Identification of genes associated with heat tolerance in Arctic charr exposed to acute thermal stress

Nicole L. Quinn, Colin R. McGowan, Glenn A. Cooper, Ben F. Koop, and William S. Davidson

1Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia; 2Icy Waters Incorporated, Whitehorse, Yukon; 3Department of Biology, University of Victoria, Victoria, British Columbia, Canada

Submitted 7 January 2011; accepted in final form 29 March 2011

Quinn NL, McGowan CR, Cooper GA, Koop BF, Davidson WS. Identification of genes associated with heat tolerance in Arctic charr exposed to acute thermal stress. Physiol Genomics 43: 685–696, 2011. First published April 5, 2011; doi:10.1152/physiolgenomics.00008.2011.—Arctic charr is an especially attractive aquaculture species given that it features the desirable tissue traits of other salmonids and is bred and grown at inland freshwater tank farms year round. It is of interest to develop upper temperature tolerant (UTT) strains of Arctic charr to increase the robustness of the species in the face of climate change and to enable production in more southern regions. We used a genomics approach that takes advantage of the well-studied Atlantic salmon genome to identify genes that are associated with UTT in Arctic charr. Specifically, we conducted an acute temperature trial to identify temperature tolerant and intolerant Arctic charr individuals, which were subject to microarray and qPCR analysis to identify candidate UTT genes. These were compared with genes annotated in a quantitative trait locus (QTL) region that was previously identified as associated with UTT in rainbow trout and Arctic charr and that we sequenced in Atlantic salmon. Our results suggest that small heat shock proteins as well as HSP-90 genes are associated with UTT. Furthermore, hemoglobin expression was significantly downregulated in tolerant compared with intolerant fish. Finally, QTL analysis and expression profiling identified Coup-TFII as a candidate UTT gene, although its specific role is unclear given the identification of two transcripts, which appear to have different expression patterns. Our results highlight the importance of using more than one approach to identify candidate genes, particularly when examining a complicated trait such as UTT in a highly complex genome for which there is no reference genome.

broodstock development; expression profiling; upper temperature tolerance; quantitative trait locus analysis

THE SALMONIDS (salmon, trout, and charr) are of substantial environmental, economic, and social value. They contribute to ecosystem health as well as to local and global economies through fisheries, aquaculture, and sport fishing. Their increasing popularity as a food choice for humans has created a demand for salmonid flesh such as that of Atlantic salmon, rainbow trout, Pacific salmon species, and, more recently, Arctic charr, which cannot be sustained by wild populations alone. This, combined with increasing environmental threats to wild populations, has fueled the demand for sustainable and effective aquaculture methods. Genomics tools have long been successful in facilitating selective breeding for genetic improvement of cultured stocks (broodstocks) in agriculture species [e.g., swine (23) and cattle (47)]. As the availability of genomics resources has increased for aquatic species, more research is being done to improve aquaculture broodstocks for species such as Atlantic salmon (3), Atlantic cod (15), rainbow trout (46), and catfish (36, 41).

Arctic charr (Salvelinus alpinus) is an especially attractive aquaculture species given that it features the desirable tissue traits of other salmonids, commands a high market value, and is bred and grown at inland freshwater tank farms year round. This circumvents some of the adverse affects of marine net pen aquaculture, the current method used for most species of salmon (2). However, Arctic charr is a cold-water species that thrives in water temperatures from 0.1 to 14°C, which presents substantial geographical limitations in terms of where this species can be grown at present. Tank farms that are otherwise equipped to grow and distribute freshwater fish species, including salmonids such as rainbow trout, often cannot accommodate Arctic charr due to an unsuitable climate during at least part of the year and the high energy cost of maintaining tanks within the optimal temperature range for their survival. In addition, fish forced to live at temperatures higher than their natural range show signs of stress, including reductions in immune function, appetite, growth and reproduction, as well as susceptibility to disease and ultimately death (30). This is a problem of increasing concern, even in temperate regions where the species is currently farmed, as temperatures are rising as a result of climate change. As temperatures continue to climb and become less predictable, Arctic charr hatcheries and tank farms throughout the world will be faced with an on-going battle to keep fish alive, healthy and comfortable, and to keep them growing and spawning at the optimal rate. Arctic charr with different genetic backgrounds show markedly different abilities to withstand thermal stress (unpublished observations, C. R. McGowan). Understanding the genes involved in upper temperature tolerance (UTT) in Arctic charr as well as other salmonids stands to benefit both the aquaculture industry by facilitating the development of more robust broodstock, as well as natural populations, as such knowledge can feed into population-based conservation initiatives against the impacts of climate change.

The common ancestor of salmonids underwent a whole genome duplication event between 20 and 120 million years ago (1). Thus, the extant salmonid species are considered pseudo-tetraploids whose genomes are in the process of reversion to a stable diploid state. This, combined with the repetitive nature of the salmonid genomes in general (10) and the lack of a fully sequenced reference genome (9) makes identifying the genes responsible for complex traits such as UTT difficult. Common approaches to such a task include the identification and analysis of quantitative trait loci (QTL), as well as expression analyses, which include microarray analysis and qPCR (22). However, individually, each of these methods has advan-
tages and shortcomings and may not provide the most accurate or comprehensive results on their own (see Discussion). In an attempt to circumvent these drawbacks while providing added confidence to QTL and expression data, it has become increasingly popular to use a combination of these methodologies, with the goal of identifying overlap between differentially expressed genes and QTL regions (17). The effectiveness of combining QTL and expression approaches for positively identifying candidate genes depends on many factors, including the resolution of the QTL analysis and the genome coverage provided by the expression analysis. In addition, the ability to detect a correlation between gene expression and QTL depends on the nature of the factor driving the QTL, an issue that is addressed within the Discussion of this paper.

