# Deep transcriptome analysis of the heat shock response in an Atlantic sturgeon (*Acipenser oxyrinchus*) cell line

Elena Santidrián Yebra-Pimentel[a,b*], Marina Gebert[c], Hans J. Jansen[a], Susanne A. Jong-Raadsen[a] and Ron P.H. Dirks[a]

a - ZF-screens B.V., 2333CH Leiden, The Netherlands.

b - Department of Basic Sciences and Aquatic Medicine, Norwegian University of Life Sciences, 0454 Oslo, Norway.

c - Working Group Aquatic Cell Technology and Aquaculture, Fraunhofer Research Institution for Marine Biotechnology and Cell Technology, 23562 Lübeck, Germany.

* - Corresponding author at: ZF-screens B.V., 2333CH Leiden, The Netherlands.

E-mail address: elena@zfscreens.com

## Abstract

Despite efforts to restore Atlantic sturgeon in European rivers, aquaculture techniques result in animals with high post-release mortality due to, among other reasons, their low tolerance to increasing water temperature. Marker genes to monitor heat stress are needed in order to identify heat-resistant fish. Therefore, an Atlantic sturgeon cell line was exposed to different heat shock protocols (30ºC and 35ºC) and differences in gene expression were investigated. In total 3,020 contigs (~1.5%) were differentially expressed. As the core of the upregulated contigs corresponded to heat shock proteins (HSP), the heat shock factor (HSF) and the HSP gene families were annotated in Atlantic sturgeon and mapped via Illumina RNA sequencing to identify heat-inducible family members. Up to 6 *hsf* and 76 *hsp* genes were identified in the Atlantic sturgeon transcriptome resources, 16 of which were significantly responsive to the applied heat shock. The previously studied *hspa1* (*hsp70*) gene was only significantly upregulated at the highest heat shock (35ºC), while a set of 5 genes (*hspc1, hsph3a*, *hspb1b*, *hspb11a,* and *hspb11b)* was upregulated at all conditions. Although the *hspc1 (hsp90a)* gene was previously used as heat shock-marker in sturgeons, we found that *hspb11a* is the most heat-inducible gene, with up to 3,296-fold higher expression in the treated cells, constituting the candidate gene markers for *in vivo* trials.

**Key words:** *Acipenser oxyrinchus*; RNAseq; temperature; heat shock proteins (HSP); transcriptome sequencing.

## 1. Introduction

Sturgeons are distributed in the northern hemisphere and are mostly anadromous, performing upriver migrations to spawn. They have been aquacultured for the last three decades due to the commercial value of their caviar [1], however wild populations are in severe decline [2,3] and currently the Atlantic sturgeon (*Acipenser oxyrinchus*, Mitchill, 1815) is extirpated from Europe [4].

In 1997, Germany and Poland started a project aiming to restore a self-sustaining Atlantic sturgeon population, derived from a Canadian broodstock, in the Baltic Sea. Rearing techniques have been improved in order to build an *ex-situ* broodstock locally. In addition, the Oder and Vistula river drainages, where more than 120,000 individuals have been released since 2005 [5], have been evaluated in search of possible spawning grounds [6]. The success of a restoration program depends both on politics (e.g., habitat preservation policies and regional and international cooperation) [7, 8] and on biology, (e.g. development of cultivation techniques adapted for restocking purposes). Aquacultured sturgeons are raised in stimulus-deprived tanks at constant temperature and photoperiod [9]; however, their distribution, abundance, and activity are determined by temperature, photoperiod and salinity fluctuations along the migration route [10,11]. An increase in water temperature, both locally due to the discharge of cooling waters or globally due to climate change, can affect swimming performance, predator avoidance, foraging behavior and shift the geographic distribution of species, especially of ectothermic animals such as fish, and result in local extinctions [12–14]. Thus, thermal adaptation is crucial for post-release survival.

As a result of physiological or chemical stress, proteins misfold and accumulate in the cytoplasm, activating a highly conserved and transient heat shock response (HSR), essential for proteostasis maintenance. HSR results from the rapid activation of the heat shock factor (HSF) by trimerization, hyperphosphorylation and translocation to the nucleus, were its DNA-binding domain (DBD) binds to the heat shock elements (HSE), located at the promoters of heat shock protein (*hsp*) genes, inducing their transcription [15]. Previous studies in *Saccharomyces cerevisiae* showed that heat-responsive gene expression attenuates within 20-40 minutes after induction [16]. Most vertebrates have multiple *hsf* genes: mammals have 4 [17], while the Spotted gar (*Lepisosteus oculatus*), sturgeon's closest living relative whose genome is sequenced, has 6 (*hsf1*, *hsf2*, *hsf3*, *hsf4*, *hsf5,* and *hsfy*).

HSPs, one of the largest and most conserved protein families, were first discovered in *Drosophila* [18] and have since then been described in all living organisms, from yeast [19] to mammals [20,21] and fish [22-24]. Its nomenclature refers to their molecular weights, however, since the human genome annotation the names have become confusing, with up to 10 different names for the same gene product. In 2009 a new nomenclature was proposed [25] based on the gene symbols that have been assigned to the HUGO Gene nomenclature committee (HGNC). This nomenclature was also used in this study, and classifies HSPs into 9 subfamilies: HspA (former Hsp70), HspH (former Hsp110), HspB (small Hsp), HspC (Hsp90), HspD (Hsp60), HspE (Hsp10) and the J domain-containing subfamilies (Hsp40) DnajA, DnajB and DnajC.

The HSP family consists of both constitutively expressed and inducible members, some of which are responsive to heat and act by binding to the exposed hydrophobic amino acid residues of the misfolded proteins preventing undesired molecular interactions. The capacity of HSPs to bind misfolded proteins is regulated through allosteric mechanisms via ATP binding and hydrolysis, with the exception of the HspB subfamily which is ATP-independent.

RNA sequencing (RNAseq) is a high-throughput tool to quantify transcriptomic changes and, unlike qPCR, is unbiased, allowing gene discovery and quantification with good correlation with qPCR when both procedures are performed well [26,27]. RNAseq has previously been used in sturgeons to study sex-related genes, developmental genes or response to infection [28-33], but never to study the heat shock response. Many studies used RNAseq to address this in teleosts [34-36], while for sturgeons only qPCR or protein analysis approaches were used [37-42]. Still, the focus has only been on *hspa1 (hsp70)* and hspc1 (*hps90)*, as the HSP family is not fully annotated in sturgeons, which are also lacking a reference genome.

Recently, the AOXlar7y cell line (***A****cipenser* ***o****xyrinchus* **lar**vae n°**7**, tr**y**psin-digestion) was established by whole larvae trypsin digestion [43], providing the possibility of reducing the use of *in vivo* sturgeon experiments. Still, the presence of stem cells in the culture was not confirmed so, characterization of the cell type needs to be assessed.

The objective of this study was to make an inventory of all the HSPs present in the Atlantic sturgeon transcriptome resources and identify the most heat shock-inducible family members. To this end, we annotated the complete

126 HSP family using transcriptomic resources from Atlantic sturgeon cell line,
127 multiple Atlantic sturgeon organs and a publicly available dataset [31] and
128 subsequently evaluated their response to heat. Although *in vivo* validation
129 during temperature-challenge trials is needed, the *in vitro* experiment provides
130 a general perspective of the HSR and candidate gene markers for the
131 selection of thermotolerant individuals with better restoration fitness.
132
133 **2. Materials and methods**
134 2.1. Cell line samples
135 The AOXlar7y sturgeon cell line [43] was used for the heat shock
136 experiments. Cells from passage 22 (P22) were thawed and seeded in a 25
137 $cm^2$ cell culture flask (Corning Life Sciences, Tewksbury, MA, USA) at 25°C,
138 the optimal growing temperature (doubling time 110h), in Leibovitz-15 medium
139 supplemented with 15% FCS (fetal calf serum), 100U/mL penicillin and
140 0.1mg/mL streptomycin.
141 Prior to the final experiment, the cells were propagated by trypsinization,
142 homogenization and subculturing at a 1:3 ratio every 4 days, up to P29. To
143 determine the heat tolerance of the AOXlar7y sturgeon cell line, the cells were
144 exposed to 28°C, 30°C, 33°C or 35°C for 1, 2, 3 or 4 hours and survival was
145 inspected under the microscope after 4, 8 and 24 hours of recovery. Based on
146 this pilot experiment, the final experiment was performed as depicted in
147 Figure 1: 24 individual flasks were seeded at 25°C, and after 24 hours the
148 medium was renewed at either 25°C (n=6), 30°C (n=9) or 35°C (n=9). The
149 flasks were then transferred to different incubators set at 25°C, 30°C and
150 35°C respectively and kept there for 1 hour. After the heat shock all the flasks

were transferred back to 25°C to recover. After 4, 8 and 24 hours of recovery, cells were lysed in QIAzol lysis reagent (Qiagen GmbH, Hilden, Germany) and stored at - 80° until further processing.

2.2 Atlantic sturgeon samples

An aquaculture-reared immature female (7,700g, 98cm) was provided by Fischzucht Rhönforelle GmbH in Gersfeld (Germany) on July 3rd, 2015. The specimen was euthanized, and tissue samples were taken in agreement with standardized fish processing methods at a licensed and registered processing unit (Fischzucht Rhönforelle GmbH). A total of 21 samples were taken through dissection, from caudal to rostral: caudal fin, dorsal fin, pelvic fin, skin, muscle, intestine (posterior), kidney, spleen, swim bladder, gonad (2 samples), intestine (anterior), stomach, liver, gallbladder, heart, gill, brain (anterior), brain (posterior), barbel and eye. Samples were preserved in RNAlater (Qiagen GmbH, Hilden, Germany) and subsequently stored at -80°C until further processing.

2.3. RNA extraction, Illumina library preparation, and RNA sequencing

RNA was extracted from the 21 juvenile Atlantic sturgeon tissue samples and from the 24 AOXlar7y cell line samples using the Qiagen miRNeasy Mini kit according to the manufacturer's instructions (Qiagen GmbH, Hilden, Germany). RNA concentration and integrity were analyzed with a Bioanalyzer 2100 total RNA Nano series II chip (Agilent, Santa Clara, USA). RNA libraries were prepared from 500 ng total RNA, using the Illumina TruSeq Stranded mRNA Sample Preparation Kit according to the

manufacturer's instructions (Illumina Inc., San Diego, USA) and the resulting libraries were evaluated with a Bioanalyzer 2100 DNA 1000 series II chip (Agilent, Santa Clara, USA).

