC changes as indicated by the color key and scale bar. **a** Expression level changes of each of the 107 gene members of the Antarctic notothenioid chaperome. **b** Changes in the subset of 11 chaperome genes that consistently responded in all Antarctic species.

to the Antarctic clade. They are (human/*E. maclovinus* /Antarctic notothenioids) L289/P/S or T, R352/N/S, and K463/K/R or G (Supplementary Fig. 4).

Evolution of the HSP70 cis-regulatory region

We isolated the upstream *cis*-regulatory regions for eight of the nine identified HSP70 gene family members. The ninth gene, *HSPA12B*, could not be included for analysis because its upstream sequence was absent in the *E. maclovinus* genome assembly, our proxy for the ancestral state. The mapped locations of putative HSEs in the upstream regions are depicted in Fig. 6 for *HSC71-like, HSPA8, HSPA6*, and *HSP70-like*, the four genes that exhibited notable changes in native expression. Three of these genes (*HSC71-like, HSPA8*, and *HSP70-like*) showed a conserved pattern of HSE locations between *E. maclovinus* and the Antarctic species. The exception is *HSPA6*, the only *HSP70* with a

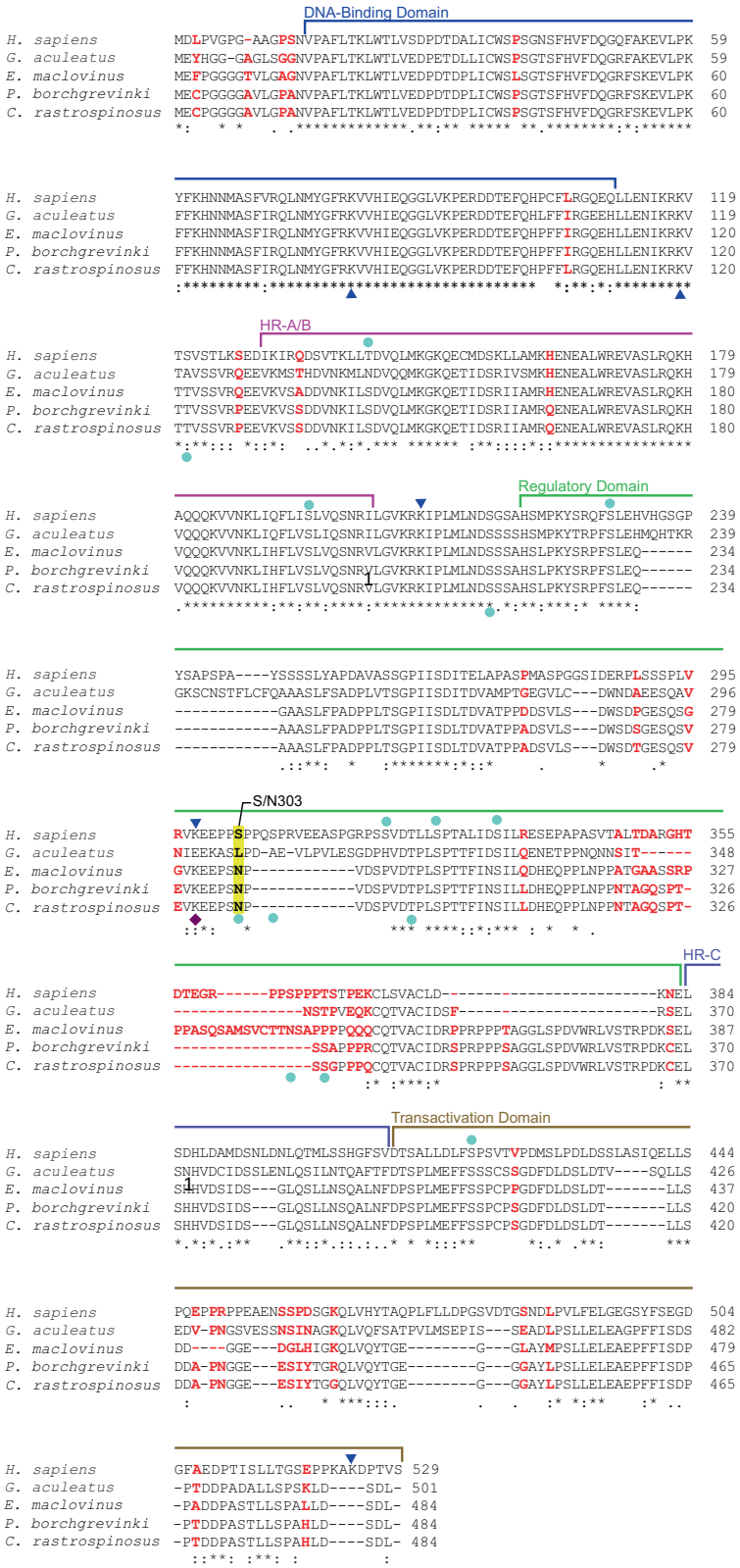


Fig. 4 Sequence alignment of HSF1 from basal *E. maclovinus* and representative derived Antarctic notothenioids. The human HSF1 sequence is included to provide a frame of reference for key domains within the gene as previously delineated (Anckar and Sistonen 2011). These include the highly conserved DNA-Binding Domain, HR-A/B (the hydrophobic heptad repeats), the regulatory domain, HR-C (C-terminal heptad repeat), and the Transactivation Domain. The known post-translational modification sites of human HSF1 (Gomez-Pastor et al. 2018) are highlighted as follows: turquois dots indicate phosphorylation, navy triangles indicate acetylation, and purple diamonds indicate sumoylation. The HSF1 sequence from *G. aculeatus* is included for a temperate model fish comparison. Sites that represent substitution in Antarctic species from the basal state in *E. maclovinus* are shown in red characters. The S/N303 substitution previously hypothesized to promote HSF1 activation (Shin et al. 2014) is highlighted in yellow.

consistent large increase in native expression across Antarctic species. While conserved with the homologous segment in *E. maclovinus HSPA6*, its transcription start site (TSS) as well as the copy number and spacing of putative HSEs are shifted further upstream (Fig. 6). The cloned upstream sequences of *HSPA6* from genomic DNA of *E.*