Here, we adopted an approach that combines previously published QTL identification with expression profiling using the 32K GRASP microarray (20) and qPCR analysis to identify genes associated with UTT in Arctic char (15a). Specifically, for UTT were previously identified in rainbow trout and Arctic char (15a, 31, 42). We mapped one of these QTL (that associated with markers SsaF43NUIG and Ssa20.19NUIG) to the same location in a homologous linkage group in Atlantic salmon (linkage group 23). We then used the Atlantic salmon genomic resources (7, 27, 39, 44) to identify nine bacterial artificial chromosomes (BACs) spanning a portion (that surrounding the SsaF43NUIG marker) of this QTL within the Atlantic salmon genome. The BACs were sequenced and annotated (38), thereby generating a list of putative UTT genes. In this study, we conducted an acute thermal trial to identify tolerant and intolerant Arctic char. RNA extracted from the gills of these fish was reverse transcribed into cDNA and used for microarray analysis, thus expanding the list of putative UTT genes identified previously. Gill tissue was chosen based on the results of a preliminary test, which indicated that of gill, liver, and muscle tissues, gill would likely identify the most differentially expressed genes. Moreover, there is evidence in the literature to indicate that gill plays an important role in stress tolerance (4, 5). We reasoned that any genes identified by either the previously conducted sequencing of the QTL region or the current microarray study would be particularly strong candidates for UTT involvement. Finally, qPCR analysis was used to examine the behaviors of specific genes and to test the results of the microarray study. This combination of approaches enabled us to conduct an examination of the Arctic char genome and thus identify genes putatively involved in UTT with higher confidence than would be provided by a single approach. In addition, our results highlighted the strengths and potential pitfalls of each of these methods, particularly when studying a complex, duplicated genome for which there is, as yet, no reference sequence (9).

MATERIALS AND METHODS

Mapping of SsaF43NUIG and Ssa20.19NUIG in Atlantic salmon. Previous reports identified UTT QTL in rainbow trout (15a, 31) and Arctic char (42), which were associated with markers Ssa20.19NUIG and SsaF43NUIG, respectively. These two markers, through their common association with marker CoCl3LAV, were found to map to the same location on homologous linkage groups (i.e., RT-10F and AC-26F in rainbow trout and Arctic char, respectively). We tested these markers for variability within the two Atlantic salmon SALMAP mapping families, Br5 and Br6, each of which contains two parents and 46 offspring (7). The forward primer for each pair contained an M13 sequence tag that was used for genotyping analysis. Genotyping results were analyzed with LINKMIFEX ver. 2.3 (8).

Experimental design and tissue collection. All experiments were conducted according to the Canadian Council for Animal Care Guidelines and were approved by the Animal Care Committee at Simon Fraser University, Canada. The temperature trial experiments were conducted at Icy Waters Inc., Whitehorse, Yukon, Canada, in September 2008, using 2006 young of the year Nauyuk Lake Arctic char. Tanks were set up with a constant flow-through system (0.33 l/s) with fresh spring water at ambient temperature (~6°C) and ambient oxygen levels (10.0–11.0 ppm). Approximately 200 fish were transferred to an experimental tank (diameter 1.86 m, depth 50 cm) and left to acclimatize for 48 h at ambient temperature. After acclimation, 10 fish were removed to act as a control group (hereafter referred to as Control fish), then water that had been diverted through a heat exchanger was added to the flow-through system to increase the water temperature in the tank by 6°C/h until it reached 22°C and then 0.5°C every 30 min until the water reached 25°C, the observed lethal temperature for these fish. Dissolved oxygen was allowed to fluctuate naturally and decreased from ~10.3 ppm to a minimum of 8.1 ppm during the trial. Fish were not fed after being transferred to the experimental tank to avoid confounding gene expression results due to food metabolism.

When the water temperature reached 25°C, the temperature was held constant and the fish were closely monitored for signs of stress. The first and last 10 individuals to show loss of equilibrium (LOE) were quickly removed from the tank for sampling, thus representing the 5% least and most temperature tolerant fish, respectively (hereafter referred to as the Intolerant and Tolerant treatment groups, respectively). This temperature regime mimicked that conducted by the previous experiments that identified the UTT QTL (15a, 31, 42) with some minor changes due to differences in the available equipment. The first LOE was observed after ~30 min at 25°C, and the last fish showed LOE ~2 h thereafter. Thus, it should be noted that Tolerant fish were exposed to lethal temperatures for up to 2 h longer than Intolerant fish, and we therefore recognize that any genes identified as differentially expressed between Tolerant and Intolerant groups of fish may reflect this unavoidable difference in exposure time, rather than UTT per se. Figure 1 is a schematic diagram of the experimental design. Fish were euthanized by a swift blow to the head and then weighed, and their fork lengths were measured. Blood was withdrawn from the caudal vein of the fish (maximum possible volume ~200 µl), the entire lower half of outer-most gill arch was removed, the entire liver was sampled, and an ~1 cm² section of muscle from above the lateral line and behind the dorsal fin of the fish was removed, in that order. Tissues were placed into RNAlater (Ambion) and were stored at room temperature for 24 h to allow RNAlater to penetrate the tissues, and then moved to ~80°C for storage until use as per the manufacturer’s instructions.

