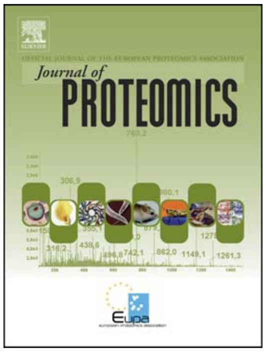


Proteome variance differences within populations of European whitefish (*Coregonus lavaretus*) originating from contrasting salinity environments

Spiros Papakostas, Anti Vasemägi, Mikael Himberg, Craig R. Primmer

PII: S1874-3919(13)00654-4
DOI: doi: 10.1016/j.jprot.2013.12.019
Reference: JPROT 1661

To appear in: *Journal of Proteomics*

Received date: 13 December 2013
Accepted date: 22 December 2013

Please cite this article as: Papakostas Spiros, Vasemägi Anti, Himberg Mikael, Primmer Craig R., Proteome variance differences within populations of European whitefish (*Coregonus lavaretus*) originating from contrasting salinity environments, *Journal of Proteomics* (2014), doi: 10.1016/j.jprot.2013.12.019

Proteome variance differences within populations of European whitefish (*Coregonus lavaretus*) originating from contrasting salinity environments

SPIROS PAPAKOSTAS[1], ANTI VASEMÄGI[1, 2], MIKAEL HIMBERG[3], CRAIG R. PRIMMER[1*]

[1]*Division of Genetics and Physiology, Department of Biology, University of Turku, 20014, Turku, Finland*

[2]*Department of Aquaculture, Institute of Veterinary Medicine and Animal Science, Estonian University of Life Sciences, 51014 Tartu, Estonia*

[3]*Laboratory of Aquatic Pathobiology, Åbo Academy University, 20520, Turku, Finland*

Keywords: Adaptation, Baltic Sea, early-life development, G-protein coupled receptors, label-free proteomics, salinity stress

[*]*Corresponding author at*: Department of Biological Sciences, Division of Genetics and Physiology, Pharmacity, Itäinen Pitkäkatu 4, University of Turku, 20014 Turku, Finland, Tel.: +358 2 333 5571, Fax: +358 2 333 6680, e-mail: craig.primmer@utu.fi (Craig Primmer).

Abstract

Variation in gene expression is an important component of the phenotypic differences observed in nature. Gene expression variance across biological groups and environmental conditions has been studied extensively and has revealed specific genes and molecular mechanisms of interest. However, little is known regarding the importance of within-population gene expression variation to environmental adaptation. To address this issue, we quantified the proteomes of individuals of European whitefish (*Coregonus lavaretus*) from populations that have previously been shown to have adapted during early development to freshwater and brackishwater salinity environments. Using MS-based label-free proteomics, we studied 955 proteins in eight hatch-stage fish embryos from each population to both freshwater and brackishwater salinity conditions. By comparing the levels of within-population protein expression variance over individuals and per protein between populations, we found that fish embryos from the population less affected by salinity had also markedly higher levels of expression variance. Gene Ontologies and molecular pathways associated with osmoregulation showed the most significant difference of within-population proteome variance between populations. Several new candidate genes for salinity adaptation were identified, emphasising the added value of combining assessments of within-population gene expression variation with standard gene expression analysis practices for better understanding the mechanisms of environmental adaptation.

Introduction

Gene expression variation has been regarded as both costly noise[1] as well as a source of variability that organisms may exploit[2, 3]. While recent studies have investigated ways that expression noise is minimised, tolerated or buffered in biological systems[4-6], high levels of gene expression variation have been suggested to be advantageous to the overall population fitness in fluctuating environments[7, 8] or under conditions of stress[9]. Such benefits have been suggested to arise from a bet-hedging strategy that uses stochastic gene expression levels in anticipation of environmental fluctuation[8, 10]. Several genetic[11-13] and non-genetic factors[4, 14] have been recognised to contribute to gene expression variation. Standard expression profiling practice focus on differences of gene expression levels between groups of biological samples exposed to different treatments or originating from differing environmental conditions[15-17]. The vast majority of such studies compare group means in expression levels, while much less attention has been given to the occurrence and potential significance of within-population gene expression variation. However, large inter-individual differences in gene expression variation within populations have in fact been observed in organisms as diverse as teleost fish[18, 19], *Drosophila melanogaster*[20], and humans[21]. Such variation can drive differences in the physiological performance between individuals[19, 22], affect disease susceptibility[23], and therefore may be of functional and biological importance. An open question is whether high within-population gene expression variation can help organisms to adapt in diverse environmental conditions. A first step towards addressing this issue would be to evaluate levels of within-population gene expression variation between populations that show differences in adaptation to a range of environmental conditions.

Here, we assessed the level of within-population protein expression variation in populations of European whitefish that have previously been shown to have divergent proteomic responses to salinity[24]. More specifically, whitefish populations of the Baltic Sea area that spawn in freshwater and brackishwater conditions, subsequently referred to as FW and BW whitefish, were shown to have marked differences in fertilisation success, early-life survival and growth with regard to salinity[24]. Remarkably, BW whitefish embryos developed equally well in salinities ranging from freshwater (0 ppt) to brackishwater conditions (6 to 10 ppt), whereas FW whitefish had much lower survival in higher salt concentrations[24]. Using high-resolution mass spectrometry, we investigated previously the protein expression response at 0 ppt and 10 ppt conditions in each population (8 biological replicates) and identified genes of interest based on analysis of differences between group means i.e. analysis of variance[24]. We found that genes in BW whitefish with significant changes in protein expression levels in relation to salinity conditions were linked with the activity of polyvalent cation receptors, G-protein coupled receptors known to function as salinity sensors in fish[25-27], and of sodium/potassium ATPases involved in sodium ion transport during osmoregulation[24]. In this study, we re-analysed the proteomic dataset specifically focusing at differences of the within-population protein expression variance between the two whitefish populations. For this purpose, we also used a recently described method to estimate significant Gene Ontology and pathway deregulations by comparing, for each annotation in each individual, levels of within-population expression variance [28]. We then compared our findings to those based on group means to assess whether analyses of within-population gene expression variation can be a usefully complementary approach to standard expression profiling methods in elucidating the molecular mechanisms of ecological adaptation.