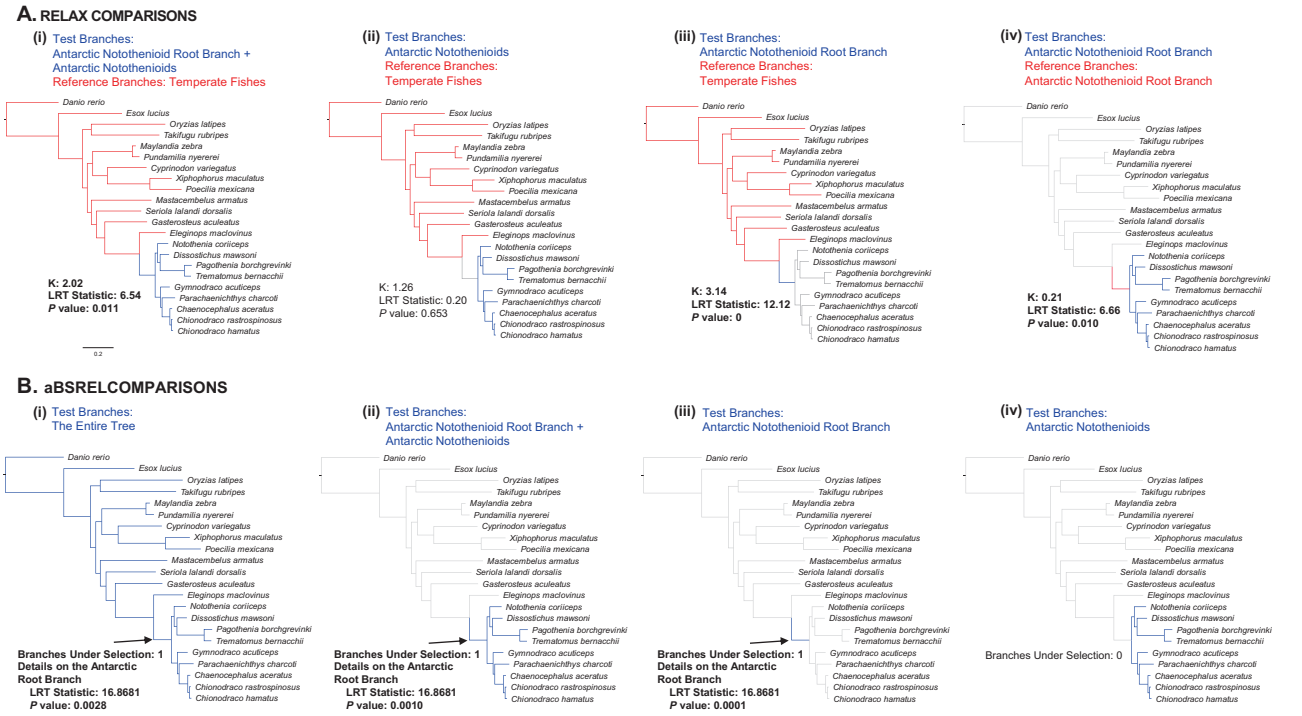


Fig. 5 Tests of selective pressure change (RELAX) and occurrence of positive selection (aBSREL) on HSF1 coding sequence. The RELAX and aBSREL analyses were run using a phylogenetic framework of temperate teleosts and Antarctic notothenioids (Supplementary Fig. 1). **a** Comparisons and findings for the RELAX analyses with the reference branches colored red and test branches colored blue; branches and taxa excluded from the analysis are colored gray. The test statistics are reported under each comparison, including the intensification parameter K where $K > 1$ indicates an intensification of selective pressure and $K < 1$ indicates a relaxation of selective pressure, the LRT Statistic indicating the goodness of fit for the test model, and the *P* value indicating if any significant change in selective pressure was detected. These results are bolded in comparisons where a significant change in selective pressure was detected. **b** Results of the aBSREL comparisons that investigated branches under significant positive selective pressure. Test branches are in blue, with excluded branches and taxa in gray. Under each comparison, the number of branches under significant positive selective pressure are reported along with the LRT Statistic and *P* value for that branch. Significant positive selective pressure was found to act on the root branch (indicated by arrow) to the Antarctic clade in three comparisons.