RNA isolation. RNA isolations and microarray analysis were conducted at the University of Victoria, Canada. Total RNA was isolated from gill, muscle, and liver tissue samples. Briefly, tissue samples were removed from RNAlater, blotted on a clean Kimwipe to remove excess solution, and disrupted and homogenized in 1 ml TRIzol reagent using a Mixer-mill (Retch MM 301) with tungsten carbide beads. Phase separation was conducted using 200 µl chloroform, and RNA was purified using the RNeasy Mini Kit (Qiagen) following the instructions. Purified RNA was treated with 1 µl RNase inhibitor (Invitrogen). RNA integrity was verified by agarose gel with ethidium bromide staining to visualize ribosomal bands and by measuring the 260/280 absorbance ratio (>1.9) using a Nano Drop (ND-1000 Spectrophotometer, Thermo Scientific) then stored at ~80°C until use.

Microarray analysis. The microarray study followed a reference design format. cDNA prepared from 1 µg gill RNA from six samples from each treatment group (Tolerant, Intolerant, and Control; 18...
slides in total) using Invitrogen’s SuperScript Indirect cDNA labeling system. Treatment groups were compared indirectly against one another using a common reference sample that was hybridized to each microarray alongside the sample cDNA. The reference sample, designed to hybridize to as many spots on the array as possible, comprised high-quality RNA isolated from Atlantic salmon gonad, brain, and spleen tissues that had been amplified using Ambion’s Amino Allyl Message Amp aRNA kit, then quantified, combined in equal amounts, and divided into per-use aliquots to avoid degradation due to repeated freeze-thawing.

The GRASP 32K cDNA microarray was used (20). Details of the microarray hybridization process can be found at the University of Victoria cGRASP website (http://web.uvic.ca/grasp/microarray/array.html) within the .pdf document entitled Invitrogen Indirect cDNA Labeling System version 3. In brief, slides were postprint processed by being rinsed in 0.2% SDS and water and dried by centrifugation, then prehybridized in 5× SSC, 0.1% SDS, 3% BSA; washed with water; dried again; and stored in a dry oven at 49°C until cDNA hybridization. cDNA (300 ng) and aRNA (500 ng) were labeled with Cy5 and Cy3 (Amersham Biosciences), respectively, using Invitrogen’s SuperScript Indirect cDNA Labeling System and then combined with 2× formamide buffer and LNA diT blocker (Genisphere) to a total volume of 60 μl, which was heated to 80°C and then loaded on to the slide in the dark. Microarrays were incubated for 16 h at 49°C in a dark, humidified chamber, then underwent a series of washes, and were dried by centrifugation. Slides were scanned at 74 and 72 PMT for Cy3 and Cy5, respectively, using a ScanArray Express Microarray Scanner (Packard BioScience BioChip Technologies, model #ASCEX00) and spot intensity was calculated with ImaGene ver. 6.5.1.

A preliminary test was done to determine which tissue type would provide the most information for the analysis. Specifically, Cy5-labeled spots with signals greater or equal to the average base/proportional malize to the median. Spots were filtered on flags present, and only genes not meeting a twofold differential expression between pairs were filtered out in the same step. Next, a Venn diagram was constructed to compare the resulting gene lists against one another. This approach enabled us to decipher transcripts that were identified by two or more of the pair-wise comparisons (i.e., the overlapping portions of the Venn diagram), and any genes showing differential expression in one of the pair-wise comparisons (i.e., the nonoverlapping portions of the Venn diagram).

PCR, cloning, and sequencing of multiple chicken ovalbumin upstream promoter transcription factor II transcripts. As described in RESULTS, chicken ovalbumin upstream promoter transcription factor II (COUP-TFII) was identified by both sequencing of an UTT QTL region in Atlantic salmon and microarray analysis. As putatively playing a role in UTT. A search for the sequence of the COUP-TFII 596 bp expressed sequence tag (EST) from the microarray (GenBank accession number DW547089; hereafter referred to as transcript A) within the Atlantic salmon EST database (http://web.uvic.ca/grasp/Project: Salmo salar - All 100/99) revealed two similar transcripts, the first corresponding to the EST spotted on to the array,
and a second, contig19531 (hereafter transcript B), which consisted of a single read of 710 bp and showed 90% sequence identity with transcript A. Given the duplicated nature of the Atlantic salmon genome, we suspected that these two ESTs may represent duplicated genes. Although transcript B was not present in the 1 Mb region of the UTT QTL previously sequenced (38), and therefore the two transcripts do not reflect a tandem duplication, we do not know whether they are located further apart on the same chromosome, or whether they are located on separate chromosomes. Using the full Atlantic salmon genomic sequence of the COUP-TFI gene from our previous report (38) (GenBank accession number EU481821.1), we designed PCR primers such that they would amplify the entire 596 bp segment of the EST that is on the microarray (transcript A). These primers were used to amplify the EST region by PCR using Arctic charr genomic DNA as a template. The PCR product was cloned using the pETBlue-1 AccepTor Vector kit (Novagen), individual clones were cultured, and the PCR inserts sequenced. This revealed that there were indeed two slightly different products in Arctic charr: one that was highly similar (97%) to the EST on the microarray (transcript A) and one that was more similar to EST contig19531 (98%), or highly similar (97%) to the EST on the microarray (transcript B). Thus, qPCR reverse primers were specifically designed across and around an 11 bp gap in transcript B to specifically amplify transcripts A and B, respectively (the same forward primer was used for both transcripts; see Supplemental Table S1 for primer sequences). This ensured that only one product that was specific to a particular transcript was amplified by qPCR, which was verified by the presence of a single dissociation curve for each product.