Materials and methods

Sample collection, common garden experiment, and effect of salinity in early-life development

Full details about the common garden experiment and the effect of salinity on survival and growth are described in Papakostas et al[24]. Briefly, sea-spawning adult whitefish (BW whitefish: *Coregonus lavaretus widegreni*) derived from the Åland Islands in the Baltic Sea (60^{o}18’40.84” N, 20^{o}12’4.46” E). Lake Päijänne-spawning adult whitefish (FW whitefish: *C. lavaretus pallasi*) were collected from individuals kept in the Laukaa hatchery in central Finland. Eggs and milt of nine females and nine males from each population were used for *in vitro* artificial fertilisation. The mating design involved three full-factorial mating matrices in which eggs of three females were fertilised by three different males. Fertilisation was carried out in five different salinities, namely 0, 2, 4, 6, and 10 ppt, to resemble the environments of FW (~0 ppt) and BW (~4 to ~8 ppt) populations. Fertilised eggs from each cross were subsequently split into three replicate clutches, then placed and reared at 6 ^{o}C on Petri dish well plates until hatching. Hatched fry were anaesthetised by an MS-222 overdose and snap-frozen in liquid nitrogen and stored at -80 °C for the proteomic experiment.

Protein extraction, fractionation, digestion, MS acquisition, and protein quantification

Full details are described in Papakostas et al[24]. Briefly, eight hatch-stage samples per salinity per population were randomly selected from two of the common garden conditions (0 ppt and 10 ppt) in both populations. Protein extraction was performed according to an SDS-based protocol. For in-solution digestion, proteins were first digested with LysC and re-digested with trypsin. Peptides were purified on C18-StageTips[29], and fractionated into six fractions (pH 3, 4, 5, 6, 8, and 11) using a SAX-C18 StageTip–based protocol[30]. In nano-LC-coupled mass-spectrometry, each fraction was analysed twice. We used a LTQ Orbitrap classic mass spectrometer (Thermo Electron, Bremen, Germany) equipped with a nanoelectrospray ion source (Proxeon, Odense, Denmark). Peptides were separated with 90 min gradients as follows: pH 3–5 fractions: 8–36% solution B; pH 6 fraction: 8–35% solution B; pH 8 fraction: 5–33% solution B; pH 11 fraction: 2–30% solution B (A: 0.5% acetic acid; B: 0.5% acetic acid/80% acetonitrile). In order to minimise technical variation[31], all samples from different populations and salinity treatments were equally distributed into just two batches and protein extraction, trypsin digestion, fractionation, and Nano-LC-MS/MS were performed at the same time for all the samples of each batch. Samples inserted the mass spectrometer in the order of FW - 0 ppt, FW - 10 ppt, BW - 0 ppt, BW - 10 ppt on a replicate-by-replicate basis as recommended for label-free experiments[32]. Experimental variation, e.g. shifts in the chromatographic gradients, would therefore negate across populations and salinity treatments allowing comparisons of protein expression variation[32, 33].

Proteins were identified and quantified using the MaxQuant v.1.1.1.36 software[34]. The minimum peptide length was set to six amino acids and the maximum false discovery rate (FDR) to 1% for both peptides and proteins. The Atlantic salmon (*Salmo salar*) protein sequences submitted to UniProt (www.uniprot.org, release 2010_11, 9497 entries) were used as a search database. Salmonid UniProt sequences are an efficient resource for this purpose in salmonid species[35]. A list of common contaminants provided with MaxQuant v.1.1.1.36 was included in the search. Protein quantification was based on both unique and 'razor' peptides. 'Razor' peptides are the shared peptides that are most parsimoniously associated with the group that has the highest number of identified peptides[34, 36]. The use of both unique and 'razor' peptides for protein quantification has been suggested to be a good compromise between unequivocal peptide assignment and more accurate quantification[34, 37]. For each fraction, peptides were matched across different LC-MS/MS runs based on mass and retention time ('match between runs' option in MaxQuant) using the default time window of 2 min. Manual inspection of the chromatographic shifts between samples verified the suitability of this option. To increase the quantification accuracy, co-fragmented peptides were distinguished by enabling the second peptide option in MaxQuant[38].

Statistical analyses

The label-free quantification (LFQ) algorithm of MaxQuant performs normalisation across samples after retention time alignment, matching between runs and assembling protein identifications. However, normalisation of the reported LFQ values across biological replicates can be employed to minimise technical variation in label-free experiments (J. Cox, pers. com.). Therefore, the LFQ intensities in each whitefish population were log2-tranformed, loess-normalised using the median values across biological replicates as a reference set.