All the libraries were sequenced using an Illumina HiSeq2500 instrument. Tissue libraries were sequenced as paired-end 2x151 nucleotides (nt) reads up to a minimum of ~ 20 million reads for each tissue sample, while AOXlar7y libraries were sequenced as 1x51nt single-reads up to a minimum of ~10 million reads for each experimental condition.

*2.4. De novo* transcriptome assemblies and annotation

Three reference transcriptomes were produced in the present study. Firstly, all combined juvenile sturgeon organ reads were *de novo* assembled into cDNA contigs using the De Bruijn graph-based *de novo* assembler implemented in the CLC Genomics Workbench version 4.4.1 (CLC bio, Aarhus, Denmark), resulting in an Organ assembly. Secondly, the same software was used to produce a Cell assembly with the AOXlar7y sequence reads only. Finally, all sequence reads were combined into a Cell-Organ-Embryo (COE) transcriptome, including the reads from the AOXlar7y, the juvenile sturgeon organs and a set of embryonic reads previously published by others [31].

The quality of the *de novo* assemblies was assessed by the assembly contiguity (contig N50) and the percentage of AOXlar7y mapped reads. In order to link the best assembly to zebrafish Ensembl protein identifiers, BLASTX 2.2.31+ [44] similarity searches were conducted locally against the UniProt zebrafish using an E-value cut off of 1E-5.

201

202 2.5. Gene expression and gene ontology (GO) analysis

203 For the gene expression analysis, Bowtie2 (version 2.2.5) [45] was
204 used to align Illumina reads from the 24 experimental samples against the 3
205 *de novo* assembled reference transcriptomes (Cell, Organ and COE contigs)
206 and against the embryonic transcriptome available at NCBI [31]. The
207 assembly with the highest mapping percentage was used as a reference for
208 downstream analysis. The resulting files were filtered using SAMtools (version
209 1.2.) [46] to exclude secondary aligned reads. Then, the aligned reads were
210 counted from the SAM alignment files using Python package HTSeq (version
211 0.5.3p9) [47], and the resulting TSV files were used for identification of
212 differentially expressed contigs (DECs) using the Bioconductor package
213 DESeq (version 1.30.0) [48] in R software (version 3.3.0). The top 100
214 expressed contigs in the unstressed cells were investigated to characterize
215 the AOXlar7y cell line. Each experimental condition (30°C and 35°C) and
216 timepoint (4, 8 and 24 hours) was compared to the control group, resulting in
217 6 DESeq pairwise comparisons. P-values were adjusted for multiple testing
218 with the Benjamini-Hochberg procedure, which controls the false discovery
219 rate (FDR). Contigs with an adjusted p-value (padj) < 0.05 were considered
220 differentially expressed at each pairwise comparison between different
221 treatments and timepoints. The set of zebrafish protein identifiers linked to
222 these contigs was investigated in order to elucidate the broad transcriptomic
223 changes. GO enrichment analysis was performed using the Fisher's Exact
224 test with Bonferroni correction implemented by the PANTHER
225 Overrepresentation test (released on 2017-12-05) using the GO Ontology

database (released on 2018-07-03) [49]. As recommended by the Gene Ontology Consortium [50, 51], a custom reference list containing all the top zebrafish Ensembl gene identifiers linked to the expressed COE contigs was used for the analysis. GO terms with a p-value < 0.05 were considered overrepresented in each dataset.

2.6. Annotation and expression of Hsp and Hsf genes

Spotted gar, zebrafish and human HSP and HSF protein sequences [25] were retrieved from NCBI and used as queries to find the corresponding sturgeon orthologues. Proteins were blasted against the translated COE transcriptome assembly using the CLC Main Workbench (version 7.7.3.). When full open reading frames (ORF) were not found in the COE assembly, the Organ, Cell or embryo assembly [31] were used. The human HSP protein nomenclature according to Kampinga et al. [25] was used. After the annotation of all Atlantic sturgeon *hsp* and *hsf* genes, the AOXlar7y Illumina reads were mapped to the annotated genes and those with more than 5 mismatches were filtered out in order to improve stringency. DESeq analysis was performed, following the aforementioned pipeline, in order to find candidate *hsp* markers for *in vivo* trials. Significantly upregulated transcripts (padj>0.05) with at least 3 fold change (FC≥3) between the control and at least 5 experimental conditions were considered good markers in this experiment.

The AOXlar7y and organ sequencing reads were deposited at the NCBI Short Read Archive (SRA) database under the accession numbers SRP161542 and SRP161601 respectively. The Cell, Organ, and Cell-Organ-

Embryo (COE) transcriptome shotgun assembly projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers GGWJ00000000, GGZX00000000 and GGZT00000000 respectively. The versions described in this paper are the first versions: GGWJ01000000, GGZX01000000, and GGZT01000000, respectively. The Atlantic sturgeon *hsp* and *hsf* ORFs were submitted to BankIT under the accession numbers MH777912-MH777987 and MH917287-MH917292 respectively.

## 3. Results

### 3.1. Heat shock treatment and RNA sequencing

The pilot experiment has shown that cells survived after being exposed to 28°C, 30°C or 33°C for up to 4 hours or to 35°C for 1 hour; however, acute mortality was found in cells exposed to 35°C for 2 hours or longer (data not shown). Based on this, the final experiment was performed applying a mild or a severe heat shock at either 30ºC or 35ºC for 1 hour, and cells were harvested at 4, 8 and 24 hours after heat shock (Figure 1). Subsequently, high-quality RNA (average RIN value of 9.1) was isolated from all cell samples and from multiple organs of an aquacultured juvenile sturgeon specimen, and Illumina libraries were prepared and sequenced, resulting in altogether more than 1 billion reads (Table S1 in Supplementary data).

### 3.2. Heat shock-responsive genes in Atlantic sturgeon cell line AOXlar7y

To select the best reference for alignment of the sequencing reads, 3 Atlantic sturgeon *de novo* transcriptome assemblies were performed and evaluated. The set of reads derived from the AOXlar7y cell line was

assembled to a 34.80 Mb Cell assembly containing 53,624 contigs with an N50 of 1.08 Kb. In parallel, reads from the Atlantic sturgeon organs were assembled to a 342.28 Mb Organ assembly containing 641,485 contigs with an N50 of 0.60 Kb. Finally, all sequencing reads were combined with a previously published set of ~ 380 million reads derived from Atlantic sturgeon embryos (SRA Accession number SRP069853) [31] and used in a *de novo* assembly that resulted in a 254 Mb Cell-Organ-Embryo (COE) assembly containing 203,131 contigs with an N50 of 1.87 Kb (Table 1).

Reads from the 24 AOXlar7y samples were aligned against the three assembled transcriptomes and the COE reference, which gave the highest overall mapping percentage (89.6%), was used for the downstream analysis (Table 1). In total 86,021 contigs (~ 42.3%) could be linked to zebrafish proteins (E-value <1E-5), corresponding to 23,436 unique zebrafish proteins and 17,319 genes (Table S2 in Supplementary data). The read counts and contig lengths were used to calculate the corresponding RPKM values (Table S3 in Supplementary data). A total of 168,739 contigs (83.07%) showed expression (RPKM value > 0) in at least one of the cell samples, and 27,425 (13.50%) showed expression in all of them.

Since it was hitherto unknown what tissue type was represented by the embryonic AOXlar7y cell line, the zebrafish proteins linked to the 100 most highly expressed contigs (highest RPKM mean) were examined. In total, 34 contigs corresponded to ribosomal proteins, 5 to keratins, 5 to actin genes, 5 to tubulin and other genes were also found. Overall, keratin 4 was the most expressed contig (Table S4 in Supplementary data).

300 Overall, DESeq analysis showed that 3,020 unique contigs (~ 1.5%)
301 were differentially expressed in at least 1 of the treatments compared with the
302 untreated cells, of which 2,302 were upregulated, 714 downregulated and 4
303 were up or downregulated depending on the experimental condition (Table S5
304 in Supplementary data).
305 At 4 and 8 hours after the 30°C heat shock, 79 and 91 contigs were
306 upregulated and 23 and 64 were downregulated respectively, while after 24 h
307 only 12 were downregulated and none upregulated. After the 35°C heat
308 shock, respectively 1,338, 1,081 and 1,143 contigs were upregulated and
309 411, 123 and 198 were downregulated (Figure 2).
310 Venn diagrams were used to determine overlaps between the sets of
311 heat shock-responsive contigs. After the 30°C heat shock, 140 unique contigs
312 (0.07%) were upregulated, 30 of which at both 4 and 8 hours after heat shock
313 (Figure 3a), while only 87 unique contigs (0.04%) were downregulated (Figure
314 3b). After the 35°C heat shock 2,257 unique contigs (1.11%) were
315 upregulated, 398 of which were upregulated at all 3 timepoints (Figure 3c),
316 and 664 unique contigs (0.33%) were downregulated, of which only 9 were
317 downregulated at all timepoints (Figure 3d).
318 Considering that one of our goals was to provide a list of consistently
319 upregulated genes that need to be validated as markers *in vivo*, we examined
320 the core set of 27 upregulated contigs (Figure 3e). Up to 18 corresponded to 5
321 heat shock proteins (*hspb1, hspb11a, hsph3a, hsph2*, and *hspc1*), while 5
322 corresponded to 4 different proteins: clusterin (*clu*), growth factor receptor
323 bound protein 2b (*grb2b*), atrial natriuretic peptide receptor 2-like (*npr2*) and
324 coiled-coil domain-containing protein 17 (*ccdc17*) (Tables 2 and S7 in

325 Supplementary data). The remaining 4 contigs did not output any blast hit or
326 were uncharacterized proteins.

327 To link the differentially expressed contigs (DECs) to GO terms, the
328 3,020 unique contigs were blasted (BLASTx) against zebrafish proteins (E-
329 value <1E-2), which resulted in 1,383 hits (45.8%) corresponding to 1,068
330 unique proteins and 1,017 zebrafish genes (Table S6 in Supplementary data).
331 GO enrichment analysis for biological processes (BP) and molecular functions
332 (MF) was performed (Table S8 in Supplementary data). With respect to
333 upregulated genes after exposure to 30°C, cellular processes involved in the
334 response to stimulus (e.g. 'cellular response to heat', 'cellular response to
335 unfolded protein', 'response to temperature stimulus', 'response to unfolded
336 protein'), protein folding (e.g. ' chaperone-mediated protein folding', 'protein
337 refolding') and protein binding (e.g. 'heat shock protein binding', 'protein
338 folding', 'misfolded protein binding') were overrepresented (Figure 4),
339 however, no overrepresented terms were found after 8 or 24 hours. After the
340 35ºC heat shock overrepresented terms were found at all the timepoints: while
341 GO terms related with cellular response to stimulus and stress (e.g. 'cell
342 chemotaxis', 'leukocyte chemotaxis', 'inflammatory response') and cell
343 migration (e.g. 'granulocyte migration', leukocyte migration', 'neutrophil
344 migration') are overrepresented after 4 and 8 hours (Figure 5a and 5b), after
345 24 hours (Figure 5c) many more GO terms are overrepresented, being the
346 'response to stimulus' the predominant term (122 genes).