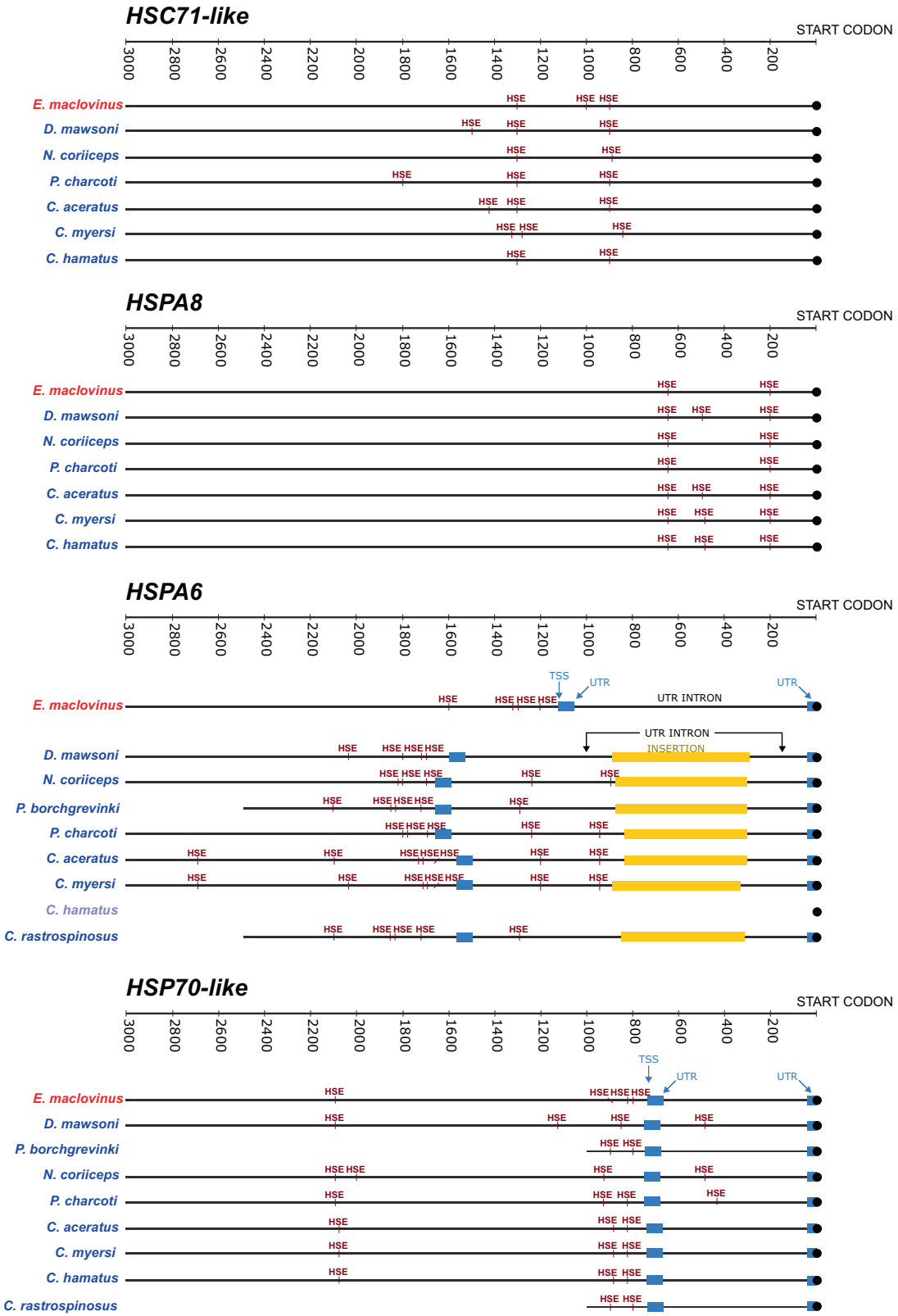
maclovinus, *P. borchgrevinki* and *C. rastrospinosus* validated this positional shift deduced from genome assemblies (Supplementary Fig. 5). The shift was found to result from an insertion in the intron that intervenes in the 5' UTR of *HSPA6*.

HSPA6 and its closest paralog, *HSP70-like*, both contain an intron in their respective 5' UTRs (Fig. 6). Of these two genes, only *HSPA6* shows an expansion of this upstream intron sequence, moving the TSS and the cluster of four HSEs further upstream (Fig. 6; Supplementary Fig. 5). The inserted sequence causing the expansion of the *HSPA6* intron was found to be a hAT Charlie type transposon. Its presence across Antarctic lineages suggests it was inserted early in the diversification of the Antarctic clade.

The upstream regions with mapped putative HSEs for the other four HSP70 gene family members (*HSPA4*, *HSPA5*, *HSPA13*, and *HSPA14*) are presented in Fig. 7. They show variable gains or contractions of TEs in the Antarctic group compared to *E. maclovinus*. However, unlike *HSPA6*, the presence of TEs in these paralogs is not correlated with changes in HSE location and copy number or in native gene expression.

Discussion

The extreme cold of the Southern Ocean and the constancy of this cold condition represent two strong selective pressures acting in concert on the endemic Antarctic notothenioid fishes. Cold temperatures are challenging to the cellular maintenance of proteostasis, while the absence of high temperature excursions would remove the selective pressure on the maintenance of the cellular apparatus that mitigates heat stress. Thus far, very little is known on how these dual selective regimes have influenced the functional character of the full membership of the chaperome. Changes in gene expression at native temperatures have not been explored beyond the heat shock protein *HSP70*, with evidence of co-option of an inducible *HSP70* into "high" constitutive expression. Additionally, evolutionary changes in HSF1 and HSEs that regulate the response are not well



understood. Utilizing the ancestral proxy *E. maclovinus* and a wide sampling of Antarctic notothenioids, this comparative study deduced how the severe cold of the Southern Ocean waters has reshaped chaperome native gene expression and to what level, and identified sequence changes or conservation in *HSF1* and HSEs that are

◀ **Fig. 6 Comparison of *cis*-regulatory regions of *HSC71-like*, *HSPA8*, *HSPA6*, and *HSP70-like*.** Detailed view of the upstream *cis*-regulatory region for the four primary cytosolic HSP70 members. Gene information is excluded from species where the upstream region could not be found or was incomplete in genome assemblies. The locations of HSEs are relative to the start codon. Conserved HSE motifs persist for all four genes, except the cluster of 4 HSEs of *HSPA6* upstream of TSS (transcription start site) was shifted further upstream from the start codon due to a TE insertion within the 5' UTR intron. The 5' UTR exons are represented by blue rectangles, and the TE insertion by yellow bar. With the UTR isolated for this species, the transcription start site is demarcated by TSS within the figure. No TE insertion was detected in the three other *HSP70* paralogs, *HSC71-like*, *HSPA8*, and *HSP70-like*. The 5' UTR and TSS for *HSPA6* and *HSP70-like* were delineated by using cloned and Sanger sequenced upstream sequences of *P. borchgrevinki* and *C. rastrospinosus*.

relevant to their functional role in regulating HSR and heat-responding genes.

Native expression of notothenioid chaperome genes

The freezing water temperatures of the Southern Ocean are thought to present a challenge to proper protein folding and thus cellular proteostasis (Peck 2016). Supporting this is the elevated protein ubiquitination levels observed in the Antarctic notothenioids compared to temperate relatives, suggesting a greater cellular load of misfolded proteins in the cold (Todgham et al. 2007). As the Antarctic notothenioids are also reported to have co-opted the inducible *HSP70* expression into high constitutive expression, this has been interpreted as an adaptive change to support proteostasis at freezing water temperatures (Place and Hofmann 2004; Place et al. 2004). These earlier studies led to a generalized notion of high constitutive heat shock protein expression in Antarctic notothenioids (Peck 2016), but lacked precise knowledge of the responsible gene members and their quantitative levels.