To estimate the proteome variance within each whitefish population per salinity condition we used eight data points from calculating the variance over all quantified proteins in each biological replicate. Comparisons were then performed using the non-parametric unpaired Wilcoxon rank test (two-tailed). We also compared the variances between groups of samples on a protein-by-protein basis using an F test, as implemented in the *stats* package of R v. 2.15.3. FDR in this case was calculated with the program QVALUE[39]. To compare proteome variances using functional information, we used the PathVar software[28]. For each Gene Ontology (GO) and molecular pathway annotation that involves at least 10 of the quantified proteins, the protein expression variance is calculated for each individual and then group variances are compared using an F test corrected for multiple comparisons with the FDR method[28]. We focused on comparisons between FW and BW whitefish. To ensure equal and maximum amount of valid information, proteins with missing expression values were stripped from the dataset. *Homo sapiens* orthologs, to use in the PathVar software, were identified by performing blastp searches against the Human reference proteome submitted in the UniProt database (release 2013_10, www.unirprot.org). To ensure correct ortholog identification we applied a rather conservative *E* value threshold for blastp ($\leq 6*E^{-14}$)[40]. Proteins assigned to the same human ortholog were not included in the analysis as a precaution against expression variation from alternative isoform regulation[41]. GO and pathway annotations were as of 2012_12. We used pathway information as submitted in the Reactome database (www.reactome.org), a free online database of manually curated, peer-reviewed molecular pathways[42]. To summarise the functional information of significant GO terms we used the program REVIGO[43]. SimRel was used to calculate semantic similarity (default allowed similarity = 0.7), and UniProt as database for GO term sizes. To explore the matrix of semantic similarities we used multidimensional scaling. We used this approach to summarise also the functional information contained in the molecular pathways of interest. For each quantified protein in the pathway we retrieved the GO annotations using the human orthologs and then summarised this information with REVIGO. Allowed similarity was set to a lower level in this case, 0.5, due to the larger amount of GO annotations. Results from the F test for each protein used to assign significance with regard to variance next to each GO term. GO annotations for the human orthologs obtained from the official site of the GO consortium (release 22 November 2013, www.geneontology.org).

Results

Protein expression variation in FW and BW whitefish

Overall, 955 proteins were used to calculate the proteome variance in FW and BW whitefish. Protein quantification based on average on 5.33 unique and 0.49 'razor' peptides (Supplementary Table 1). BW whitefish showed significantly higher proteome variance than FW whitefish in both salinity conditions (Fig. 1; Wilcoxon test: 0 ppt: $P = 0.0011$, 10 ppt: $P = 0.0002$). On a protein-by-protein basis, 176 proteins showed significantly different levels of expression variance between populations (F test: $P < 0.05$, q-value = 0.2038; Supplementary Table 1). Of these proteins, 154 showed higher expression variance in BW whitefish and just 22 had higher expression variance in FW whitefish (Fig. 2; Supplementary Table 1).

There were 21 GO terms associated with proteins that had significantly different levels of within-population expression variance between FW and BW whitefish ($P \leq 0.016$, FDR < 0.05; Supplementary Table 2). *Ion transport, metal ion binding,* and *mitochondrion* were the GO terms for Biological Process, Molecular Function, and Cellular Component GO categories with the highest significance (Table 1). Summarised for semantic similarity, *ion transport* showed further minimum dispensability (Supplementary Table 2) and characterised of one of the groups of semantically similar GO terms for Biological Process (Fig. 3).

We found 31 pathways in the Reactome database that had significantly different levels of within-population expression variance between FW and BW whitefish (Supplementary Table 3). The three most significant pathways were *G-protein coupled receptor downstream signaling* (comprising 138 proteins), *Metabolism* (142 proteins), and *Signaling by G-protein coupled receptors* (144 proteins) (Table 1). These pathways highly overlapped in terms of the proteins they contained (135 proteins in common, Supplementary Table 4) and expression variance was significantly higher in BW whitefish (not shown). GO summation for the proteins in the *G-protein coupled receptor downstream signaling* pathway returned *positive regulation of transferase activity*, *cellular zinc ion homeostastis*, and *protein ubiquitination* as the most significant GO terms (Fig. 4). These GO terms summarised several others including *MAPK cascade* and *regulation of apoptotic process* or *release of sequestered calcium ion into cytosol* and *regulation of membrane potential* (Supplementary Table 5).

Discussion

Based on 955 proteins, we found that BW whitefish had significantly higher levels of proteome variance compared to FW whitefish, regardless of salinity condition (Figs. 1). This result is confirmed by the analysis performed on individual proteins, as 154 proteins showed higher expression variance in BW whitefish vs. 22 proteins with higher variance in FW whitefish (Fig. 2). Functional annotations associated with osmoregulation showed the highest difference of within-population proteome variance between the two populations and higher variance in BW whitefish. *Ion transport*, the most significant and least dispensable GO term for biological process (Fig. 3; Table 1; Supplementary Table 2), and *GPCR downstream signaling*, the most significant molecular pathway (Table 1), are two representative cases. Regulation of ion flux is quintessential for fish during salinity adaptation[44] as GPCR proteins sense environmental salinity and activate downstream signaling [25, 26].

What additional insights were revealed about salinity adaptation in whitefish populations from the assessment of within-population proteome variance? First, we identified new candidate genes for salinity adaptation. The most striking case involved key proteins for ion regulation which in fish are thought to be the Na^+-K^+-ATPase (NKA), the $Na^+/K^+/2Cl^-$ cotransporter (NKCC), and the V-H^+-ATPase ion pumps [44]. Previously, we found that a NKA protein was significantly overexpressed in higher salinity in BW whitefish[24]. Remarkably, V-type proton ATPase subunits showed significant differences in protein expression variance between salinity conditions in BW whitefish (Atlantic salmon UniProt AC: B5X1Y1, B5X274, Supplementary Table 1). Together with NKCC (Atlantic salmon UniProt AC: A9NJG3), V-type proton ATPase subunits were included in GO term *ion transport* with the most significant difference in proteome variance between FW and BW populations (Supplementary Tables 1, and 4). In other words, the study of within-population proteome variance captured the remaining two categories of key proteins for ion regulation. Several other proteins, many of which are novel candidates for salinity adaptation in fish, had significant difference in protein expression variance (Supplementary Table 1). A particular example involves the GrpE protein (Atlantic salmon UniProt AC: B9EM46), a mitochondrial co-chaperone that modulates the function of the heat shock 70 kDa proteins, HSP70[45]. GrpE protein concentration is essential for the cellular machinery capable of repairing stress-induced protein damage [45]. This protein was among those with the highest significance in different levels of protein expression variance between FW and BW whitefish (P = 1.13E-05, q-value = 0.0012).