347 With respect to downregulated genes, overrepresentation of 'cell cycle'
348 and 'mitotic cell cycle' GO terms was found only at 4 hours after 35ºC heat
349 shock.

### 3.3. Annotation and expression of Hsp and Hsf family members in Atlantic sturgeon

Protein sequences of all known human, zebrafish and spotted gar HSP and HSF family members were retrieved from NCBI and used as queries to search for the corresponding Atlantic sturgeon orthologues in the COE, Organ, Cell or Embryo transcriptomes (Table S9 in Supplementary data).

Spotted gar HSF proteins were used to identify 6 *hsf* genes in sturgeon, which lacked the *hsf3* gene and had 2 *hsfy* genes (*hsfy1* and *hsfy2*). None of the *hsf* genes was differentially expressed at any time point after the mild and severe heat shock.

The former Hsp70 family is divided into 2 sub-families: HspA and HspA-related HspH (Hsp110). The HspA family contains 13, 8 and 7 members in humans, zebrafish and spotted gar, respectively. Nine members were found in the Atlantic sturgeon transcriptome: *hspa1, hspa5, hspa8, hspa9, hspa12a, hspa12b, hspa13, hspa14a, hspa14b*. The HspH family contains 4 members in both human and spotted gar, plus 1 duplicated gene in zebrafish. The *hsph1* member could not be retrieved in any of the available sturgeon assemblies and the *hsph3* member was found duplicated.

Humans, spotted gar and zebrafish contain the *hspb1*, *hspb2, hspb3, hspb4* (αA-crystallin), *hspb5* (αB-crystallin)*, hspb7, hspb8* and *hspb9* members of the small HSP family, whereas *hspb6* is missing in spotted gar, *hspb10* is only present in humans, and *hspb11*, *hspb12* and *hpsb15* are only present in zebrafish. With the exception of *hspb3*, orthologues of all spotted gar *hspb* genes could be found in the sturgeon transcriptome, including 2

375 *hspb1* genes (*hspb1a* and *hpsb1b*). The *hspb11* member, absent in human
376 and spotted gar, could also be retrieved duplicated in the sturgeon
377 transcriptome (*hspb11a* and *hspb11b*).

378 The HspC family (former *hsp90*) has 5 members in humans and 4 in
379 zebrafish and spotted gar, which lacks the *hspc2* member. The same fish
380 members were found in the sturgeon transcriptome, with an additional *hspc3*
381 duplication (*hspc3a* and *hspc3b*). The HspD and HspE families each contain
382 only 1 single gene in humans, zebrafish and spotted gar, and orthologues of
383 both of them could be found in sturgeon.

384 The new nomenclature for the former Hsp40 family divides it into
385 DnajA, DnajB and DnajC families. We have found 4, 10 and 32 genes for
386 each family, respectively.

387 Upon mapping the AOXlar7y Illumina reads against the annotated
388 sturgeon HSP genes (Table S10 in Supplementary data) and performing
389 differential expression analysis using DESeq, we found that 16 out of 76 HSP
390 genes (21.%) were differentially expressed between the control and some of
391 the experimental conditions. No HSPs transcripts were found differentially
392 expressed between the control and 24 hours after the 30ºC heat shock. Within
393 the HspA family, only *hspa1* was differentially expressed after 35ºC heat
394 shock (Figure 6a), being up to ~1000-fold upregulated 4 hours after the 35ºC
395 treatment.

396 Upregulation of 3 out of 4 HspH family members was observed (up to
397 ~9-fold): *hsph2* and *hsph3a* were significantly upregulated at all the
398 conditions, however, the *hsph3b* paralog was only slightly upregulated at after
399 the 30° treatment (Figure 6a).

In addition, 5 out of 10 HspB family members and the single HspC family member *hspc1* were upregulated after heat shock: while *hspb8* was only significantly upregulated at 4 hours after the 35°C heat shock and *hspb5* only at 35ºC, *hspb1b, hspb11a, hspb11b* and *hspc1* were consistently upregulated in all the treatments (excluding 24 hours after the 30°C treatment), with the *hspb11a* having the higher expression (as high as ~3296-fold compared to the untreated cells) (Figure 6b). The HspE1 and HspD1 were not differentially expressed at any condition.

Within the DnajA family (Figure 6c), only the *dnaja4* member was upregulated. Four DnajB members were upregulated: *dnajb1* and *dnajb5* were only upregulated after the 35°C heat shock, *dnajb4* was also upregulated at 4 hours after the 30°C heat shock, and *dnajb2* was only upregulated after the 30°C heat shock.

Although the DnajC family is the most extensive Dnaj subfamily, it contained only 1 heat shock inducible gene under our conditions, *dnajc3*, which was upregulated only at 24 hours after the 35ºC heat shock.

## 4. Discussion

The AOXlar7y cell line is easy to maintain and propagate, providing an excellent tool for examining the effects of different stressors. Healthy AOXlar7y cells have cubic to fibroblast-like morphology, and the RNAseq analysis in this study shows that keratin 4 is the most abundantly expressed mRNA in this cell line, suggesting that they are epithelial cells [52]. Although both the cell line and the donor embryo have been reported to be more tolerant to cold than heat [43, 53] the optimal temperature for cells is higher

than for the entire animal (25°C and 18°C-23°C, respectively). After establishing the temperature tolerance limits, cells were exposed to a mild (30°C) and severe (35°C) non-lethal heat shock for 1 hour and, after 4, 8 and 24 hours of recovery, RNA was isolated and sequenced. By RNAseq we could get a general perspective of the heat-inducible genes, which may not truly represent the metabolic state of the cells *in vivo* and therefore needs validation, but still provides the sequence of all *hsp* genes and a general perspective of which genes may be more responsive to heat.

The COE reference transcriptome had a much higher alignment rate of cellular reads (89.6%) than the transcriptomes that were assembled from the individual data sets, and was therefore selected for identification of DECs. This reference provides a considerable number of sturgeon protein sequences that contribute to sturgeon research. The set of 3,020 DECs between the control and at least one of the treatments was blasted against zebrafish proteins, the most related species for which GO analysis is available. A total of 1,383 DECs could be assigned to zebrafish proteins, while the remaining contigs were noncoding, sturgeon-specific, missing in zebrafish or lacked sufficient homology with their zebrafish orthologues.

The DESeq results showed a very intense response at 35°C, initially involving energy-consuming gene upregulation and later also gene downregulation, which is an energy-saving mechanism to direct energy towards the repair of damaged molecules [36]. After a mild 30°C heat shock there were very few transcriptional changes at early timepoints, and even zero after 24 hours of recovery. The narrow temperature tolerance range between 30°C and 35°C where sturgeon cells switch from a mild to a severe heat

450 shock response is in agreement with existing studies in green sturgeon
451 (Acipe*nser medirostris*) [53] and Kaluga (*Huso dauricus*) [39].

452 The GO analysis showed a heat shock response at 35°C which
453 increased with the recovery time. At 30°C the response was more specific
454 (overrepresentation of less GO terms); however, 24 hours after heat shock
455 the cells showed the same transcriptional activity as the untreated cells,
456 suggesting that the cells were already fully recovered after this mild heat
457 shock (30ºC). The bulk of the 27 contigs that were upregulated at all
458 conditions (excluding 30°C after 24 hours) corresponded with HSP genes,
459 and included 4 other genes: *clu*, *grb2b, npr2 and ccdc17*. Clusterin is a
460 molecular chaperone [54, 55] and, similarly to *npr2*, is involved in cell survival
461 after apoptosis induction [56-58]. Coiled-coil domains are involved in the
462 regulation of gene expression, but the specific function of *ccdc17* has not
463 been studied in detail [59]. As both the DESeq and GO results point to the
464 HSP genes as central players in the heat shock response, we have annotated
465 the entire family and analyzed their response to heat.

466 In total 76 Hsp and 6 Hsf genes could be retrieved in the Atlantic
467 sturgeon transcriptomes, including 4 partial sequences and 78 full ORFs: 33
468 (40.24%) in the COE assembly and 45 (54.87%) in either the embryo, cell or
469 organ transcriptomes.

470 In our experiment, only 16 out of 76 Hsp genes were heat-inducible, for
471 which the coefficient of variation (CV) was lower than 30% in most of the
472 cases (89.58%). HspB was found to be the most heat responsive family with
473 less than 30% CV except for *hspb5*, which had high CV at all the timepoints
474 after the 30ºC heat shock. The HspB family is ATP-independent and

475 characterized by the presence of a conserved α-crystallin domain [60]. Their
476 expression has been shown to enhance the post-stimulus survival of
477 mammalian cells [61] and, besides the molecular chaperone activity, some
478 members have additional cellular functions: *hspb1* (*hsp27*) and *hspb5*,
479 induced by heat shock in zebrafish [62], are known to inhibit apoptosis [63-
480 65], while *hspb8* has kinase activity [66]. Unfortunately, the human *hspb11*
481 gene proposed by Kampinga [25], previously known as intraflagellar transport
482 protein 24 (*ift25*), lacks the α-crystallin domain and its nomenclature hasn't
483 been approved [67]. In fact *hspb11*, also known as *hsp30*, exists in all
484 vertebrates except mammals [68], and doesn't share any amino acid
485 sequence similarity with the human *ift25*. Since *hspb11* absent in both human
486 and spotted gar genomes, the zebrafish protein sequence was used to
487 retrieve the Atlantic sturgeon orthologue.

488 The few published studies on HSPs in *Acipenserids* were performed in
489 species other than Atlantic sturgeon, focused only on *hspa1* (*hsp70)* and
490 *hspc1* (*hsp90)* and didn't use RNAseq. Using qPCR, *hspa1* and *hspc1* were
491 found expressed in both unstressed and heat-shocked Kaluga juvenile
492 tissues; however, *hspa1* was found to be more inducible by cold than heat
493 [39]. In contrast, *hspa1* showed higher expression after heat than cold stress
494 in both white (*A. transmontanus*) and green (*A. medirostris*) sturgeon larvae
495 [69]. Existing studies found a heat-dependent increase of deformities
496 accompanied by an increase of HspA1 and HspA5 (Hsp78) and a decrease of
497 HspD1 (Hsp60) protein levels [70]. If heat stress is not lethal, the
498 accumulation of HSPs may lead to the tolerance of more severe and
499 otherwise fatal stresses [71,72]. Some studies suggest that heat shock

experienced by the parental fish or gametes could improve larvae thermotolerance, resulting in higher survival rates and lower incidence of deformities linked to high HspD1 and HspC1 protein levels [53]. After heat shock, Hsp levels were higher in a heat-adapted subspecies of doctor fish (*Garra rufa*) than in the non-adapted, indicating that Hsp levels provide thermotolerance [73].