Expression analysis using qPCR. We prepared cDNA for qPCR from 1 µg of total RNA using Invitrogen’s SuperScript III Reverse Transcriptase kit following the manufacturer’s instructions. The six RNA samples per treatment group that were used for microarray analysis were used along with RNA from three additional fish per treatment group (i.e., nine individuals per treatment group were tested with qPCR). qPCR primer pairs were designed for 24 genes selected based on interest in function as well as the degree of fold-change observed in the microarray analysis. qPCR primer pairs were designed from the Atlantic salmon EST sequences used on the GRASP 32K array using Primer 3 version 0.4.0 (http://frodo.wi.mit.edu/primer3/). Primers were tested for amplification efficiency using cDNA generated from a single gill RNA sample using the qPCR conditions described below followed by a dissociation curve analysis to test for a single product for each primer pair and that no primer dimers were generated during the 40 amplification cycles. The 13 primer pairs meeting these criteria and showing the highest efficiencies (range 80.7–109.7, Supplemental Table S1) were used for expression analysis of cDNAs from nine individuals from each treatment group. qPCR was conducted using the ABI 7900HT system with Sybr green (Quanta Biosciences) under the following conditions: 95°C for 3 min followed by 40 cycles of 95°C for 1 s, 60°C for 30 s and 72°C for 15 s, with one 96-well plate run per individual cDNA sample, which included, in triplicate, all 13 primers with corresponding no-template controls, and two primer pairs for the endogenous control gene, EF1A. Specifically, we used the EF1A primers designed by Olsvik et al. (29), which cross exons 5 and 6, and also designed primers that span exons 3 and 4 (EF1A3to4) (Supplemental Table S1). Having two primer sets within one gene serves as a control for the quality of the cDNA reverse transcription reaction because the expression of EF1A should be the same using both primer pairs. The amplification efficiencies of the endogenous control primer sets EF1A and EF1A3to4 were 98.6% and 98.8%, respectively. Each plate also contained a no-reverse-transcriptase control (i.e., RNA from the individual being tested that had gone through the steps of the cDNA preparation but lacking reverse transcriptase) to test for genomic DNA contamination of the cDNA. Also included was a linker sample, i.e., cDNA from a single individual amplified with EF1A, which was compared across all plates to test for technical variations between plates (average CT = 21.54, SD = 0.44).

Statistical analysis of qPCR results. qPCR results were analyzed using the ΔΔCt method (32) and calibrated for individual primer amplification efficiencies, producing a relative quantification (RQ) compared with a calibrator individual (a selected untreated control individual). RQ values were tested for outliers using the box-whisker method such that any data points falling outside of the 90% range were eliminated from the analysis. Remaining RQ values were log transformed to meet the assumptions of the statistical tests. Pair-wise Student’s t-tests were performed between pairs among the three groups, Tolerant, Intolerant, and Control. This statistical approach was used to maintain consistency with the microarray analyses, and enabled us to determine the degree to which the gene of interest (GOI) behaved similarly or differently between the microarray and qPCR analyses. All statistical analyses were performed using Graphpad Prism ver. 5.

RESULTS

Mapping of SsaF43NUIG and Ssa20.19NUIG in Atlantic salmon. The microsatellite markers SsaF43NUIG and Ssa20.19NUIG were informative (i.e., variable) in both of the Atlantic salmon SALMAP mapping families, Br5 and Br6 (7), and mapped within 1.1 centimorgan of each other on linkage group 23 of Atlantic salmon, which corresponds to Atlantic salmon chromosome 16 (33) (Fig. 2). This provided strong evidence for a UTT QTL in this region of the Atlantic salmon genome. Thus, tiled Atlantic salmon BACs surrounding these markers were identified, nine of which spanning SsaF43NUIG were sequenced and annotated, as described in detail in our previous report (38).

Fish sizes. Fish weights per treatment group were as follows (average ± standard deviation): Tolerant 46.92 ± 10.62 g, Intolerant 42.89 ± 10.62 g, Control 33.14 g ± 14.0 g. Fish length were: Tolerant 18.67 ± 0.97 cm, Intolerant 18.26 ± 14.72 cm, Control 16.32 ± 2.53 cm. There was no significant difference size between Tolerant and Intolerant fish (both weight and length), although Tolerant fish were larger than Control fish (1-way ANOVA followed by Tukey’s post hoc test P = 0.0292 and P = 0.0151 for weight and length, respectively; Supplemental Fig. S1).