Second, by comparing levels of within-population pathway expression variance we confirmed the importance of *GPCR downstream signaling* in salinity adaptation for the whitefish populations. Previously, interaction network analysis on the proteins with significantly different levels of protein expression between salinities provided only indirect evidence on the activity of GPCR pathway. Specifically, we recognised a central position in the network for a highly conserved pleiotropic cytokine, tumour necrosis factor, suggested to play a critical role in salinity adaptation and regulated by the activity of GPCR proteins that act as salinity sensors in fish[24]. In this study, *GPCR downstream signaling* showed the most significant difference in pathway expression variance between whitefish populations (Table 1). Studying further the proteins in this pathway in terms of their levels of difference in expression variance, we found more evidence suggesting that *GPCR downstream signaling*, and in this regard differences in expression variance, are of relevance for salinity adaptation in the whitefish populations. *Cellular zinc ion homeostasis* and *positive regulation of transferase activity* (Fig. 4) summarised *stress-activated MAPK cascade*, *apoptotic process*, *regulation of membrane potential*, and *release of sequestered calcium ion into cytosol* (Supplementary Table 5) with known roles during osmoregulation. For example, salinity sensors in fish are calcium receptors [25, 26] and these receptors have evolutionary conserved functional domains for MAPK activation[27]. Apoptosis is a fundamental process in fish during osmoregulation[46, 47].

Although the assessment of just two populations precludes drawing firm conclusions about the potential origins of the clear difference in within population expression variation, several hypotheses can be proposed for testing in future research. Firstly, variable environments may select for higher gene expression variance in organisms[9], particularly in molecular mechanisms that sense environmental fluctuations. In yeast for example, sequence variants in environmental sensor genes have been identified as loci that can dramatically increase cell-to-cell gene expression variability of downstream pathways without necessarily changing mean expression [48, 49]. Following these observations, our data are in line with the notion that BW whitefish may have adapted to the fluctuating salinity conditions in the Baltic Sea compared to FW whitefish that inhabit a more stable environment in terms of salinity. In any case, a greater number of freshwater and marine populations would need to be tested before any conclusions can be drawn. It would also be interesting to include anadromous whitefish populations that experience both brackish- and freshwater environments during different phases of their life history, in such a study. Secondly, the FW whitefish studied here originate from a hatchery-reared stock, which could also potentially have had an effect on protein expression variance. There are in fact several factors related to hatchery rearing that may be expected to have contrasting effects on expression variance. On the one hand, lower effective population sizes and drift in farmed whitefish could reduce levels of gene expression variation. On the other, fish of hatchery origin may accumulate slightly deleterious mutations because of relaxed purifying selection in benign farmed environment[50] which would increase gene expression variation. The latter, however, is not supported by our findings. Thirdly, natural differences in the level of genetic diversity could also potentially affect within-population expression variation[51]. Microsatellite data from nine markers (data from [52]) show slightly higher levels of mean genetic variation in BW than in FW whitefish i.e. towards the same direction as proteome variance, albeit non-significant (Allelic richness = 8.411 and 7.525, respectively, $P = 0.09$) thus suggesting further research to assess this possibility in more detail may be warranted.

We assessed gene expression variance directly at the protein level as that it is ultimately the most relevant functional measurement[53, 54]. Proteins are the typical effectors of biological function and variance in protein levels has been shown repeatedly that cannot fully explained by mRNA abundance[53, 55, 56]. However, label-free methods are known to have low accuracy in protein quantification and are prone to the introduction of experimental variation that may be a possible confounding factor[32, 33]. Regardless, it is very unlikely that our observations were the result of any systematic error in quantification. Mixing between samples as recommended for label-free experiments and good depth in biological replication coupled with technical replication ensured that any technical variation would have the same impact on each population and condition.

In conclusion, by quantifying hundreds of proteins in a well-described system of salmonid fish with differences in early development to salinity environments, we demonstrate that within-population gene expression variance can provide useful information as a complementary approach to more standard methods of gene expression profiling. Our analysis returned several interesting candidates, many of which with a previously undescribed role, for salinity adaptation in fish. Pathway expression variance seems a promising approach as pathway perturbations may be of potential evolutionary and ecological relevance.

Acknowledgements

Benny, Hans and Helge Holmström assisted in getting the fish samples as well as Juha-Pekka Vähä helped in the common garden experiment and the survival and growth analyses. Liisa Arike performed the Nano-LC-MS ⁄MS runs and Lauri Peil assisted in protein quantification. Matthieu Bruneaux has helped with some of the bioinformatics analyses used in this work. The study was supported by the European Community’s Seventh Framework Programme (FP ⁄ 2007-2013) under grant agreement no. 217246, made with the joint Baltic Sea research and development programme of the Baltic Organisations Network for Funding Science (BONUS), as well as by the Academy of Finland and the Estonian Science Foundation (grant 6802).