Overall, 5 HSPs are consistently upregulated (FC≥3) after all the treatments (excluding 24 h after the 30ºC heat shock) and are candidate markers for *in vivo* validation: *hsph3a*, *hspb1b*, *hspb11a*, *hspb11b* and *hspc1.* Interestingly, this list excludes *hspa1*, which is only upregulated after severe heat shock in our study. Although *hspc1*, previously found upregulated *in vivo,* is also included in the list included in the list, *hspb11a* has as much as ~3296-fold upregulation compared to the control cells, and is the best candidate marker for *in vivo* validation and trials.

*Hspb11* is an intron-less gene, enabling fast expression without major splicing events [74]. Heat shock was shown to induce accumulation of hspb11 mRNA in Atlantic salmon [75], Chinook salmon (1250-fold) [76], red band trout (200-fold) [77], zebrafish [67], clawed frog [70], heat-tolerant Arctic charr [78] and killfish [79], with the latter showing more upregulation in heat-tolerant southern populations than the northern counterparts.

In conclusion, the AOXlar7y cell line provides the opportunity to reduce *in vivo* experiments on Atlantic sturgeon, an extirpated species in Europe. We have performed RNAseq on heat-shocked cells to get a general perspective of heat inducible genes, and identified and annotated 6 HSF and 76 HSP genes. Only 16 *hsp* transcripts were significantly upregulated after the applied

525 treatment of which 5 were common to all treatments and timepoints, excluding
526 24h after the 30ºC heat shock treatment. These genes had at least a 3-fold
527 increase in expression and one of them, *hspb11a*, had as much as a 3,296-
528 fold increase. These genes are candidate markers for the selection of
529 thermotolerant individuals and should be validated *in vivo*.
530
531 **Funding**
532 This work was supported by the European Union's Horizon 2020 research
533 and innovation program under the Marie Skłodowska-Curie Actions:
534 Innovative Training Network "IMPRESS" [Grant Agreement No. 642893].
535
536 **Acknowledgements**
537 The authors wish to sincerely thank Peter Gross and Vojta Kaspar for the help
538 provided during sturgeon sampling and Elena Sarropoulou for her permission
539 and support providing the Atlantic sturgeon embryonic reads and reading
540 through the manuscript.
541
542 **References**
543 [1] Bronzi P, Rosenthal H, Gessner J. 2011. Global sturgeon aquaculture
544 production: An overview. J Appl Ichthyol, 27. 169-175.
545 https://doi.org.10.1111/j.1439-0426.2011.01757.x
546 [2] Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, et al.
547 2006. Freshwater Biodiversity: Importance, Threats, Status and
548 Conservation Challenges. Biol Rev Camb Philos Soc, 81(2). 163–82.
549 https://doi.org/10.1017/S1464793105006950.

[3] Freyhof J and Brooks E. 2011. European Red List of Freshwater Fishes. Luxembourg: Publications Office of the European Union. https://doi.org/10.2779/85903.

[4] Elvira B, Leal S, Doadrio I, Almodóvar A. 2015. Current Occurrence of the Atlantic Sturgeon *Acipenser Oxyrinchus* in Northern Spain: A New Prospect for Sturgeon Conservation in Western Europe. Plos One, 10(12): e0145728. https://doi.org/10.1371/journal.pone.0145728.

[5] Kolman R, Kapusta A, Duda A, Wiszniewski G. 2011. Review of the Current Status of the Atlantic Sturgeon *Acipenser Oxyrinchus Oxyrinchus* Mitchill 1815, in Poland: Principles, Previous Experience, and Results. J Appl Ichthyol, 27(2). 186–91. https://doi.org/10.1111/j.1439-0426.2011.01680.x.

[6] Gessner J and Bartel R. 2000. Sturgeon Spawning Grounds in the Odra River Tributaries : A First Assessment. Boletin Instituto Espanol de Oceanografia, 16 (1–4). 127–37.

[7] Gessner J. 2000. Reasons for the Decline of *Acipenser Sturio* L., 1758 in Central Europe, and Attempts at Its Restoration. Boletin Instituto Espanol de Oceanografia, 16(1–4). 117–26.

[8] Kirschbaum F, Ludwig A, Hensel E, Wurtz S, Kloas W, et al. 2004. Status of the Project on Protection and Restoration of Atlantic Sturgeon in Germany: Background, Current Situation, and Perspectives. 36-53 in Species differentiation and population identification in the sturgeons Acipenser sturio L. and Acipenser oxyrinchus. J. Gessner and J. Ritterhoff editors, BfN Skripten 101: 122pp.

573 [9] Mohler JW. 2000. Early Culture of the American Atlantic Sturgeon *Acipenser*
574 *oxyrinchus oxyrinchus* Mitchill, 1815 and Preliminary Stocking Trials. Boletin
575 Instituto Español de Oceanografia, 16(1–4). 203–8.
576 [10] Beitinger TL, Bennett WA, McCauley RW. 2000. Temperature Tolerances
577 of North American Freshwater Fishes Exposed to Dynamic Changes in
578 Temperature. Environ Biol Fishes, 58. 237–75.
579 https://doi.org/10.1023/A:1007676325825.
580 [11] Gradil AM, Wright GM, Speare DJ, Wadowska DW, Purcell S, et al. 2014.
581 The Effects of Temperature and Body Size on Immunological Development
582 and Responsiveness in Juvenile Shortnose Sturgeon (*Acipenser*
583 *brevirostrum*). Fish Shellfish Immunol, 40(2). 545–55.
584 https://doi.org/10.1016/j.fsi.2014.07.036.
585 [12] Fields PA, Graham JB, Rosenblatt RH, Somero GN. 1993. Effects of
586 Expected Global Climate Change on Marine Faunas. Trends Ecol Evol,
587 8(10). 361–67. https://doi.org/10.1016/0169-5347(93)90220-J.
588 [13] Paragamian, VL, Beamesderfer RCP, Ireland SC. 2005. Status, population
589 dynamics, and future prospects of the endangered Kootenai River white
590 sturgeon population with and without hatchery intervention. Trans Am Fish
591 Soc, 134. 518-532. https://doi.org/10.1577/T03-011.1
592 [14] Hofmann GE and Todgham AE. 2010. Living in the Now: Physiological
593 Mechanisms to Tolerate a Rapidly Changing Environment. Annu Rev
594 Physiol, 72(1). 127–45. https://doi.org/10.1146/annurev-physiol-021909-
595 135900.
596 [15] Andrews GK, Harding MA, Calvet JP, Adamson ED. 1987. The Heat-
597 Shock Response in Hela-Cells Is Accompanied By Elevated Expression of

598 the C-Fos Protooncogene. Mol Cell Biol, 7(10). 3452–58.
599 https://doi.org/10.1128/MCB.7.10.3452
600 [16] Morano KA, Grant CM, Moye-Rowley WS. 2012. The Response to heat
601 shock and Oxidative stress in Saccharomyces cerevisiae. Genetics, 190(4).
602 1157-1195.
603 [17] Åkerfelt M, Morimoto RII, Sistonen L. 2010. Heat Shock Factors:
604 Integrators of Cell Stress, Development and Lifespan. Nat Rev Mol Cell Biol,
605 11(8). 545–55. https://doi.org/10.1038/nrm2938.
606 [18] Ritossa F. 1962. A New Puffing Pattern Induced by Temperature Shock
607 and DNP in drosophila. Experientia, 18(12). 571–73.
608 https://doi.org/10.1007/BF02172188.
609 [19] Miller MJ, Xuong NH, Geiduschek EP. 1979. A Response of Protein
610 Synthesis to Temperature Shift in the Yeast *Saccharomyces Cerevisiae*.
611 Proc Natl Acad Sci U S A, 76(10). 5222–25.
612 https://doi.org/10.1073/pnas.76.10.5222.
613 [20] Slater A, Cato ACB, Sillar GM, Kioussis J, Burdon RH. 1981. The Pattern
614 of Protein Synthesis Induced by Heat Shock of HeLa Cells. Eur J Biochem,
615 117(2). https://doi.org/10.1111/j.1432-1033.1981.tb06343.x.
616 [21] Wu B, Hunt C, Morimoto R. 1985. Structure and Expression of the Human
617 Gene Encoding Major Heat Shock Protein HSP70. Mol Cell Biol, 5(2). 330–
618 41. https://doi.org/10.1128/MCB.5.2.330.21
619 [22] Krone PH, Lele Z, Sass JB. 1997. Heat Shock Genes and the Heat Shock
620 Response in Zebrafish Embryos. Biochem Cell Biol, 75. 487–97.
621 https://doi.org/10.1139/o97-083.

622 [23] Råbergh CMI, Airaksinen S, Soitamo A, Björklund HV, Johansson T, et al.
623 2000. Tissue-Specific Expression of Zebrafish (*Danio rerio*) Heat Shock
624 Factor 1 MRNAs in Response to Heat Stress. J Exp Biol, 1824. 1817–24.
625 [24] Hori TS, Gamperl AK, Afonso LOB, Johnson SC, Hubert S, et al. 2010.
626 Heat-Shock Responsive Genes Identified and Validated in Atlantic Cod
627 (*Gadus Morhua*) Liver, Head Kidney and Skeletal Muscle Using Genomic
628 Techniques. BMC Genomics, 17. 72-93. https://doi.org/10.1186/1471-2164-
629 72.
630 [25] Kampinga HH, Hageman J, Vos MJ, Kubota H, Tanguay RM, et al. 2009.
631 Guidelines for the Nomenclature of the Human Heat Shock Proteins. Cell
632 Stress Chaperones, 14(1). 105–11. https://doi.org/10.1007/s12192-008-
633 0068-7.
634 [26] Wu AR, Nef NF, Kalisky T, Dalerba P, Treutlein B, et al. 2014. Quantitative
635 assessment of single-cell RNA-sequencing methods. Nat Methods, 11(1).
636 41-46.
637 [27] Shi Y and He M.2014. Differential gene expression identified by RNA-Seq
638 and qPCR in two sizes of pearl oyster (*Pinctada fucata*). Gene, 538(2). 313-
639 322. https://doi.org/10.1016/j.gene.2014.01.031
640 [28] Cheng Y, Xia Y, Shao C, Han L, Chen X, et al. 2016. Discovery and
641 identification of candidate sex-related genes based on transcriptome
642 sequencing of Russian sturgeon (*Acipenser gueldenstaedtii*) gonads.
643 Physiol Genomics, 48(7). 464-476. doi:
644 10.1152/physiolgenomics.00113.2015.
645 [29] Yue H, Li C, Du H, Zhang S, Wei Q. 2015. Sequencing and *de novo*
646 assembly of the gonadal transcriptome of the endangered Chinese sturgeon

(*Acipenser sinensis*). PLoS One, 10(6):e0127332. https://doi.org/10.1371/journal.pone.0127332

[30] Song W, Jiang K, Zhang F, Lin Y, Ma L. 2016. RNA-sequencing of the sturgeon *Acipenser baerii* provides insights into expression dynamics of morphogenic differentiation and developmental regulatory genes in early versus late developmental stages. BMC Genomics,17. 564 doi: 10.1186/s12864-016-2839-3.