Expression profiling of Tolerant, Intolerant, and Control fish by microarray analysis. We were primarily interested in identifying genes that are associated with UTT, rather than genes that were differentially expressed regardless of the capacity of the individual fish to withstand thermal stress (i.e., general heat response genes). Thus, for the analysis of the microarray data, we focused on the gene lists contained in the three nonoverlapping regions of the Venn diagram (Fig. 3), which represented the genes that were present in only one of the gene lists generated. However, all gene lists, including P values and fold-change values for the pair-wise comparisons are available in Supplemental Table S2. In addition, all microarray data (normalized as well as raw data) were deposited within Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE26306. From the pair-wise comparisons among gene lists, genes were deemed noteworthy or interesting based on suspected function (i.e., genes that have been indicated in previous studies as playing a role in stress response, or those that play substantial roles in major biological pathways) as well as fold-change and significance level
A total of 86 genes were differentially expressed only between Tolerant and Intolerant fish given the parameters assigned (i.e., twofold differential expression, $P < 0.01$). Noteworthy genes that were upregulated in Tolerant compared with Intolerant fish were pituitary homeobox 2a (10.3-fold), actin-alpha cardiac muscle 1 (3.4-fold), and three heat shock proteins (Hsps), all of which belonged to the Hsp-beta family (2.03- to 2.8-fold). Genes that were downregulated in Tolerant compared with Intolerant fish included seven beta-hemoglobin subunits (2.7- to 6.6-fold) and seven alpha-hemoglobin subunits (3.4- to 6.7-fold) (Fig. 4).

A total of 105 genes were differentially expressed only between Tolerant and Control fish. Note that the Control group comprised randomly sampled fish that had not been subjected to heat stress and thus contained an unknown mixture of genotypes. The genes on this list, therefore, may be associated with increased tolerance to heat stress but do not necessarily distinguish temperature tolerant individuals from intolerant ones. Noteworthy genes upregulated in Tolerant fish compared with untreated Controls included six Hsps [five Hsp90 genes (2.1- to 3.5-fold) and Hsp30 (3.41-fold)] as well as ubiquitin (2.5-fold) (Fig. 5), and 78 kDa glucose-regulated protein precursor (3.1-fold). Genes that were downregulated in Tolerant vs. Control fish included two beta-hemoglobin subunits (3.4- and 5.3-fold) and two alpha-hemoglobin subunits (2.7- and 4.1-fold).

A total of 105 genes were differentially expressed only between Intolerant fish and untreated Controls contained 196 genes. Of interest was that COUP-TFII was upregulated in Intolerant fish compared with Controls (2.7-fold) as well as myosin heavy-chain-fast skeletal muscle and myosin heavy polypeptide 11-smooth muscle (3.6- and 2.2-fold, respectively) and Heat shock 70 kDa protein (5.7-fold). Apoptosis-stimulating of p53 protein 1 was downregulated in Intolerant compared with Control fish (4.2-fold).

Comparison of qPCR and microarray results among treatment groups. Supplemental Table S3 lists the genes tested by qPCR with their corresponding GenBank accession numbers, the microarray gene list that the genes were found on, and the corresponding fold change. Also shown in Supplemental Table S3 are the results of the qPCR analysis with the average RQ value for each gene tested per treatment group and results ($P$ values) of pair-wise $t$-tests between groups. Results from the qPCR analysis that are in significant agreement with the microarray analysis are highlighted in yellow, while those showing the same trend as the microarrays are in orange, and those with no trend are left white. Of the 13 genes examined by qPCR, six showed differential expression in the same direction as indicated by the microarray analysis, six showed no trend and one showed significant differential expression in the opposite direction. Specifically, Hsp90-beta showed highly significant ($P < 0.0024$) upregulation in Tolerant and Intolerant fish compared with Controls, with significantly higher expression in Tolerant compared with Intolerant fish (Fig. 6). Additionally, 78 kDa glucose-regulated protein precursor showed downregulation in both Tolerant and Intolerant fish compared with controls ($P < 0.0001$ and $P = 0.0378$, respectively), and Hsp11-beta was significantly upregulated in Tolerant vs. Intolerant fish ($P = 0.03$). Also worth noting is that Hsp90-beta, which was one of two heat shock proteins in the centre of the Venn diagram and (i.e., significantly differentially expressed in
all comparisons in the microarray analysis), showed the same patterns of significant differential expression by qPCR and microarray analyses.

The *COUP-TFII* transcript *A* (i.e., the transcript present on the microarray, see MATERIALS AND METHODS) showed significant upregulation in Intolerant fish compared with Controls (*P* < 0.01; fold-change >2.0), which is in accordance with the results from the microarray analysis. However, interestingly, the second transcript of *COUP-TFII*, transcript *B*, which was not on the microarray, exhibited significant differential regulation in the opposite direction, with decreased expression in Intolerant fish compared with Control fish as well as Tolerant fish (*P* = 0.0018 and *P* = 0.0003, respectively; Fig. 7). Finally, *Hsp7-beta* and actin-alpha cardiac muscle 1 showed trends of differential expression in the same direction as the microarrays but did not reach statistical significance. Conversely, for apoptosis-stimulating of p53 protein 1 there was significant differential expression in the opposite direction as seen from the microarray analysis (i.e., upregulated in Intolerant vs. Control fish; *P* = 0.0022). The remaining genes showed very high *P* values in all comparisons, (i.e., *P* > 0.31), indicating that all treatment groups exhibited similar expression patterns.