Figure legends

Fig. 1 Box plots of the calculated variance of normalised expression levels across 955 proteins in eight hatch-stage larvae from the FW and BW whitefish population grown at 0 ppt and 10 ppt salinities. $^{**}P < 0.01$, $^{***}P < 0.001$

Fig. 2 Scatter plot of normalised protein expression variance in FW and BW whitefish for 955 proteins. Colours indicate significantly higher variance in FW (green, 22 proteins) and BW (red, 154 proteins) whitefish ($P < 0.05$).

Fig. 3 Scatter plot of the GO terms for biological process summarised by employing multidimensional scaling to the matrix of semantic similarities. Colour is proportional to the *P* value of difference in protein expression variance between FW and BW whitefish. Size is proportional to the term frequency in the GO annotation database, also indicates general terms. Names present the most significant terms in each group of semantically similar terms.

Fig. 4. Scatter plot of the GO terms for biological process for the genes of the *GPCR downstream signaling* pathway with significant difference in protein expression variance between FW and BW whitefish, summarised by multidimensional scaling on the semantic similarity matrix. Colour is proportional to the *P* value of difference in protein expression variance between FW and BW whitefish. Size is proportional to the term frequency in the GO annotation database. Names indicate the most significant terms in each part of two-dimensional semantic space.

Tables

Table 1. The three most significant GO terms per GO category and molecular pathways with different expression variance between FW and BW whitefish

Annotation	Proteins in the dataset	*P*-value	FDR
GO Biological Process			
Ion transport (GO: 0006811)	14	1.30E-04	0.02
RNA splicing (GO: 0008380)	49	4.90E-04	0.00
Proteolysis (GO: 0006508)	25	8.10E-04	0.01
GO Molecular Function			
Metal ion binding (GO: 0046872)	77	4.20E-05	0.00
Protein binding (GO: 0005515)	355	6.40E-05	0.00
Isomerase activity (GO: 0016853)	21	2.00E-04	0.00
GO Cellular Component			
Mitochondrion (GO: 0005739)	129	1.00E-05	0.01
Endoplasmic reticulum (GO: 0005783)	55	1.50E-05	0.01
Cytoplasm (GO: 0005737)	294	4.00E-04	0.00
Molecular Pathway			
GPCR downstream signaling	138	1.20E-04	0.02
Metabolism	142	1.20E-04	0.02
Signaling by GPCR	144	1.40E-04	0.02

References

[1] Wang Z, Zhang JZ. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:E67-E76.

[2] Whitehead A, Crawford DL. Variation within and among species in gene expression: raw material for evolution. Molecular Ecology. 2006;15:1197-211.

[3] Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM. Common pattern of evolution of gene expression level and protein sequence in Drosophila. Molecular Biology and Evolution. 2004;21:1308-17.

[4] Raj A, van Oudenaarden A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell. 2008;135:216-26.

[5] Queitsch C, Sangster TA, Lindquist S. Hsp90 as a capacitor of phenotypic variation. Nature. 2002;417:618-24.

[6] Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et al. Single-cell proteomic analysis of S-cerevisiae reveals the architecture of biological noise. Nature. 2006;441:840-6.

[7] Thattai M, van Oudenaarden A. Stochastic gene expression in fluctuating environments. Genetics. 2004;167:523-30.

[8] Kussell E, Leibler S. Phenotypic diversity, population growth, and information in fluctuating environments. Science. 2005;309:2075-8.

[9] Blake WJ, Balazsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, et al. Phenotypic consequences of promoter-mediated transcriptional noise. Molecular Cell. 2006;24:853-65.

[10] Muller J, Hense BA, Fuchs TM, Utz M, Potzsche C. Bet-hedging in stochastically switching environments. Journal of Theoretical Biology. 2013;336:144-57.

[11] Ge B, Pokholok DK, Kwan T, Grundberg E, Morcos L, Verlaan DJ, et al. Global patterns of cis variation in human cells revealed by high-density allelic expression analysis. Nature Genetics. 2009;41:1216-U78.

[12] Borneman AR, Gianoulis TA, Zhang ZDD, Yu HY, Rozowsky J, Seringhaus MR, et al. Divergence of transcription factor binding sites across related yeast species. Science. 2007;317:815-9.

[13] Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature. 2010;464:768-72.

[14] Chalancon G, Ravarani CNJ, Balaji S, Martinez-Arias A, Aravind L, Jothi R, et al. Interplay between gene expression noise and regulatory network architecture. Trends in Genetics. 2012;28:221-32.

[15] Gibson G. Microarrays in ecology and evolution: a preview. Molecular Ecology. 2002;11:17-24.

[16] Ranz JM, Machado CA. Uncovering evolutionary patterns of gene expression using microarrays. Trends in Ecology & Evolution. 2006;21:29-37.

[17] Gibson G. The environmental contribution to gene expression profiles. Nature Reviews Genetics. 2008;9:575-81.

[18] Oleksiak MF, Churchill GA, Crawford DL. Variation in gene expression within and among natural populations. Nature Genetics. 2002;32:261-6.

[19] Rees BB, Andacht T, Skripnikova E, Crawford DL. Population Proteomics: Quantitative Variation Within and Among Populations in Cardiac Protein Expression. Molecular Biology and Evolution. 2011;28:1271-9.

[20] Hutter S, Saminadin-Peter SS, Stephan W, Parsch J. Gene expression variation in African and European populations of Drosophila melanogaster. Genome Biology. 2008;9.

[21] Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Akey JM. Gene-expression variation within and among human populations. American Journal of Human Genetics. 2007;80:502-9.

[22] Oleksiak MF, Roach JL, Crawford DL. Natural variation in cardiac metabolism and gene expression in Fundulus heteroclitus. Nature Genetics. 2005;37:67-72.