In this study, we found that evolution in chronic cold has not markedly altered the native expression of most chaperome genes in the Antarctic notothenioids. While most Antarctic species tend towards increased expression across chaperome genes relative to the ancestral proxy *E. maclovinus*, this increase is at low ≤ 1 $\log_2$FC levels (Figs. 1 and 2). Furthermore, only 11 chaperome genes show a consistent response across species, and these are comprised of mostly a single chaperone class, the cytosolic members of the HSP70 gene family (Fig. 3b). Thus, evolution in chronic cold appears to have primarily affected these particular members of the chaperome in regards to native expression levels.

Members of the HSP70 chaperone family fulfill a generalist role as they do not have specific client proteins. They directly recognize hydrophobic residues exposed in all proteins in non-native state, thus serving as sentinels of general cellular proteostasis (Balchin et al. 2016; Murphy 2013; Preissler and Deuerling 2012). Apart from these cytosolic HSP70 members, the generally small increases in expression across the chaperome are not consistent with the hypothesis that the HSR in Antarctic notothenioids is constitutively activated as suggested by Shin et al. (2014), or with high constitutive expression of HSPs (Peck 2016). Instead, it is possible that a wide range effect on proteostasis could arise from augmenting the expression of a small number of generalist molecular chaperones, delivering broad protection of the native cellular protein pool in the cold. A relevant regulatory change that would enhance expressions of a few HSP70 genes could provide this broad benefit and limit demands for changes in the rest of the chaperome. Increased expression was specific to three cytosolic HSP70 members: *HSPA6*, *HSP70-like*, and *HSC71-like*. Of these, the ancestrally inducible *HSPA6* is the sole HSP70 member to show a large (> 4 $\log_2$FC) expression increase that was consistent across the Antarctic species (Figs. 2 and 3). The trait alteration of this key formerly-inducible *HSP70* reflects a gain of function, and is consistent with the hypothesis of evolutionary regulatory change. The strong and consistent native upregulation of *HSPA6* across examined Antarctic notothenioid lineages also indicates this inducible *HSP70* paralog is the key contributor to the heightened constitutive *HSP70* expression reported in prior studies.

Along with the generalist HSP70 members, HSPA14 a key chaperone in the ribosome-associated complex shows a small but consistent increase across Antarctic species. This plays an essential role in assisting with early folding of nascent polypeptides as they exit the ribosome (Jaiswal et al. 2011; Preissler and Deuerling 2012; Otto et al. 2005). While the increase in expression of this gene is relatively small, this increase remains consistent with past expectations that initial protein folding is a challenge near freezing water temperatures.

One exception to regulatory change causing enhanced *HSP70* expression is *HSPA8*. Normally the primary constitutively expressed cytosolic HSP70 member, *HSPA8* showed a modest reduction in expression in five of the six examined Antarctic notothenioids (Figs. 2 and 3a). It seems counterintuitive that *HSPA8* expression should decrease in the face of perpetual challenges of protein folding in the cold. Whether this change in *HSPA8* expression persists across tissues, and whether it preceded or followed the evolutionary changes in expression in the other HSP70 members will require further investigation.

Evolution of the HSF1 coding sequence

The lack of thermal variability in the Southern Ocean over geological time would have relaxed evolutionary selective