[31] Kaitetzidou E, Ludwig A, Gessner J, Sarropoulou E,. 2017. Expression Patterns of Atlantic Sturgeon (*Acipenser oxyrinchus* ) During Embryonic Development. G3 (Bethesda), 7(2). 533–542. https://doi.org/10.1534/g3.116.036699.

[32] Li S, Zhang Y, Cao Y, Wang D, Liu H, et al. 2017. Transcriptome profiles of Amur sturgeon spleen in response to infection. Fish Shellfish Immunol, 70. 451-460.

[33] Li S, Wang D, Cao Y, Zhang Y, Liu H, et al. 2018. Transcriptome profile of Amur sturgeon (*Acipenser schrenckii*) liver provides insights into immune modulation in response to *Yersinia ruckeri* infection. Aquaculture, 492. 137-146.

[34] Kassahn KS, Crozier RH, Pörtner HO, Caley MJ. 2009. Animal Performance and Stress: Responses and Tolerance Limits at Different Levels of Biological Organisation. Biol Rev Camb Philos Soc, 84(2). 277–92. https://doi.org/10.1111/j.1469-185X.2008.00073.x.

[35] Komoroske LM, Connon RE, Jeffries KM, Fangue NA. 2015. Linking Transcriptional Responses to Organismal Tolerance Reveals Mechanisms of

671 Thermal Sensitivity in a Mesothermal Endangered Fish. Mol Ecol, 24(19).
672 4960–81. https://doi.org/10.1111/mec.13373.
673 [36] Jesus TF, Grosso AR, Almeida-Val VMF, Coelho MM. 2016.
674 Transcriptome Profiling of Two Iberian Freshwater Fish Exposed to Thermal
675 Stress. J Therm Biol, 55. 54–61.
676 https://doi.org/10.1016/j.jtherbio.2015.11.009.
677 [37] Ni M, Wen H, Li J, Chi M, Ren Y, et al. 2014. Two HSPs Gene from
678 Juvenile Amur Sturgeon (*Acipenser schrenckii*): Cloning, Characterization
679 and Expression Pattern to Crowding and Hypoxia Stress. Fish Physiol
680 Biochem, 40(6). 1801–16. https://doi.org/10.1007/s10695-014-9969-9.
681 [38] Lee S, Hung SSO, Fangue NA, Haller L, Verhille CE, et al. 2016. Effects of
682 Feed Restriction on the Upper Temperature Tolerance and Heat Shock
683 Response in Juvenile Green and White Sturgeon. Comp Biochem Physiol A
684 Mol Integr Physiol, 198. 87–95. https://doi.org/10.1016/j.cbpa.2016.04.016.
685 [39] Peng G, Zhao W, Shi Z, Chen H, Liu Y, et al. 2016. Cloning HSP70 and
686 HSP90 Genes of Kaluga (*Huso dauricus*) and the Effects of Temperature
687 and Salinity Stress on Their Gene Expression. Cell Stress Chaperones,
688 21(2). 349–59. https://doi.org/10.1007/s12192-015-0665-1.
689 [40] Zhang Y, Loughery JR, Martyniuk CJ, Kieffer JD. 2016. Physiological and
690 Molecular Responses of Juvenile Shortnose Sturgeon (*Acipenser*
691 *Brevirostrum*) to Thermal Stress. Comp Biochem Physiol A Mol Integr
692 Physiol, 203. 314–21. https://doi.org/10.1016/j.cbpa.2016.10.009.
693 [41] Yuan X, Zhou YH, Huang YP, Guo WT, Johnson D, et al. 2017. Effects of
694 Temperature and Fatigue on the Metabolism and Swimming Capacity of

695 Juvenile Chinese Sturgeon (*Acipenser sinensis*). Fish Physiol Biochem,
696 43(5). 1279-1287. https://doi.org/10.1007/s10695-017-0371-2
697 [42] Zhang Y and Kieffer JD. 2017. The Effect of Temperature on the Resting
698 and Post-Exercise Metabolic Rates and Aerobic Metabolic Scope in
699 Shortnose Sturgeon *Acipenser brevirostrum*. Fish Physiol Biochem 43(5),
700 1245-1252. https://doi.org/10.1007/s10695-017-0368-x.
701 [43] Grunow B, Noglick S, Kruse C, Gebert M. 2011. Isolation of Cells from
702 Atlantic Sturgeon *Acipenser Oxyrinchus Oxyrinchus* and Optimization of
703 Culture Conditions. Aquat Biol, 14. 67–75. https://doi.org/10.3354/ab00383.
704 [44] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. 2009.
705 BLAST+: Architecture and Applications. BMC Bioinformatics,10. 1–9.
706 https://doi.org/10.1186/1471-2105-10-421.
707 [45] Langmead B and Salzberg SL. 2012. Fast Gapped-Read Alignment with
708 Bowtie 2. Nat Methods, 9(4). 357–59.
709 [46] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The
710 Sequence Alignment/Map Format and SAMtools. Bioinformatics, 25(16).
711 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
712 [47] Anders S, Pyl PT, Huber W. 2015. HTSeq-A Python Framework to Work
713 with High-Throughput Sequencing Data. Bioinformatics, 31(2). 166–69.
714 https://doi.org/10.1093/bioinformatics/btu638.
715 [48] Anders S and Huber W. 2010. Differential Expression Analysis for
716 Sequence Count Data. Genome Biol, 11. R106. https://doi.org/10.1186/gb-
717 2010-11-10-r106.
718 [49] Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, et al. 2003.
719 PANTHER: A Library of Protein Families and Subfamilies Indexed by

720 Function. Genome Res, 13(9). 2129–2141.
721 https://doi.org/10.1101/gr.772403.
722 [50] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. 2000. Gene
723 Ontology: Tool for the Unification of Biology. Nat Genetics, 25(1). 25–29.
724 https://doi.org/10.1038/75556.
725 [51] Carbon S, Dietze H, Lewis SE, Mungall CJ, Munoz-Torres MC, et al. 2017.
726 Expansion of the Gene Ontology Knowledgebase and Resources: The Gene
727 Ontology Consortium. Nucleic Acids Res, 45(D1). D331–38.
728 https://doi.org/10.1093/nar/gkw1108.
729 [52] Hudson DL. 2002. Keratins as Markers of Epithelial Cells. In: Wise C. (eds)
730 Epithelial Cell Culture Protocols. Methods in Molecular Medicine, 188.
731 Humana Press.
732 [53] Linares-Casenave J, Werner I, van Eenennaam JP, Doroshov SI. 2013.
733 Temperature Stress Induces Notochord Abnormalities and Heat Shock
734 Proteins Expression in Larval Green Sturgeon (*Acipenser medirostris*, Ayres
735 1854). J Appl Ichthyol, 29(5). 958–67. https://doi.org/10.1111/jai.12220.
736 [54] Humphreys DT, Carver JA, Easterbrook-Smith SB, Wilson MR. 1999.
737 Clusterin has chaperone-like activity similar to that of small heat shock
738 proteins. J Biol Chem, 274(11). 6875-6881.
739 https://doi.org/10.1074/jbc.274.11.6875.
740 [55] Carver JA, Rekas A, Thorn DC, Wilson MR. 2003. Small heat-shock
741 proteins and clusterin: intra- and extracellular molecular chaperones with a
742 common mechanism of action and function. IUBMB Life, 55(12): 661-668.
743 https://doi.org/10.1080/15216540310001640498

744 [56] Viard I, Wehrli P, Jornot L, Bullani R, Vechietti JL, Schifferli, Tschopp J,
745 French LE. 1999. Clusterin gene expression mediates resistance to
746 apoptotic cell death induced by heat shock and oxidative stress. J Invest
747 Dermatol, 112(3). 290-296. https://doi.org/10.1046/j.1523-
748 1747.1999.00531.x
749 [57] Jarolim S, Ayer A, Pillay B, Gee AC, Phrakaysone A, et al. 2013.
750 *Saccharomyces Cerevisiae* Genes Involved in Survival of Heat Shock. G3
751 (Bethesda), 3(12). 2321–33. https://doi.org/10.1534/g3.113.007971.
752 [58] Phithakrotchanakoon C, Puseenam A, Phaonakrop N, Roytrakul S,
753 Tanapongpipat S, et al. 2018. Hac1 Function Revealed by the Protein
754 Expression Profile of a OtHAC1 Mutant of Thermotolerant Methylotrophic
755 Yeast Ogataea Thermomethanolica. Mol Biol Rep, 45(5). 1311–1319.
756 https://doi.org/10.1007/s11033-018-4287-4.
757 [59] Nobeyama Y and Nakagawa H. 2016. Aberrant DNA Methylation in
758 Keratoacanthoma. PLoS ONE, 11(10). 1–14.
759 https://doi.org/10.1371/journal.pone.0165370.
760 [60] Narberhaus F. 2002. Alpha-Crystallin-Type Heat Shock Proteins:
761 Socializing Minichaperones in the Context of a Multichaperone Network.
762 Microbiol Mol Biol Rev, 66(1). 64–93. https://doi.org/10.1128/MMBR.66.1.64-
763 93.2002.
764 [61] Jäättelä, M. 1999. Escaping cell death: survival proteins in cancer. Exp
765 Cell Res, 248(1). 30-43.
766 [62] Long Y, Li L, Li Q, He X, Cui Z. 2012. Transcriptomic Characterization of
767 Temperature Stress Responses in Larval Zebrafish. PLoS ONE, 7(5).
768 https://doi.org/10.1371/journal.pone.0037209.