[23] Li JJ, Liu Y, Kim T, Min RQ, Zhang ZL. Gene Expression Variability within and between Human Populations and Implications toward Disease Susceptibility. Plos Computational Biology. 2010;6.

[24] Papakostas S, Vasemagi A, Vaha J-P, Himberg M, Peil L, Primmer CR. A proteomics approach reveals divergent molecular responses to salinity in populations of European whitefish (Coregonus lavaretus). Molecular Ecology. 2012;21:3516-30.

[25] Nearing J, Betka M, Quinn S, Hentschel H, Elger M, Baum M, et al. Polyvalent cation receptor proteins (CaRs) are salinity sensors in fish. Proceedings of the National Academy of Sciences of the United States of America. 2002;99:9231-6.

[26] Chang WH, Shoback D. Extracellular Ca2+-sensing receptors - an overview. Cell Calcium. 2004;35:183-96.

[27] Loretz CA. Extracellular calcium-sensing receptors in fishes. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology. 2008;149:225-45.

[28] Glaab E, Schneider R. PathVar: analysis of gene and protein expression variance in cellular pathways using microarray data. Bioinformatics. 2012;28:446-7.

[29] Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols. 2007;2:1896-906.

[30] Wisniewski JR, Zougman A, Mann M. Combination of FASP and StageTip-Based Fractionation Allows In-Depth Analysis of the Hippocampal Membrane Proteome. Journal of Proteome Research. 2009;8:5674-8.

[31] Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nature Reviews Genetics. 2010;11:733-9.

[32] Bantscheff M, Lemeer S, Savitski MM, Kuster B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Analytical and Bioanalytical Chemistry. 2012;404:939-65.

[33] Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Analytical and Bioanalytical Chemistry. 2007;389:1017-31.

[34] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology. 2008;26:1367-72.

[35] Papakostas S, Vollestad LA, Primmer CR, Leder EH. Proteomic Profiling of Early Life Stages of European Grayling (Thymallus thymallus). Journal of Proteome Research. 2010;9:4790-800.

[36] Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data - The protein inference problem. Molecular & Cellular Proteomics. 2005;4:1419-40.

[37] Hubner NC, Mann M. Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). Methods. 2011;53:453-9.

[38] Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. Journal of Proteome Research. 2011;10:1794-805.

[39] Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:9440-5.

[40] Primmer CR, Papakostas S, Leder EH, Davis MJ, Ragan MA. Annotated genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and evolution research. Molecular Ecology. 2013;22:3216-41.

[41] Wang ET, Sandberg R, Luo SJ, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008;456:470-6.

[42] Croft D, O'Kelly G, Wu GM, Haw R, Gillespie M, Matthews L, et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Research. 2011;39:D691-D7.

[43] Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. Plos One. 2011;6.

[44] Hwang PP, Lee TH. New insights into fish ion regulation and mitochondrion-rich cells. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology. 2007;148:479-97.

[45] Kampinga HH, Craig EA. `The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nature Reviews Molecular Cell Biology. 2010;11:579-92.

[46] Sakamoto T, McCormick SD. Prolactin and growth hormone in fish osmoregulation. General and Comparative Endocrinology. 2006;147:24-30.

[47] Whitehead A, Zhang SJ, Roach JL, Galvez F. Common functional targets of adaptive micro- and macro-evolutionary divergence in killifish. Molecular Ecology. 2013;22:3780-96.

[48] Fehrmann S, Bottin-Duplus H, Leonidou A, Mollereau E, Barthelaix A, Wei W, et al. Natural sequence variants of yeast environmental sensors confer cell-to-cell expression variability. Molecular Systems Biology. 2013;9.

[49] Brown KM, Landry CR, Hartl DL, Cavalieri D. Cascading transcriptional effects of a naturally occurring frameshift mutation in Saccharomyces cerevisiae. Molecular Ecology. 2008;17:2985-97.

[50] Araki H, Berejikian BA, Ford MJ, Blouin MS. Fitness of hatchery-reared salmonids in the wild. Evolutionary Applications. 2008;1:342-55.

[51] Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. Nature Genetics. 2007;39:1217-24.

[52] Lovisa Wennerström LL, Nils Ryman, Fred M. Utter, Nurul Izza Ab Ghani, Carl André, Jacquelin DeFaveri, Daniel Johansson, Lena Kautsky, Juha Merilä, Natalia Mikhailova, Ricardo Pereyra, Annica Sandström, Amber G. F. Teacher, Roman Wenne, Anti Vasemägi, Małgorzata Zbawicka, Kerstin Johannesson, Craig R. Primmer. Genetic biodiversity in the Baltic Sea: species-specific patterns challenge management. Biodiversity and Conservation. 2013;22:3045-65.

[53] Cox J, Mann M. Is proteomics the new genomics? Cell. 2007;130:395-8.

[54] Altelaar AFM, Munoz J, Heck AJR. Next-generation proteomics: towards an integrative view of proteome dynamics. Nature Reviews Genetics. 2013;14:35-48.

[55] Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, Goodlett DR, et al. Genetic basis of proteome variation in yeast. Nature Genetics. 2007;39:1369-75.

[56] Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. Nature. 2011;473:337-42.