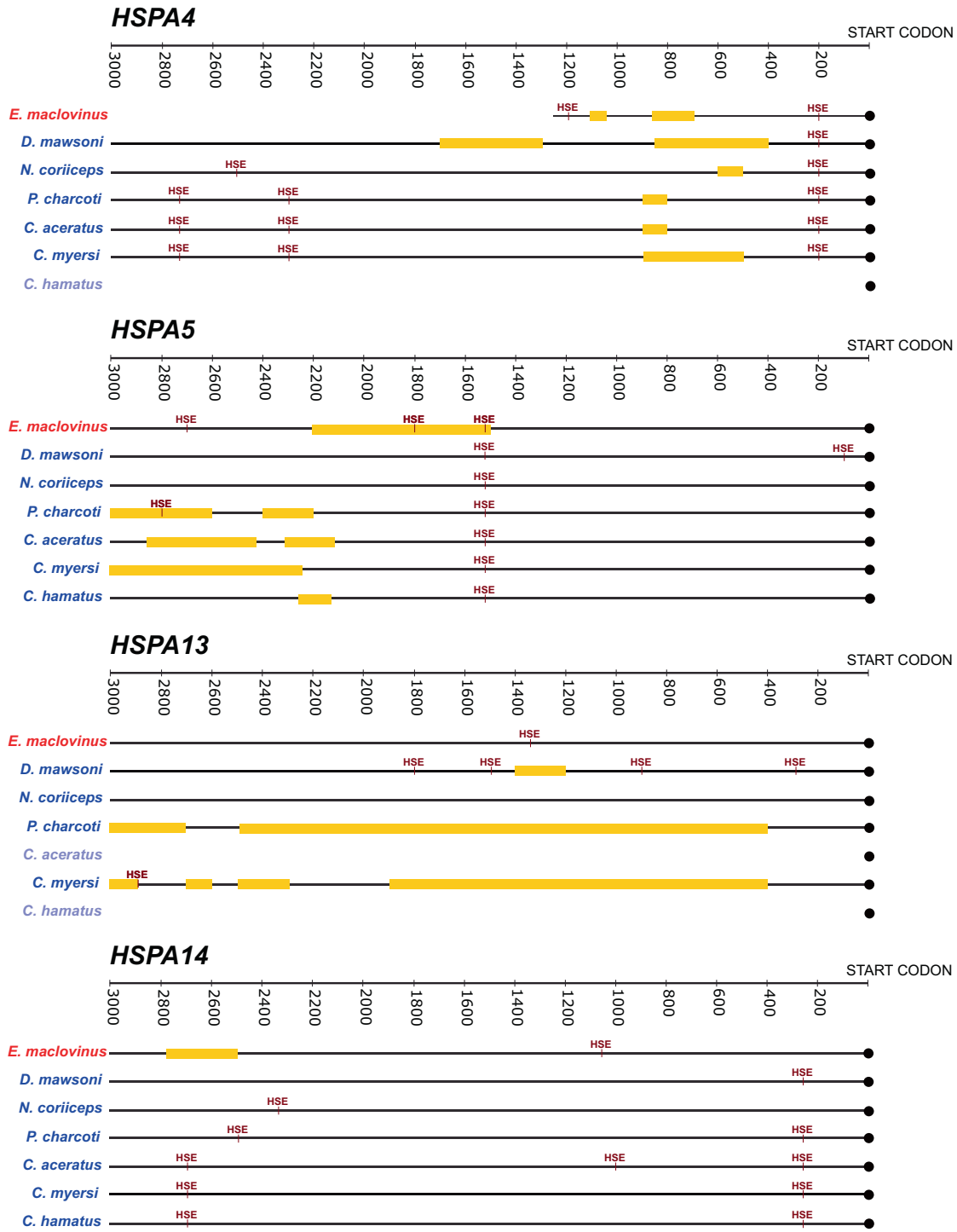


Fig. 7 Comparison of *cis*-regulatory regions of *HSPA4*, *HSPA5*, *HSPA13*, and *HSPA14*. Detailed view of the upstream *cis*-regulatory region for four other *HSP70* members in Antarctic notothenioids using available sequences from *E. maclovinus* as the template for alignment and mapping. Absent gene information for orthologs in some species (e.g. *HSPA5* for *C. hamatus*) was due to upstream regions being either incomplete or unreported. The final member, *HSPA12B*, has an incomplete upstream region in the *E. maclovinus* genome assembly, precluding its use as a template to map its ortholog in the Antarctic species. Lineage-specific variation in TE insertion (yellow bar) gains or contractions were observed.

pressure on maintaining the cellular capability to respond to acute heat stress. Ostensibly, the relaxed selection is reflected in the loss of the HSR in the modern Antarctic notothenioid fishes. Yet besides the overt loss of the HSR, it has remained unclear whether and how this expected relaxed selection might have affected the network of genes

that once responded to heat stress in the Antarctic notothenioid ancestor. As the master controller of the HSR (Åkerfelt et al. 2010), HSF1 is a natural candidate to evaluate for the occurrence of relaxed selection.

A previous study reported a non-conservative, Antarctic-specific amino acid substitution at a putative phosphorylation site in HSF1, from S303 to N303 (Shin et al. 2014), and posited that it has crucial functional consequences. S303 in mammalian HSF1 is one of the known constitutively phosphorylated sites, which negatively regulates (represses) HSF1 in unstressed conditions. Substitutions that preclude phosphorylation of S303 would prevent phosphorylation-dependent sumoylation of the nearby Lys298 (Hietakangas et al. 2006), de-repressing HSF1 and leading to high constitutive activity (Kline and Morimoto 1997). If a similar outcome had occurred in the notothenioids, it would be consistent with strong activation of HSF1 and constitutive high levels of molecular chaperone expression previously proposed (Shin et al. 2014; Peck 2016).

However, we found evidence to the contrary in this study. First, increases in native expression were muted for most molecular chaperone genes, not indicative of a constitutively activated HSR. Second, we found the S/N303 substitution is not Antarctic-specific, but actually predated the Antarctic notothenioid radiation, as it already exists in the HSF1 sequence of the basal *E. maclovinus* (Fig. 3). The substitution has not altered native chaperone gene expression in *E. maclovinus*, as this species has been shown to mount a robust inducible HSR (Bilyk et al. 2018). Thus, the S/N303 substitution could not be responsible for the de-repression of HSF1 or alteration of the HSR in the Antarctic species as previously proposed. The other putative PTM sites based on the human HSF1 template are largely conserved between notothenioids and human, as well as between *E. maclovinus* and Antarctic species. However, whether any of these putative PTM sites serve similar post-modification roles as mammalian HSF1 is yet unknown, as experimental determination of PTM sites in teleost fish HSF1 is completely lacking at present.

On testing for selective pressure change on the HSF1 CDS, we found a consistent signature of change on this transcriptional factor in the Antarctic notothenioids, localized to the root branch leading to the Antarctic clade (Fig. 5, Supplementary Tables 4 and 5). Contrary to the intuitive expectation of relaxed selective pressure on HSF1 due to the absence of environmental thermal variation, the detected change was in fact an intensification of selective pressure on HSF1 (Fig. 5a, Supplementary Table 4) due to positive selection (Fig. 5b, Supplementary Table 5). The HSF1 sequences of Antarctic clade species are extremely conserved (>98% to 100% identities) at both the nucleotide and protein levels. The nine examined species span three Antarctic families, from the least derived Nototheniidae to the most derived icefish family Channichthyidae. Thus, the homogeneity of their HSF1 sequences indicates a strong likelihood of purifying selection operating in the clade. Taken together, these results suggest that the greatest evolutionary influence on the HSF1 CDS occurred during the earlier cooling of the Southern Ocean in the ancestor of the Antarctic notothenioids. Subsequently purifying selection acted to maintain the tinkered HSF1 during the subsequent Antarctic notothenioid radiation.

Three residues in HSF1 with changes specific to the Antarctic clade were identified as experiencing positive selection. None of these correspond to known human HSF1 PTM sites, but they have undergone substitution with respect to the basal *E. maclovinus* HSF1. They are S(T)/P/L289 (in the order of: Antarctic notothenioids/*E. maclovinus*/human; human numbering), S/N/R352, and R(G)/K/K463. The first two substitutions, Ser(Thr)/Pro and Ser/Asn, could potentially result in the gain of a novel phosphorylation site. These two sites are located in the regulatory domain of HSF1 and may affect the activation or repression of HSF1 in the Antarctic notothenioids, depending on whether their regulatory effect is positive or negative. Addressing this hypothesis will require systematic empirical determination of the function of all potential PTM sites in teleost fishes.

While most individual amino acid substitutions were conservative or somewhat conservative (Supplementary Fig. 2), the great majority of protein sequence changes were concentrated in the C-terminal half of the HSF1 that contains the regulatory and transactivation domains (Supplementary Fig. 2). This region also shows a prominent 17-residue Antarctic-specific deletion and a 3-residue Antarctic-specific indel. The high concentration of amino acid changes in the two domains suggests that the evolutionary tinkering of Antarctic HSF1 may pertain to the control (regulatory) of the progression and the magnitude (transactivation) of the HSR (Anckar and Sistonen 2011; Dai and Sampson 2016; Vihervaara and Sistonen 2014), both of which are testable hypotheses for future functional investigations.