**DISCUSSION**

**Upregulation of Hsps in thermo-tolerant fish.** Hsps comprise a well-studied group of highly conserved, ubiquitously distributed proteins that are upregulated when exposed to various stresses. In general, Hsps function as molecular chaperones, acting to maintain protein integrity during cellular stress conditions, including, but not limited to, elevated temperatures (reviewed in Ref. 40). In eukaryotes, Hsps are grouped into families according to their specific functions, sequence similarity and size (e.g., Hsp100, Hsp90, Hsp70, and Hsp60 with molecular weights of 100, 90, 70, and 60 kDa, respectively),

---

**Fig. 3.** Venn diagram of three pair-wise comparisons of treatment groups (i.e., Tolerant vs. Intolerant, Tolerant vs. Control and Intolerant vs. Control; large circles). Numbers refer to the number of genes in that section. We were particularly interested in the gene lists in the nonoverlapping regions, as these genes were more likely to be associated with differences in temperature tolerance, rather than a response to heat stress in general. The center of the diagram contained Hsp90-beta and Heat shock cognate 71 kDa protein (GenBank accession numbers CA062155 and EG813231, respectively). C, Control; I, Intolerant; T, Tolerant.

**Fig. 4.** The mean (6 individuals per treatment group) normalized (reference vs. raw) expression values for alpha (*A*) and beta (*B*) hemoglobin genes significantly differentially expressed between Intolerant and Tolerant groups as determined by the microarray analysis.

**Fig. 5.** The mean (6 individuals per treatment group) normalized (reference vs. raw) expression values plotted against treatment group for all Hsp90-beta genes as well as ubiquitin. *Differential expression between groups meet significance parameters for microarray analysis (i.e., *P* < 0.01, fold-change >2.0).
along with the small Hsp (sHsp) group, which includes members that are 12–43 kDa (12, 40). Not surprisingly, the microarray analysis revealed several Hsps as differentially expressed in Arctic charr subjected to elevated temperatures. However, an interesting pattern emerged when the gene lists were analyzed using a Venn diagram.

First, sHsps were specifically associated with temperature tolerance as two Hsp11-beta transcripts and one Hsp7-beta were upregulated in Tolerant vs. Intolerant fish by microarray analysis, and the qPCR results supported this relationship for one Hsp11-beta and the Hsp7-beta (Supplemental Table S3), whereas no other Hsps were present in this gene list. An additional sHsp11 (also known as Hsp30-beta) gene was upregulated in Tolerant vs. Control fish (microarray only, not tested by qPCR). sHsps exhibit a diverse range of structures but share a conserved sequence of ~80 amino acid residues, the a-crystallin domain, located at the COOH terminus of the protein (14). Functionally, the sHsps display chaperone activities by interacting with unfolding proteins to maintain the folded state (16), while individual sHsps have been reported to play roles in cellular stress resistance and the inhibition of apoptosis (14). It has also been reported that Hsp11-beta was upregulated in heat-shocked zebrafish (Danio rerio) embryos, whereas the expression Hsp7-beta did not change in response to heat shock (12, 24). Finally, Hsp30 (a.k.a. Hsp11) mRNA expression was elevated in the heart, brain, white muscle, red muscle, and liver but not the blood of heat-shocked adult rainbow trout (6). Thus, these sHsps appear to be transcribed in a tissue-specific manner in heat-shocked fish. However, no reports of the activity of either of these genes in response to thermal stress in adult fish gills are currently available. Clearly, further analysis of these sHsps and their roles in heat tolerance in Arctic charr, including expression profiling at various life-stages in different tissues in response to a variety of heat stress regimes is warranted.

Second, the microarray analysis revealed that several members of the Hsp90 family followed the same pattern of expression as exhibited by the sHsps (i.e., Tolerant > Intolerant > Control).
controls). This pattern was validated by qPCR for the Hsp90-beta transcript (GenBank accession number CA062155) that is located in the center of the Venn diagram (Fig. 3). For all other Hsp90 genes, the increased expression between Tolerant and Control fish was statistically significant at P < 0.01 (Fig. 5). In general, Hsp90 genes function to maintain protein integrity, but they have also been reported to play roles in immune function, apoptosis and varying aspects of the inflammatory response in fish (reviewed in Ref. 40). Our results with the Tolerant fish showing elevated Hsp90 expression compared with Intolerant and Control fish indicate that these particular genes may be affiliated with temperature tolerance, and thus they merit further examination in terms of both determining the specific physiological roles of these Hsps as well as deciphering how, functionally and genetically, they differ from other Hsp90s.