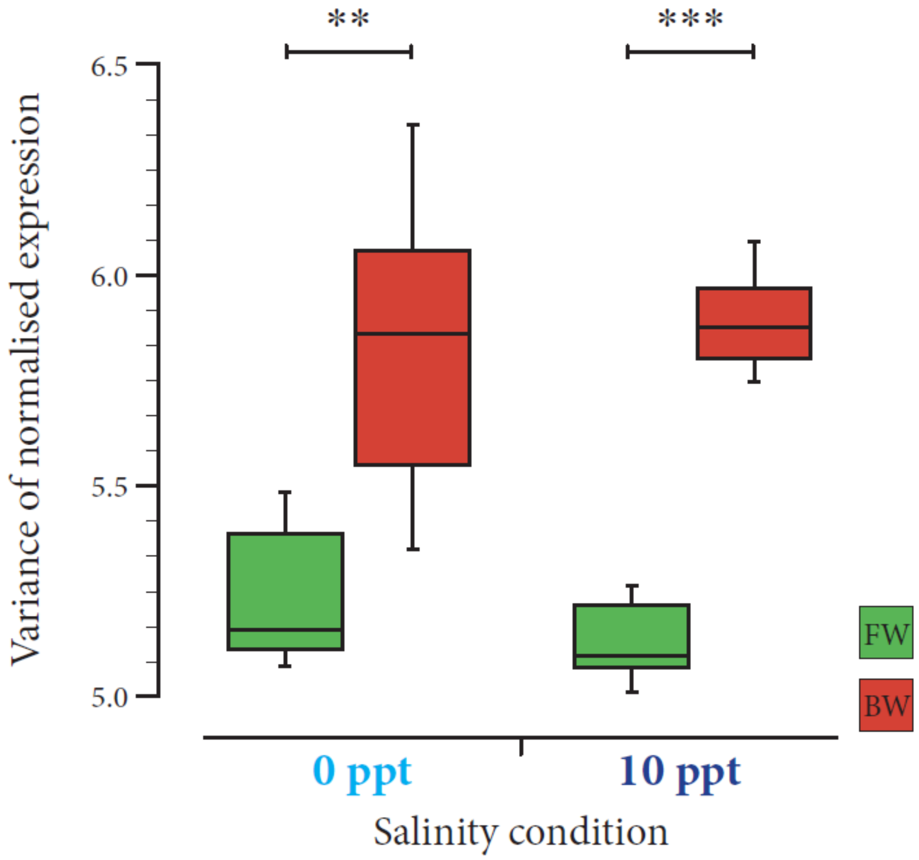

**
***
6.5
6.0
5.5
5.0
Variance of normalised expression
FW
BW
0 ppt
10 ppt
Salinity condition

Figure 1

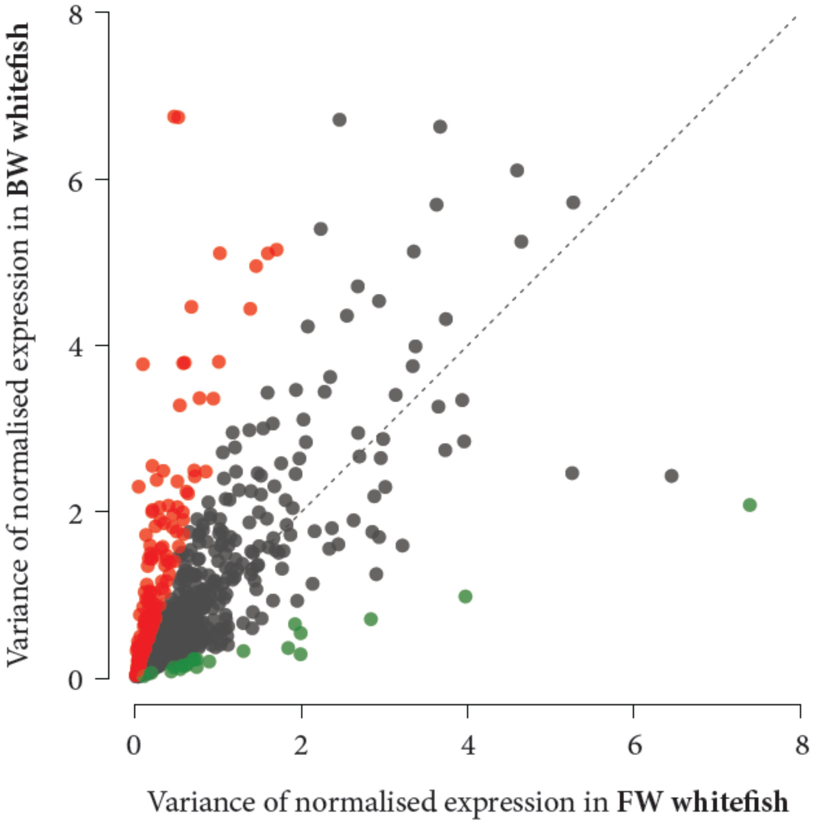

Variance of normalised expression in **BW whitefish**
Variance of normalised expression in **FW whitefish**
0
2
4
6
8
0
2
4
6
8