Even with the loss of the HSR in Antarctic notothenioids, the continued expression of HSF1 at levels comparable to the basal *E. maclovinus* in most Antarctic species (Fig. 3a) suggests its continued participation in a non-HSR capacity. While HSF1 is best known as the master control of the HSR, it has other essential functions including stress-independent regulatory activity over a wide range of genes involved in development, metabolism, and longevity (Brunquell et al. 2016). It also appears to serve an essential role in maintaining proteostasis even outside of heat shock (Solís Eric et al. 2016). The persistence of HSF1 expression combined with the presence of sites under positive selective

pressure may reflect a process of optimization for some of these non-stress related tasks.

Evolution of the HSP70 *cis*-regulatory region

The *cis*-regulatory or promoter regions of genes contain regulatory elements that transcriptional factors bind to and control gene expression. The HSEs of heat-responsive genes are specific targets of the activated HSF1, the master control of HSR. Alteration to the HSE sequence during notothenioid evolution in chronic cold could alter binding specificity, potentially leading to trait changes such as the observed silencing of HSR and the prevalent modest upregulation of native chaperome genes. We tested this hypothesis by comparative analyses of the *cis*-regulatory regions of the HSP70 gene family between the ancestral proxy *E. maclovinus* and derived Antarctic notothenioids.

Contrary to the anticipated relaxation of selective pressure in constant cold that may lead to decay of HSEs, we found that the HSEs of the member genes of the *HSP70* family were generally conserved in copy number between *E. maclovinus* and the Antarctic notothenioids, even among ancestrally inducible genes (Figs. 6 and 7). Recent findings by Bogan and Place (2019) did identify that cryonotothens (Antarctic notothenioids) exhibited fewer HSEs per base pair relative to temperate notothens (*Cottoperca gobio* and *E. maclovinus*). Here we found that this decrease in HSEs was very modest. Moreover, using the two ancestrally inducible *HSP70* members as representatives, we found nucleotide changes within the HSE nGAAn/nTTCn motif are also minimal. This conservation is illustrated in the alignment of 5' sequence with the identified HSE motifs therein, upstream of the transcription start site (TSS) in *HSPA6* and *HSP70*-like from *E. maclovinus* and the Antarctic species (Supplementary Fig. 5). Substitutions relative to *E. maclovinus* were mostly in the naturally variable first and fifth positions of the HSE motifs. Changes in the outer nucleotides nevertheless have been shown to reduce heat shock factor binding affinity, but not to the extent of changes to the central GAA or TTC of the motif (Cunniff and Morgan 1993). Therefore, the anticipated relaxation of selective pressure in chronic cold appeared not to have resulted in loss or degradation of the HSEs of the *HSP70* family, the crucial effector of the HSR.

In parallel with the loss of inducibility, *HSPA6* was unique in its consistently large (>4 $\log_2$FC) differential increase across all examined Antarctic lineages, as mentioned above. Correlating with this prominent increase is an expansion of the 5' UTR intron of *HSPA6* due to the insertion of a hAT Charlie type transposon. This TE insertion likely occurred at the dawn of the Antarctic notothenioid radiation, as it is present in the *HSPA6* of all seven examined Antarctic species spanning three families, from the basal Antarctic Nototheniidae family to the most derived icefish family Channichthyidae (Fig. 6).

TEs could function as an important source of evolutionary innovation. Various lines of evidence showed the correlation of TE activity increase with periods of abiotic stress, with the ensuing TE insertions serving to create genetic variability upon which selective pressure could act (Belyayev 2014; Capy et al. 2000; Cappucci et al. 2019). Comparative genome analyses between the basal temperate *E. maclovinus* and Antarctic notothenioid *D. mawsoni* provided evidence that the latter underwent a burst of TE expansion, estimated to occur in the late Miocene as Southern Ocean temperatures sharply declined, when Antarctic notothenioid diversification and radiation began (Chen et al. 2019). TE insertions have also been found to be associated with increased expression of nearby genes following activation from stressful conditions (Horváth et al. 2017), including examples of formerly-inducible genes switching to constitutive expression in yeast (Williamson et al. 1981) and wheat (Tovkach et al. 2013). The insertion of a hAT Charlie TE in the regulatory region of Antarctic notothenioid *HSPA6* could have occurred during the burst of TE expansion in the Antarctic notothenioids, leading to altered gene expression as manifest in the strong increase in modern native expression, or loss of inducibility. If this TE insertion is found responsible for increased native *HSPA6* transcription, then the increased levels of this particular generalist molecular chaperone could be responsible for bolstering proteostasis in Antarctic notothenioids as the Southern Ocean rapidly chilled during mid-Miocene. This hypothesis is amenable to empirical investigation using a transgenic model fish or a fish cell line system.

Conclusion

This study examined potential molecular changes in the shared cellular apparatus underlying proteostasis and heat stress response associated with Antarctic notothenioid evolution in constant cold. We use RNAseq and transcript profiling to examine, for the first time, the native expression levels of a broad swath of chaperone and co-chaperone genes that comprise the chaperome responsible for cellular proteostasis. We detected only subtle changes in the native expression of most chaperome genes in Antarctic species compared to their temperate sister notothenioid that represents the ancestral condition. Among the small number of genes that exhibited larger differential expression, only *HSPA6* – an ancestrally inducible *HSP70* member – showed a consistent strong >4 $\log_2$FC increase across all Antarctic species. The strong native expression of *HSPA6* indicates it is the key *HSP70* paralog contributing to the heightened constitutive *HSP70* expression in earlier reports. Its change

from inducible to strong constitutive expression is correlated with a hAT-transposon insertion within its 5' UTR intron, unique within the HSP70 gene family.