Finally, we found that ubiquitin, a small regulatory protein found in all cells that tags proteins for recycling (18), showed the same pattern of expression as the Hsp90 genes (i.e., Tolerant fish > Intolerant > Controls; Fig. 5). On one hand, this is not surprising given that ubiquitin is regularly used as an index of misfolded or damaged proteins and thus is often found at levels similar to Hsps (35). On the other hand, it is interesting to note that Tolerant fish showed the highest levels of ubiquitin gene expression with significantly higher levels than Control fish.

It is worth noting again that, as a result of their prolonged survival, the Tolerant fish were exposed to the final temperature for a longer period than the Intolerant fish; therefore, it cannot be ruled out that some of the elevated expression of these stress-response genes in the Tolerant group may be a consequence of their increased time of exposure to this temperature. There is evidence that warm-adapted individuals of various taxa, including Drosophila and desert lizards, tend to show elevated constitutive levels of Hsps rather than the extreme spikes that tend to be exhibited by cooler-adapted individuals (reviewed in Ref. 43). Thus, future research should focus on the particular role of ubiquitin as well as Hsps, and their relationships, in thermal tolerance (vs. thermal stress in general), as well as the genetic differences in upper temperature Tolerant compared with Intolerant Arctic char.

Combining QTL and expression data. The effectiveness of combining QTL and expression approaches for positively identifying candidate genes depends on many factors, including the resolution of the QTL analysis and the genome coverage provided by the expression analysis. Additionally, the ability to detect a correlation between an expression QTL (eQTL) and a phenotypic QTL (pQTL) depends on the nature of the factor driving the QTL. Modifiers of gene expression can be cis- or trans-acting, which dictates whether the eQTL and pQTL coincide, and thus, whether a combined QTL/expression approach will result in Type II errors (i.e., false negative results) (48). Figure 8 presents three scenarios to illustrate this point. Figure 8A shows a situation in which the allele (represented by *) associated with the trait in question occurs in the coding region of the gene, which translates into a modified gene product. In this case, the mutation driving the pQTL is in cis (i.e., colocated) with the measured gene. This situation results in different gene products, one of which is associated with the trait of interest, but there is no change in the overall expression of the protein. Thus, the pQTL and qQTL will be independent of one another. B: allelic variation in the promoter of a gene. The variation driving the pQTL and the gene in question are in cis; however, the binding affinity of a transcription regulator would be altered, resulting in differential expression of the gene, while the gene product itself remains unchanged. In this case, the pQTL and eQTL colocalize, and the combination QTL/expression approach would be a powerful tool for identifying the gene of interest using these two independent methods. C: the situation in which the variation responsible for differential gene expression is in trans (i.e., a change in a transcription factor results in differential expression in a gene at a separate locus). Although the eQTL and pQTL are not colocated, an analysis of the genes that are differentially expressed may lead to the identification of a common pathway controlled at the transcription level. This form of pathway analysis could link the qQTL and eQTL.

Fig. 8. Possible associations between expression (e) QTL and phenotypic (p) QTL. A: a situation in which the variation associated with the trait in question (represented by *) occurs in the coding region of the gene, which translates into a modified gene product. In this case, the mutation driving the pQTL is in cis (i.e., colocated) with the measured gene. This situation results in different gene products, one of which is associated with the trait of interest, but there is no change in the overall expression of the protein. Thus, the pQTL and qQTL will be independent of one another. B: allelic variation in the promoter of a gene. The variation driving the pQTL and the gene in question are in cis; however, the binding affinity of a transcription regulator would be altered, resulting in differential expression of the gene, while the gene product itself remains unchanged. In this case, the pQTL and eQTL colocalize, and the combination QTL/expression approach would be a powerful tool for identifying the gene of interest using these two independent methods. C: the situation in which the variation responsible for differential gene expression is in trans (i.e., a change in a transcription factor results in differential expression in a gene at a separate locus). Although the eQTL and pQTL are not colocated, an analysis of the genes that are differentially expressed may lead to the identification of a common pathway controlled at the transcription level. This form of pathway analysis could link the qQTL and eQTL.
at a separate locus. Although the eQTL and pQTL are not colocalized, an analysis of the genes that are differentially expressed may lead to the identification of a common pathway controlled at the level of transcription (13), and this form of pathway analysis could link the eQTL and pQTL. Therefore, combining or cross-referencing the results of genomics approaches, such as QTL and expression analyses, can be a powerful way to correlate results, particularly given the relatively high rates of Type I errors inherent in both of these approaches when conducted on their own, but one must be aware of the potential of Type II errors, and data interpretation as well as follow-up studies should be conducted accordingly.

We were especially interested in the results for COUP-TFII, a member of the steroid/thyroid hormone receptor superfamily, because it is not yet possible due to the lack of sequenced hemoglobin genes in any other fish genome studied thus far, and there are several transcript-specific hemoglobin primers for Arctic charr, which would require transcript-specific hemoglobin primers for Arctic charr, which could facilitate the identification of specific hemoglobin genes associated with UTT. Therefore, given that Arctic charr are a cold-adapted species and the results of the study suggest that reduced hemoglobin expression is associated with tolerance to acute heat stress, a full investigation of the Arctic charr hemoglobin repertoire, including sequencing of the genes as well as physiological studies examining the gill capacities of temperature Tolerant and Intolerant fish.