Figure 2

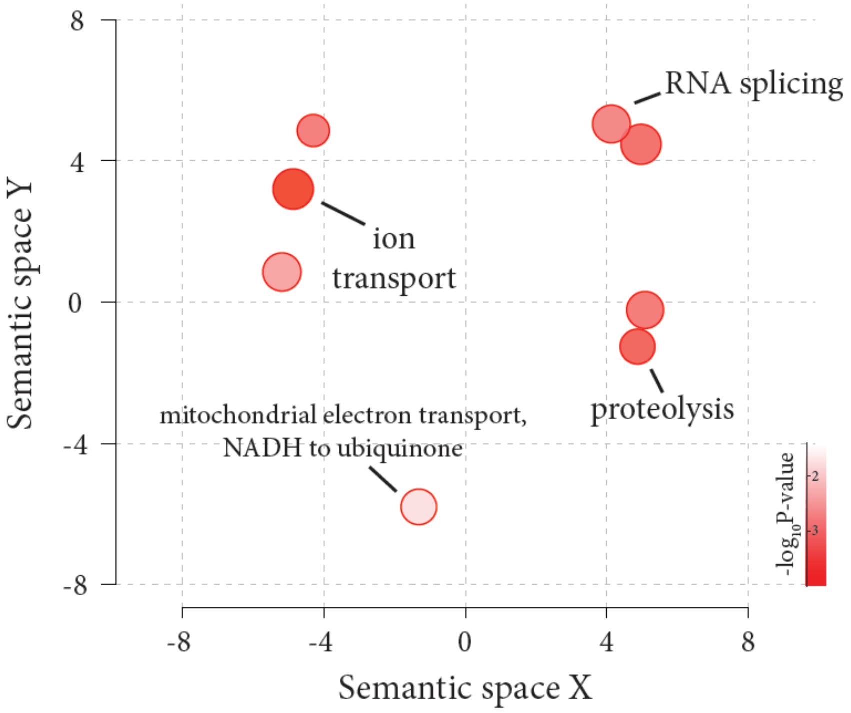
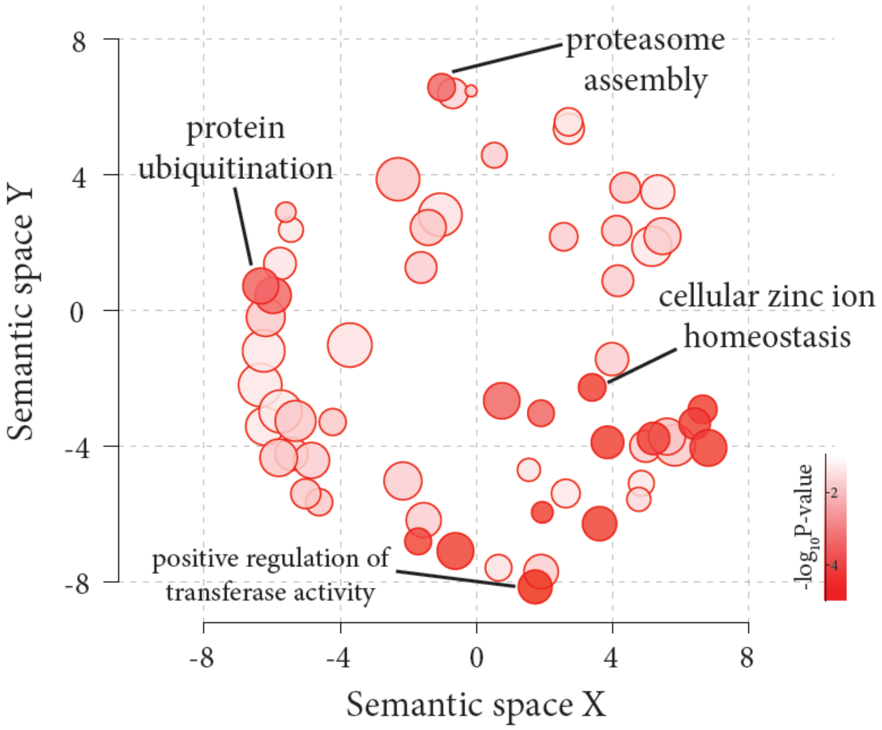

8
4
0
-4
-8
Semantic space Y
RNA splicing
ion
transport
proteolysis
mitochondrial electron transport,
NADH to ubiquinone
-log10P-value
-2
-3
-8
-4
0
4
8
Semantic space X

Figure 3

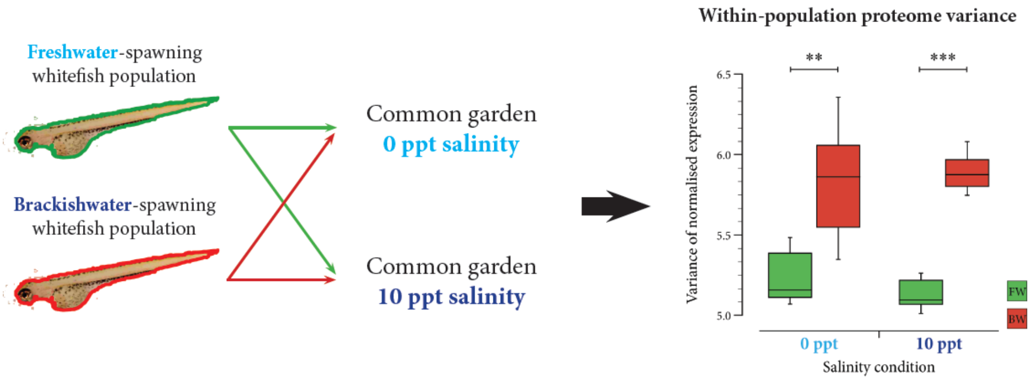


proteasome assembly
protein ubiquitination
cellular zinc ion homeostasis
positive regulation of transferase activity
Semantic space Y
Semantic space X
8
4
0
-4
-8
-8
-4
0
4
8
$-\log_{10}$P-value
-2
-4

Figure 4

Biological significance

We demonstrate the benefits of studying within-population gene expression variance together with more typical methods of gene expression profiling. Proteome variance differences within European whitefish populations originating from different salinity environments allowed us to identify several new candidate genes for salinity adaptation in teleost fish and generate many further hypotheses to be tested.



Graphical Abstract

Highlights

- We study proteome expression in whitefish larvae from contrasting salinity environments
- Larvae were raised in a common garden setting at different salinities
- Proteome variance is higher in larvae originating from brackishwater than freshwater
- We identify new candidate genes for salinity adaptation in teleost fish
- Within-population gene expression variation complements gene expression profiling