Contrary to the anticipated relaxation of selective pressure in constant cold on maintaining the regulators that once responded to large thermal variations, we found an intensification of selective pressure on HSF1, the master control of the HSR. This intensification included contributions from positive selective pressure on the ancestral branch leading to the Antarctic clade of notothenioids, suggesting that following the silencing of the HSR the roles HSF1 play in cellular well-being outside of the HSR are being selected upon. In examining the HSEs, the canonical targets of HSF1, we find only limited sequence variations in these *cis*-regulatory elements, suggesting a relatively limited impact of relaxed selective pressure following the loss of the HSR. Consequently, evolutionary changes to their expression patterns must have involved other factors besides the HSE motifs. In summary, in this study we detected both subtle and distinctive changes in the chaperome genes and the regulatory HSF1 and HSEs, associated with Antarctic notothenioid evolution in chronic cold. Our study therefore opens salient questions amenable to experimentation that would stimulate further research into addressing the evolutionary thermal plasticity and adaptive potential of Antarctic notothenioids to a changing environment.

Data availability

Sequencing results for the genomic region upstream of the HSP70 genes have been archived on Genbank through accession numbers MN275890-MN275895. The RT-PCR verified sequence of HSF1 are similarly archived on Genbank with accession numbers MT813024-MT813026.

Acknowledgements The authors thank Katherine R. Murphy for her thoughtful comments on the manuscript in its in initial preparation. A special thanks to Jeffery Haas, Director of Information Technology for the Schools of Integrative Biology and Molecular and Cellular Biology at the University of Illinois, Urbana Champaign, who unfailingly helped through a near endless series of software installation requests. This work was supported by the US National Science Foundation Division of Polar Programs grant ANT-1142158 to C-HCC and AL DeVries. LV-C was supported by Fondap-IDEAL 15150003. The funding agencies of this study played no role in the design, collection, analysis, and interpretation of data and writing of this manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Ahn D-H, Shin SC, Kim B-M, Kang S, Kim J-H, Ahn I et al. (2017) Draft genome of the Antarctic dragonfish, *Parachaenichthys charcoti*. GigaScience 6(8):1–6

Åkerfelt M, Morimoto RI, Sistonen L (2010) Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol 11(8):545–555

Amin J, Ananthan J, Voellmy R (1988) Key features of heat shock regulatory elements. Mol Cell Biol 8(9):3761–3769

Anckar J, Sistonen L (2011) Regulation of HSF1 function in the heat stress response: implications in aging and disease. Annu Rev Biochem 80(1):1089–1115

Balchin D, Hayer-Hartl M, Hartl FU (2016) In vivo aspects of protein folding and quality control. Science 353(6294):aac4354

Bargelloni L, Babbucci M, Ferraresso S, Papetti C, Vitulo N, Carraro R et al. (2019) Draft genome assembly and transcriptome data of the icefish *Chionodraco myersi* reveal the key role of mitochondria for a life without hemoglobin at subzero temperatures. Commun Biol 2:443

Barnes DKA, Fuentes V, Clarke A, Schloss IR, Wallace MI (2006) Spatial and temporal variation in shallow seawater temperatures around Antarctica. Deep Sea Res II 53(8):853–865

Beers JM, Jayasundara N (2015) Antarctic notothenioid fish: what are the future consequences of 'losses' and 'gains' acquired during long-term evolution at cold and stable temperatures? J Exp Biol 218(12):1834

Belyayev A (2014) Bursts of transposable elements as an evolutionary driving force. J Evol Biol 27(12):2573–2584

Berthelot C, Clarke J, Desvignes T, Detrich HW,III, Flicek P, Peck LS et al. (2019) Adaptation of Proteins to the Cold in Antarctic Fish: A Role for Methionine? Genome Biol Evol 11(1):220–231. https://doi.org/10.1093/gbe/evy262

Bilyk KT, DeVries AL (2011) Heat tolerance and its plasticity in Antarctic fishes. Comp Biochem Physiol A Mol Integr Physiol 158(4):382–390

Bilyk KT, Vargas-Chacoff L, Cheng CHC (2018) Evolution in chronic cold: varied loss of cellular response to heat in Antarctic notothenioid fish. BMC Evol Biol 18(1):143

Bogan S, Place SP (2019) Accelerated evolution at chaperone promoters among Antarctic notothenioid fishes. BMC Evol Biol 19 (1):205

Brehme M, Voisine C, Rolland T, Wachi S, Soper James H, Zhu Y et al. (2014) A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. Cell Rep. 9(3):1135–1150

Brunquell J, Morris S, Lu Y, Cheng F, Westerheide SD (2016) The genome-wide role of HSF-1 in the regulation of gene expression in Caenorhabditis elegans. BMC Genomics 17(1):559

Buckley BA, Somero GN (2009) cDNA microarray analysis reveals the capacity of the cold-adapted Antarctic fish *Trematomus bernacchii* to alter gene expression in response to heat stress. Polar Biol 32(3):403–415

Capy P, Gasperi G, Biémont C, Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? Heredity 85:101

Cappucci U, Noro F, Casale AM, Fanti L, Berloco M, Alagia AA et al. (2019) The Hsp70 chaperone is a major player in stress-induced transposable element activation. PNAS 116(36):17943–17950

Chen L, Lu Y, Li W, Ren Y, Yu M, Jiang S et al. (2019) The genomic basis for colonizing the freezing Southern Ocean revealed by Antarctic toothfish and Patagonian robalo genomes. GigaScience 8(4):giz016