We have recently completed sequencing and annotating the full hemoglobin repertoire of the Atlantic salmon genome (37). There are more hemoglobin transcripts in Atlantic salmon than in any other fish genome studied thus far, and there are several non-Bohr beta-hemoglobin genes in Atlantic salmon. We propose that the gene products act as emergency oxygen supplies under conditions of high stress, such as that of increased temperature, decreased oxygen availability or in conditions where the fish is exerting higher than normal levels of energy (37). Therefore, given that Arctic charr are a cold-adapted species and the results of this study suggest that reduced hemoglobin expression is associated with tolerance to acute heat stress, a full investigation of the Arctic charr hemoglobin repertoire, including sequencing of the genes as well as expression profiling, should be a priority for future research. Such an investigation would allow gene-specific hemoglobin primers to be designed, which could facilitate the identification of specific hemoglobin genes associated with UTT, a discovery with numerous implications for both cultured and wild Arctic charr and other salmonids.

Benefits and limitations of QTL analysis, microarrays, and qPCR for identifying genes governing complex traits. QTL are useful as a starting point for identifying genes that govern a complex trait as they can provide information with regards to the chromosomal region of participating genes, as well as an estimate of the overall contribution of that region to the phenotype in question. However, gaining insight to the actual genes responsible for the QTL is resource intensive, particu-
larly if there is no reference genome sequence available. Even if a genomic sequence is available and annotated, there are usually many genes within a QTL region and further experimentation is required to determine which, if any, of these genes is contributing to the trait in question. Expression profiling (microarray or qPCR analysis) provides a means of identifying eQTL. Cross-referencing pQTL analysis with expression data may reveal whether cis- and trans-controlling elements determine the relative abundance of mRNA for a given gene (11). When pQTL and eQTL coincide, as is the case with the SsaF43NUIG UTT QTL (38, 42) and COUP-TFII in this study, this provides cross-validation that the region, and the element identified therein, indeed contributes to the phenotype in question. However, as illustrated in Fig. 8, there are a number of situations in which a combination of approaches using QTL and expression analysis may not be informative, or may require further examinations, such as pathway analysis, to extract meaningful information.

On its own, microarray analysis acts as an excellent exploratory tool because thousands of genes can be tested at once. However, microarrays based on cDNA clones have inherent drawbacks, such that they are resource intensive and highly susceptible to technical variations as well as statistical pitfalls (25). The latter may result in Type I errors (false positive results), a particular concern when using a large microarray, as the false-positive rate increases as the number of spots increases. Indeed, this was a concern in this study, and might, at least in part, explain the lack of consistency between some of our microarray and qPCR data. Another challenge that we encountered was the potential for cross-hybridization between similar transcripts. It is possible, if not likely, that this occurred with the Hsp genes in the same families, the two COUP-TFII transcripts and the hemoglobin genes identified in our analysis. This may have muted evidence of differential expression between individual transcripts and, instead, produced the genetic expression data (microarray and qPCR) from temperature-tolerant and -intolerant individuals. Our results suggest that a number of sHsps as well as larger HSP90 genes may be associated with tolerance to acute heat exposure. Furthermore, the microarray analysis provided clear evidence that hemoglobin genes (both alpha and beta) are significantly differentially expressed between Tolerant and Intolerant fish. COUP-TFII was identified by QTL sequencing as well as the microarray analysis as a candidate gene, although its specific role is unclear given the subsequent identification of two transcripts, which show different expression patterns. Our results highlight the importance of using more than one approach to identify candidate genes, particularly when examining a complicated trait such as UTT in a species whose genome is highly complex and for which there is no genome sequence or even a suitable reference genome. Specifically, the lack of consistency between our microarray analysis and qPCR results strongly suggests that the results of microarray analysis should be further supported either by qPCR or QTL association and that the use of microarray analysis should be limited to gene explorations by looking for groups of associated genes (e.g., gene families or those with shared pathways) that show similar trends in expression. Further examination of the physiological roles of these genes and their variants is necessary to be able to develop genomic markers associated with them. The results of this study can be incorporated into ongoing broodstock development programs to develop commercial strains of Arctic charr that can withstand warmer growing conditions, as well as used to screen wild populations for sensitivity to climate change, and potentially implement population-specific conservation initiatives.

ACKNOWLEDGMENTS

The authors thank S. Pavey, as well as the members of the Koop Laboratory, particularly K. von Schalburg and B. Sutherland, and members of the Davidson laboratory, particularly K. Lubieniecki and K. Johnstone, for technical assistance and help with data analysis. K. Johnstone designed the EF1A3to4 primers.

GRANTS

This work was supported by funding for the Consortium for Genomic Research on All Salmonids Project (cGRASP) from Genome Canada and Genome BC and by a Strategic Grant from the Natural Sciences and Engineering Research Council of Canada.

DISCLOSURES

C. R. McGowan is the Broodstock Development Manager for Icy Waters Inc. He accepts full responsibility for the conduct of the trial, has full access to all the data, and control over the decision to publish.

REFERENCES

ACUTE HEAT TOLERANCE IN ARCTIC CHARR


42. Somorjai IM, Danzmann RG, Ferguson MM. Distribution of temperature tolerance quantitative trait loci in Arctic char (Salvelinus alpinus) and inferred homologies in rainbow trout (Oncorhynchus mykiss). *Genetics* 165: 1443–1456, 2003.