769 [63] Arrigo AP. 2007. The Cellular ‘Networking’ of Mammalian Hsp27 and Its
770 Functions in the Control of Protein Folding, Redox State and Apoptosis. Adv
771 Exp Med Biol, 594. 14–26. https://doi.org/10.1007/978-0-387-39975-1_2.
772 [64] Wang K and Spector A. 1996. Alpha-Crystallin Stabilizes Actin Filaments
773 and Prevents Cytochalasin-Induced Depolymerization in a Phosphorylation-
774 Dependent Manner. Eur J Biochem, 242(1). 56–66.
775 https://doi.org/10.1111/j.1432-1033.1996.0056r.x.
776 [65] Kamradt MC, Chen F, Sam S, Cryns VL. 2002. The Small Heat Shock
777 Protein AB-Crystallin Negatively Regulates Apoptosis during Myogenic
778 Differentiation by Inhibiting Caspase-3 Activation. J Biol Chem, 277(41).
779 38731–38736. https://doi.org/10.1074/jbc.M201770200.
780 [66] Smith CC, Yu YX, Kulka M, Aurelian L. 2000. A Novel Human Gene
781 Similar to the Protein Kinase (PK) Coding Domain of the Large Subunit of
782 Herpes Simplex Virus Type 2 Ribonucleotide Reductase (ICP10) Codes for
783 a Serine-Threonine PK and Is Expressed in Melanoma Cells. J Biol Chem,
784 275(33). 25690–25699. https://doi.org/10.1074/jbc.M002140200.
785 [67] Kappé G, Boelens WC, de Jong WW. 2010. Why Proteins without an α-
786 Crystallin Domain Should Not Be Included in the Human Small Heat Shock
787 Protein Family HSPB. Cell Stress Chaperones, 15(4). 457–61.
788 https://doi.org/10.1007/s12192-009-0155-4.
789 [68] Marvin M, O’Rourke D, Kurihara T, Juliano CE, Harrison KL, et al. 2008.
790 Developmental Expression Patterns of the Zebrafish Small Heat Shock
791 Proteins. Dev Dyn, 237(2). 454–63. https://doi.org/10.1002/dvdy.21414.
792 [69] Wang W, Deng DF, de Riu N, Moniello G, Hung SSO. 2013. Heat Shock
793 Protein 70 (HSP70) Responses in Tissues of White Sturgeon and Green

794 Sturgeon Exposed to Different Stressors. North Am J Aquac, 75(2). 164–69.
795 https://doi.org/10.1080/15222055.2012.747457.
796 [70] Werner I, Linares-Casenave J, van Eenennaam JP, Doroshov SI. 2007
797 The Effect of Temperature Stress on Development and Heat-Shock Protein
798 Expression in Larval Green Sturgeon (*Acipenser medirostris*). Environ Biol
799 Fish, 79(3-4). 191–200. https://doi.org/10.1007/s10641-006-9070-z.
800 [71] Iwama GK, Vijayan MM, Forsyth RB, Ackerman PA. 1999. Heat Shock
801 Proteins and Physiological stress in fish. Am Zool, 39(6). 901-909.
802 [72] Demeke A and Tassew A. 2015. Heat Shock Protein and their Significance
803 in Fish Health. Research & Reviews: J Veterinary Sci, 2(1). 66-75.
804 [73] Oksala NKJ, Ekmekçi FG, Özsoy E, Kirankaya S, Kokkola T, et al. 2014.
805 Natural thermal adaptation increases heat shock protein levels and
806 decreases oxidative stress. Redox Biol, 3. 25-28.
807 [74] Heikkila JJ. 2017. The expression and function of hsp30-like small heat
808 shock protein genes in amphibians, birds, fish and reptiles. Comp Biochem
809 Physiol A Mol Integr Physiol, 203. 179-192.
810 [75] Zarate J and Bradley TM. 2003. Heat shock proteins are not sensitive
811 indicators of hatchery stress in salmon. Aquaculture, 223. 175-187.
812 [76] Tomalty KMH, Meek MH, Stephens MR, Rincon G, Fangue NA, et al.
813 Transcriptional response to acute thermal exposure in juvenile Chinook
814 salmon determined by RNAseq. G3 (Bethesda) 5, 1335-1349.
815 [77] Chen Z, Farrel AP, Matala A, Narum SR. 2018. Mechanisms of thermal
816 adaptation and evolutionary potential of conspecific population to changing
817 environments. Mol Ecol 27(3), 659-674.

818 [78] Quinn NL, McGowan CR, Cooper GA, Koop BF, Davison WS. 2011.
819 Identification of genes associated with heat tolerance in Arctic char exposed
820 to acute thermal stress. Physiol Genomics, 43. 685-696.
821 [79] Healy TM, Tymchuk WE, Osborne EJ, Schulte PM. 2010. Heat shock
822 response of killfish (*Fundulus heteroclitus*): candidate gene and
823 heterologous microarray approaches. Physiol Genomics, 41. 171-184.
824 .

## Figures and tables

**Table 1.** Overview of assemblies' statistics and read mapping. List and number of input sequence reads for each assembly, number of resulting contigs, assembly length in Megabases (Mb), contig N50 in Kilobases (Kb), maximum contig length in Kilobases (kb) and percentage of cell-derived sequencing reads mapped to each assembly.

| **Assembly** | **Cell** | **Organ** | **Embryo** | **COE** |
|---|---|---|---|---|
| **Input reads** | cell | organs | embryo | cell + organs + embryo |
| **n. reads (million)** | 364,86 | 639,59 | 380,08 | 1424,41 |
| **Contigs** | 53,624 | 641,485 | 179,564 | 203,131 |
| **Assembly size (Mb)** | 34.80 | 342.28 | 166.71 | 254.00 |
| **N50 (Kb)** | 1.08 | 0.60 | 1.94 | 1.87 |
| **Max (Kb)** | 15.63 | 16.64 | 54.44 | 34.02 |
| **Accession number** | GGWJ01000000 | GGZX01000000 | GEUL01000000 | GGZT01000000 |
| **Mapped reads (%)** | 74.0 | 54.0 | 75.0 | 89.6 |

**Table 2.** List of the core 27 up-regulated contigs and corresponding gene description, sturgeon gene name and other names present in the literature.

| COE contig/s ID | Gene description | Sturgeon gene | Other names |
|---|---|---|---|
| 75185, 75186, 75188 | Heat shock protein beta-11 | *hspb11a/b* | *hsp30* |
| 81106, 81108 | Heat shock protein beta-1 | *hspb1* | *hsp27, hsp25* |
| 68968 | Heat shock 70kDa protein 4L | *hsph3a* | *hspa4l, apg1* |
| 66436, 3989, 4740, 133784, 161880, 133767, 63068, 79530, 69039, 69847, 169012 | Heat shock protein 90kDa alpha family class A member 1 | *hspc1* | *hsp90aa1.2, hsp86, hsp90, hsp89* |
| 113709 | Heat shock 70kDa protein 4 | *hsph2* | *hspa4, apg2, hsp110* |
| 68283, 7082 | Clusterin | *clu* | *clu* |
| 22979 | Growth factor receptor-bound protein 2 | *grb2* | *grb2* |
| 166695 | atrial natriuretic peptide receptor 2 | *npr2* | *npr2* |
| 64489 | Coiled-coil domain-containing protein 17 | *ccdc17* | *ccdc17* |
| 46154, 75199 | n.a. | n.a. | n.a. |
| 62264, 62266 | Uncharacterized protein | n.a. | n.a. |