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34(17):i884–i890

Cunniff NF, Morgan WD (1993) Analysis of heat shock element recognition by saturation mutagenesis of the human HSP70.1 gene promoter. J Biol Chem 268(11):8317–8324
Cziko PA, DeVries AL, Evans CW, Cheng C-HC (2014) Antifreeze protein-induced superheating of ice inside Antarctic notothenioid fishes inhibits melting during summer warming. P Nat Acad Sci 111(40):14583
Dai C, Sampson SB (2016) HSF1: Guardian of proteostasis in cancer. Trends Cell Biol 26(1):17–28
DeVries AL, Cheng CHC (2005) Antifreeze proteins and organismal freezing avoidance in polar fishes. Fish Physiol 22:155–201
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1):15–21
Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7(1):214
Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 32(5):1792–1797
Finka A, Goloubinoff P (2013) Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. Cell Stress Chaperon- 18(5):591–605
García-Prat L, Martínez-Vicente M, Perdiguero E, Ortet L, Rodríguez-Ubreva J, Rebollo E et al. (2016) Autophagy maintains stemness by preventing senescence. Nature 529:37
Gomez-Pastor R, Burchfiel E, Thiele D (2018) Regulation of heat shock transcription factors and their roles in physiology and disease. Nat Rev Mol Cell Biol 19:4–19
Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29(7):644–652
Hietakangas V, Anckar J, Blomste HA, Fujimoto M, Palvimo JJ, Nakai A et al. (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. Proc Nat Acad Sci 103(1):45–50
Hofmann GE, Buckley BA, Airaksinen S,EKJ, Somero GN (2000) Heat-shock protein expression is absent in the Antarctic fish Trematomus bernacchii (family Nototheniidae). J Exp Biol 203:2331–2339
Horváth V, Merenciano M, González J (2017) Revisiting the relationship between transposable elements and the eukaryotic stress response. Trends Gen 33(11):832–841
Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17(8):754–755
Hunt BM, Hoefling K, Cheng C-HC (2003) Annual warming episodes in seawater temperatures in McMurdo Sound in relationship to endogenous ice in notothenioid fish. Antarct Sci 15(3):333–338
Huth TJ, Place SP (2016) Transcriptome wide analyses reveal a sustained cellular stress response in the gill tissue of *Trematomus bernacchii* after acclimation to multiple stressors. BMC Genomics 17:127
Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28 (23):3150–3152
Jaiswal H, Conz C, Otto H, Wölfle T, Fitzke E, Mayer MP et al. (2011) The chaperone network connected to human ribosome-associated complex. Mol Cell Biol 31(6):1160
Joshi S, Wang T, Araujo TLS, Sharma S, Brodsky JL, Chiosis G (2018) Adapting to stress — chaperome networks in cancer. Nat Rev Cancer 18(9):562–575
Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R et al. (2017) JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucl Acids Res 4;46(D1):D260–D266
Kim B-M, Amores A, Kang S, Ahn D-H, Kim J-H, Kim I-C et al. (2019) Antarctic blackfin icefish genome reveals adaptations to extreme environments. Nat Ecol Evol 3(3):469–478
Klaips CL, Jayaraj GG, Hartl FU (2018) Pathways of cellular proteostasis in aging and disease. J Cell Biol 217(1):51
Kline MP, Morimoto RI (1997) Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. Mol Cell Biol 17(4):2107
Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357
Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinforma 12(1):323
Murphy ME (2013) The HSP70 family and cancer. Carcinogenesis 34 (6):1181–1188
Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL (2012) Detecting individual sites subject to episodic diversifying selection. PLOS Genet 8(7):e1002764
Otto H, Conz C, Maier P, Wölfle T, Suzuki CK, Jenö P et al. (2005) The chaperones MPP11 and Hsp70L1 form the mammalian ribosome-associated complex. P Nat Acad Sci 102 (29):10064–10069
Peck LS (2016) A Cold limit to adaptation in the sea. Trends Ecol Evol 31(1):13–26
Place SP, Hofmann GE (2004) Constitutive expression of a stress-inducible heat shock protein gene, hsp70, in phylogenetically distant Antarctic fish. Polar Biol 28(4):261–267
Place SP, Zippay ML, Hofmann GE (2004) Constitutive roles for inducible genes: evidence for the alteration in expression of the inducible hsp70 gene in Antarctic notothenioid fishes. Am J Physiol Reg Int Comp Physiol 287(2):R429–R436
Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21(5):676–679
Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25(7):1253–1256
Preissler S, Deuerling E (2012) Ribosome-associated chaperones as key players in proteostasis. Trends Biochem Sci 37(7):274–283
Richter K, Haslbeck M, Buchner J (2010) The heat shock response: life on the verge of death. Mol Cell 40(2):253–266
Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucl Acids Res 43(7): e47–e47
Schmidt M, Finley D (2014) Regulation of proteasome activity in health and disease. Biochimica et Biophysica Acta 1843 (1):13–25
Shin SC, Ahn DH, Kim SJ, Pyo CW, Lee H, Kim M-K et al. (2014) The genome sequence of the Antarctic bullhead notothen reveals evolutionary adaptations to a cold environment. Genome Biol 15 (9):468
Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL (2015) Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Mol Biol Evol 32(5):1342–1353
Solís Eric J, Pandey Jai P, Zheng X, Jin Dexter X, Gupta Piyush B, Airoldi Edoardo M et al. (2016) Defining the essential function of yeast HSF1 reveals a compact transcriptional Program for maintaining eukaryotic proteostasis. Mol Cell 63(1):60–71
Somero GN, DeVries AL (1967) Temperature tolerance of some Antarctic fishes. Science 156(3772):257–258
Song W, Li L, Huang H, Jiang K, Zhang F, Wang L et al. (2018) Tissue-based transcriptomics of *Chionodraco hamatus*: sequencing, de novo assembly, annotation and marker discovery. J Fish Biol 94(2):251–260.
Thorne MA, Burns G, Fraser KP, Hillyard G, Clark MS (2010) Transcription profiling of acute temperature stress in the Antarctic plunderfish *Harpagifer antarcticus*. Mar Genomics 3(1):35–44
Todgham AE, Hoaglund EA, Hofmann GE (2007) Is cold the new hot? Elevated ubiquitin-conjugated protein levels in tissues of Antarctic fish as evidence for cold-denaturation of proteins in vivo. J Comp Physiol B 177(8):857–866

Tovkach A, Ryan PR, Richardson AE, Lewis DC, Rathjen TM, Ramesh S et al. (2013) Transposon-mediated alteration of TaMATE1B expression in wheat confers constitutive citrate efflux from root apices. Plant Physiol 161(2):880

Vihervaara A, Sistonen L (2014) HSF1 at a glance. J Cell Sci 127 (2):261

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL (2018) Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Mol Biol Evol 35(3):773–777

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K (2015) RELAX: detecting relaxed selection in a phylogenetic framework. Mol biol evol 32(3):820–832

Williamson VM, Young ET, Ciriacy M (1981) Transposable elements associated with constitutive expression of yeast alcohol dehydrogenase II. Cell 23(2):605–614