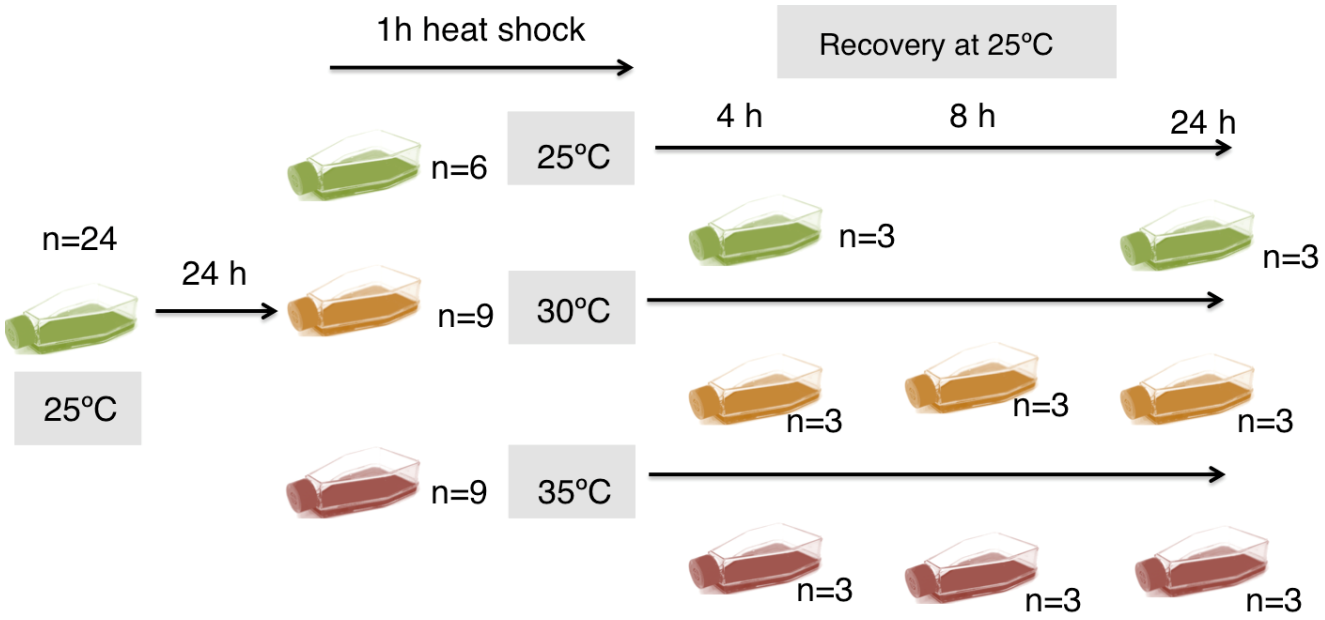
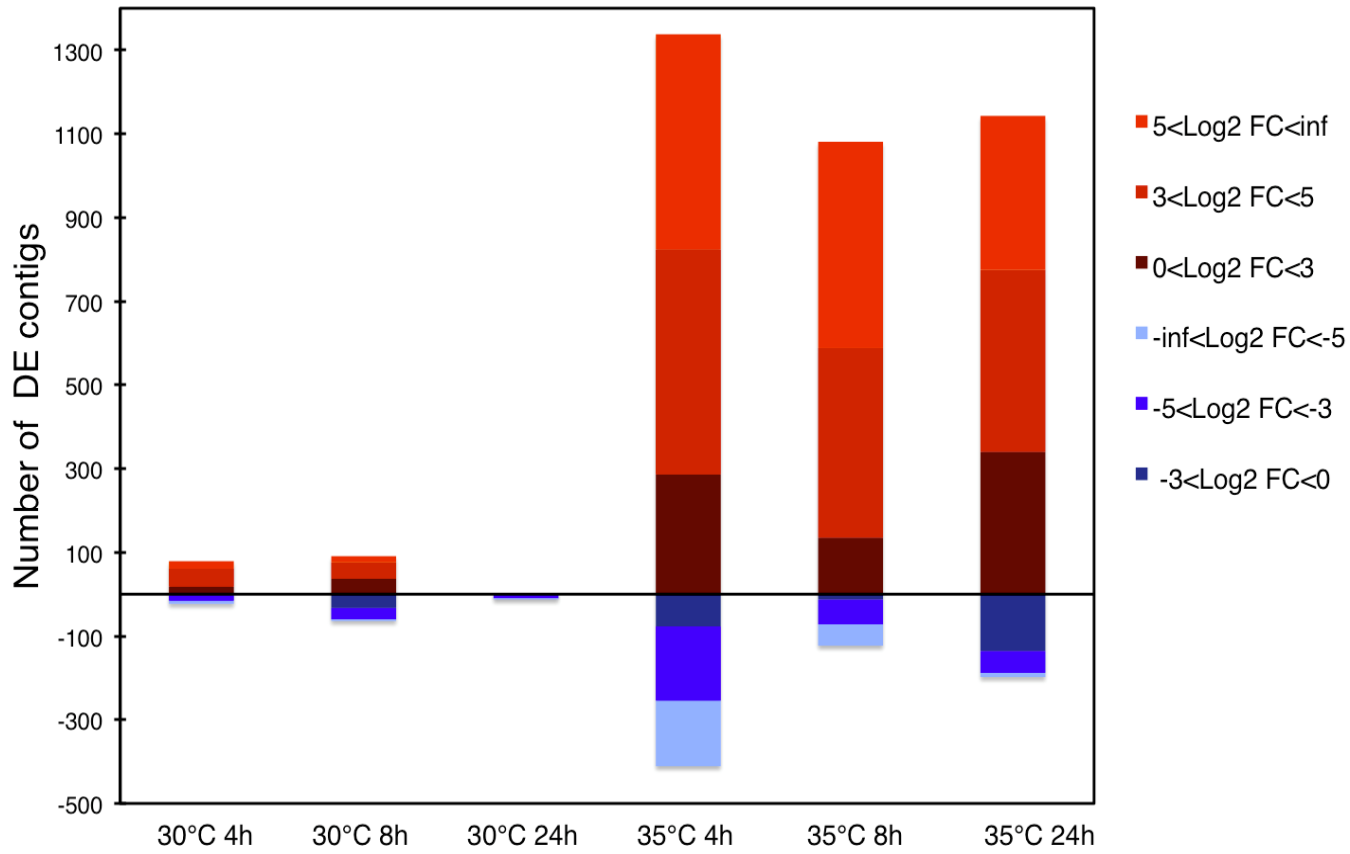



**Figure 1. Experimental design.** The AOXlar7y cell line was seeded in 25mL flasks and cultured at 25ºC for 24 hours (h). Then, the cells were treated at either 30º or 35º for 1 hour and then transferred back to 25ºC for recovery. RNA was harvested after recovery for 4, 8 and 24 hours and sequenced with Illumina HiSeq2500.




871

872 **Figure 2. Number of differentially expressed contigs (DECs) per condition**.
873 Positive values represent upregulated contigs and negative values represent
874 downregulated contigs. Different brightness represent different fold changes (FC)
875 ranges.

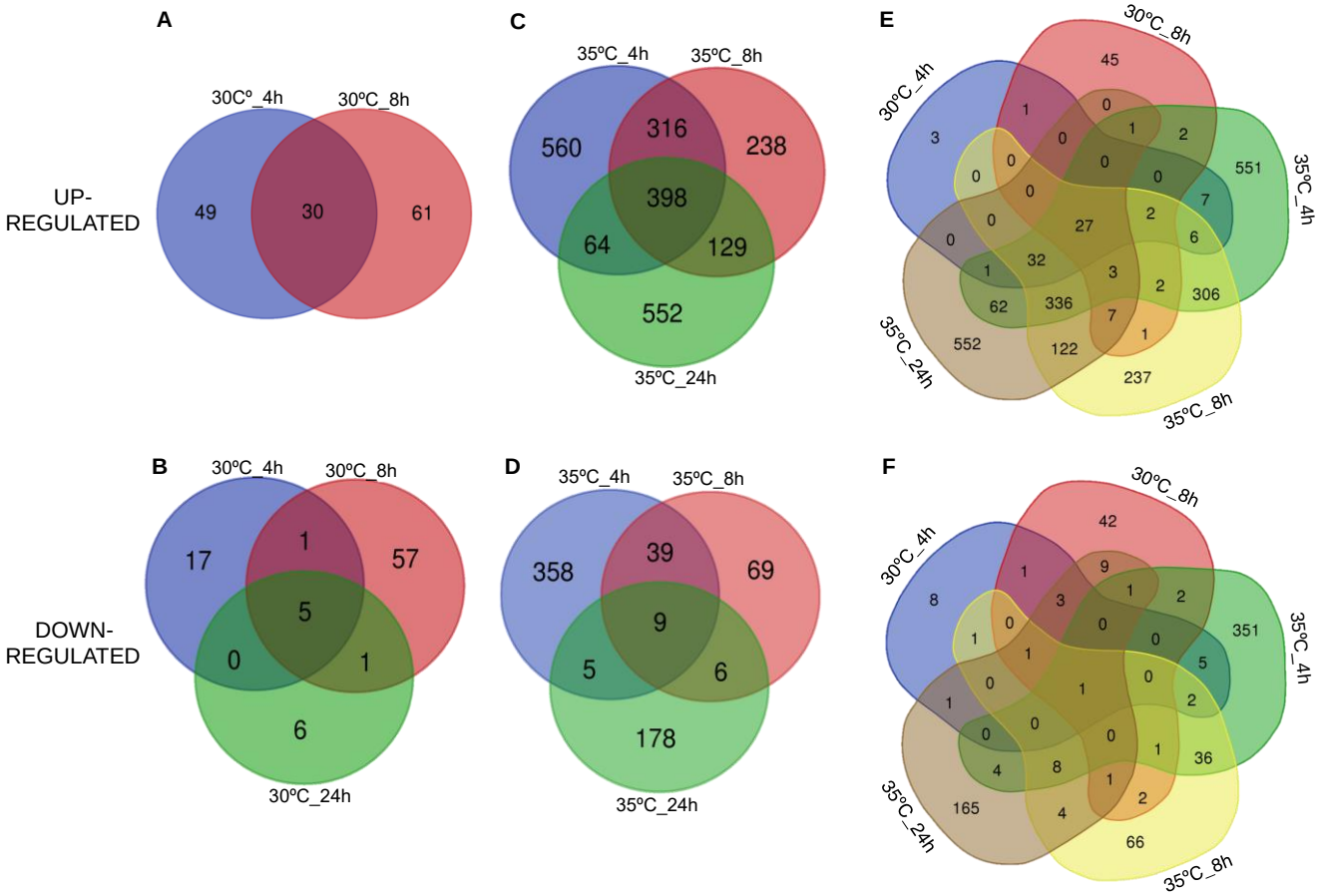



**Figure 3. Venn diagrams showing the differentially expressed contigs (DECs) overlaps per condition**. The last timepoint (24h) after the 30ºC heat shock was excluded from the diagrams as it had only 12 downregulated contigs and none upregulated. Venn diagrams at the top show the overlap of upregulated contigs at different timepoints after 30ºC (A), 35ºC (C) or both (E) treatments. At the bottom, Venn diagrams show the overlap of downregulated contigs at different timepoints after 30ºC (B), 35ºC (D) or both (F) treatments.

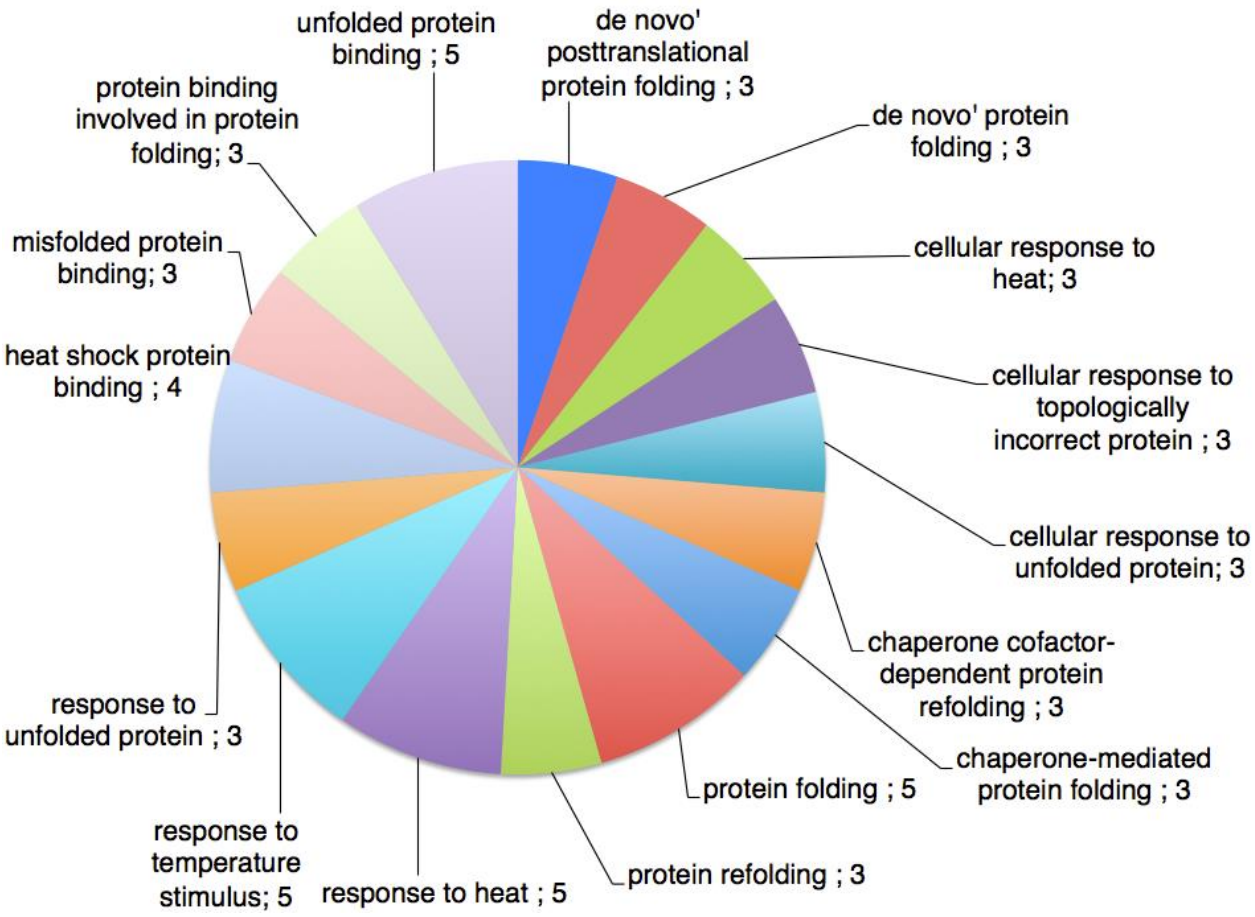



**Figure 4. Pie chart showing overrepresented gene ontology (GO) terms at 30º.** Biological process and molecular functions overrepresented at 30ºC after 4 hours, including the number of differentially expressed genes in each term.

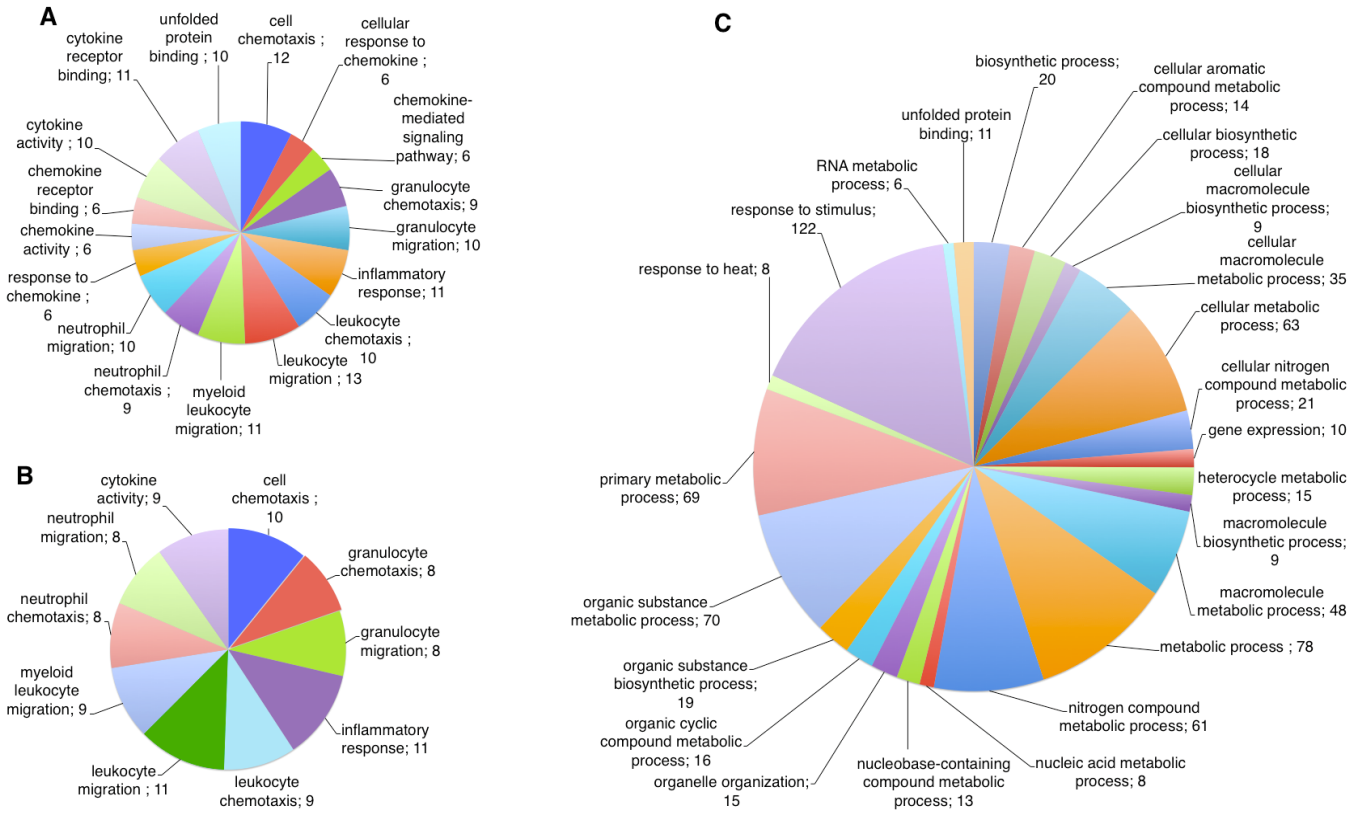



**Figure 5. Pie chart showing overrepresented gene ontology (GO) terms at 35º.** Biological process and molecular functions overrepresented at 35ºC after **A:** 4 hours; **B:** 8 hours; **C:** 24 hours, including the number of differentially expressed genes in each term

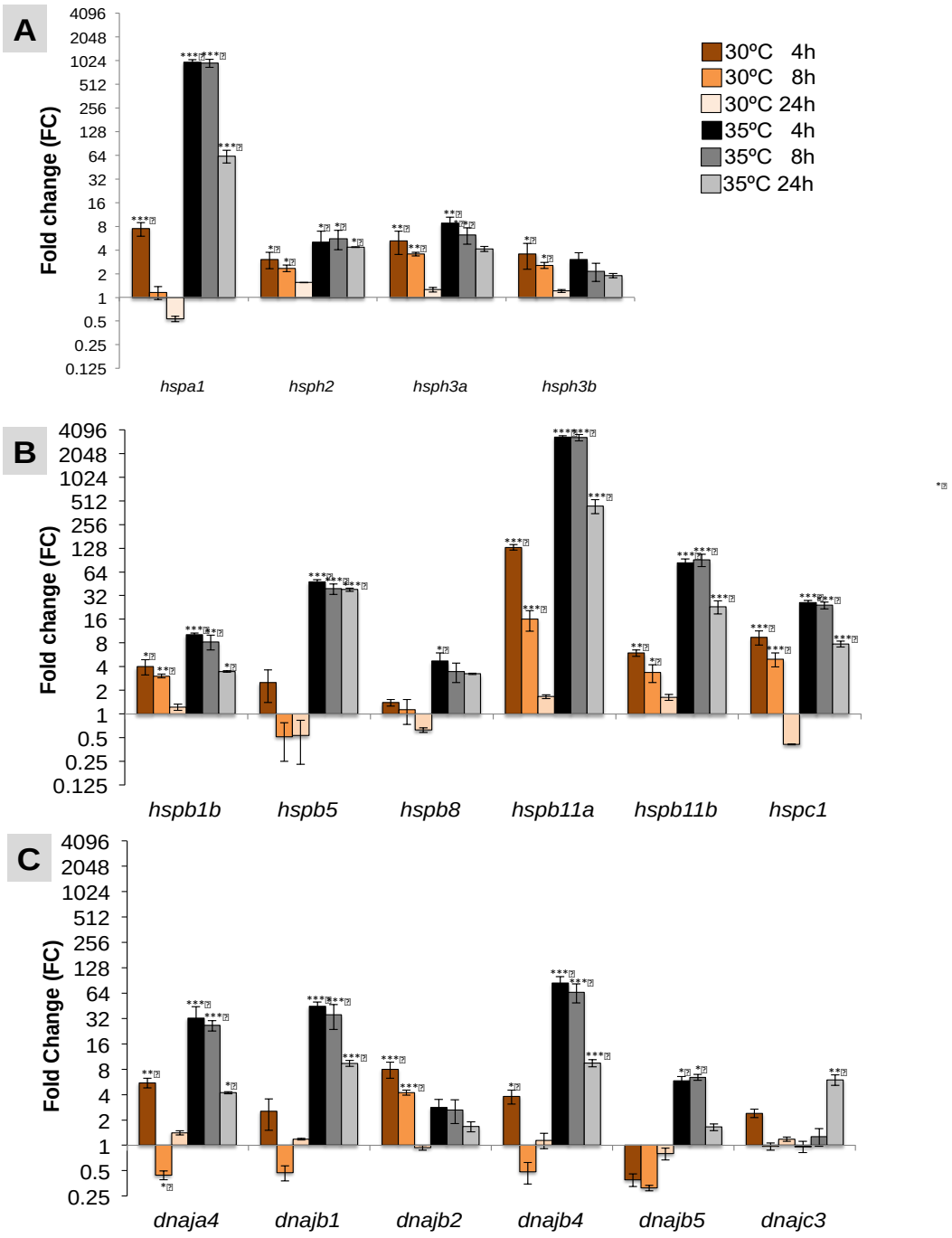



**Figure 6. Fold change (FC) of the heat shock inducible genes per condition. A:** HspA and HspH genes**; B:** HspB and HspC genes; **C:** DNAJ genes**.** Differences are considered significant when padj < 0.05. *padj < 0.05, **padj < 0.001, ***padj < 0.0001.

**Supplementary data**

**Table S1:** Sample list and corresponding number of sequencing reads per sample.

**Table S2**: List of top zebrafish BLASTx hits linked to of the COE contigs.

**Table S3:** Number of AOXlar7y sequencing reads mapped to the COE reference assembly.

**Table S4:** Top100 expressed contigs in the untreated cell line and best BLAST hits.

**Table S5**: DESeq results.

**Table S6:** List of differentially expressed contigs (DECs) in at least one condition with best zebrafish BLASTx hit.

**Table S7**: List of the always upregulated 27 COE contigs and best blastx hits.

**Table S8:** List of differentially represented GO terms per comparison.

**Table S9:** List of the annotated Hsp and Hsf genes in Atlantic sturgeon.

**Table S10:** Number of AOXlar7y sequencing reads mapped to the list of annotated genes and DESeq analysis results.

**Highlights**

- The heat shock response of an Atlantic sturgeon cell line was determined by RNAseq.
- The core of the upregulated contigs corresponded to heat shock proteins (HSP).
- 16 out of 76 HSP genes were significantly responsive to the applied heat shock.
- *hspb11a* was the most heat-inducible gene.