Fish Physiol Biochem (2015) 41:1403–1417
DOI 10.1007/s10695-015-0095-0

Proteomic analysis of early-stage embryos: implications for egg quality in hapuku (*Polyprion oxygeneios*)

Yair Y. Kohn · Jane E. Symonds · Torsten Kleffmann · Shinichi Nakagawa · Malgorzata Lagisz · P. Mark Lokman

Received: 30 December 2014 / Accepted: 6 July 2015 / Published online: 18 July 2015

Abstract In order to develop biomarkers that may help predict the egg quality of captive hapuku (*Polyprion oxygeneios*) and provide potential avenues for its manipulation, the present study (1) sequenced the proteome of early-stage embryos using isobaric tag for relative and absolute quantification analysis, and (2) aimed to establish the predictive value of the abundance of identified proteins with regard to egg quality through regression analysis. Egg quality was determined for eight different egg batches by blastomere symmetry scores. In total, 121 proteins were identified and assigned to one of nine major groups according to their function/pathway. A mixed-effects model analysis revealed a decrease in relative protein abundance that correlated with (decreasing) egg quality in one major group (heat-shock proteins). No differences were found in the other protein groups. Linear regression analysis, performed for each identified protein separately, revealed seven proteins that showed a significant decrease in relative abundance with reduced blastomere symmetry: two correlates that have been named in other studies (vitellogenin, heat-shock protein-70) and a further five new candidate proteins (78 kDa glucose-regulated protein, elongation factor-2, GTP-binding nuclear protein Ran, iduronate 2-sulfatase and 6-phosphogluconate dehydrogenase). Notwithstanding issues associated with multiple statistical testing, we conclude that these proteins, and especially iduronate 2-sulfatase and the generic heat-shock protein group, could serve as biomarkers of egg quality in hapuku.

Keywords Hapuku · *Polyprion oxygeneios* · Egg quality · Proteomics · iTRAQ · Biomarkers

Y. Y. Kohn · P. M. Lokman (✉)
Department of Zoology, University of Otago, PO Box 56, 340 Great King St., Dunedin 9016, New Zealand
e-mail: mark.lokman@otago.ac.nz

Y. Y. Kohn · J. E. Symonds
Bream Bay Aquaculture Park, NIWA, PO Box 147, Ruakaka 0151, New Zealand

Present Address:
Y. Y. Kohn
Arava Research and Development Station, Hatzeva, Israel

T. Kleffmann
Centre for Protein Research, University of Otago, PO Box 56, Dunedin 9054, New Zealand

S. Nakagawa · M. Lagisz
Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

Introduction

Existing approaches for determining egg quality focus on either measuring parameters that can predict the viability of fertilized eggs (e.g. blastomere morphology, Shields et al. 1997) or examining the abundance

and composition of certain compounds (e.g. vitamins, Palace and Werner 2006; fatty acids, Salze et al. 2005) that may affect egg quality. However, how these egg quality measurements relate to the normal development of fertilized eggs is not well understood.

The ovulated egg is transcriptionally inactive, and rapid early postfertilization development is strongly dependent on stored proteins and transcripts (see review by Li et al. 2013). As such, these stored entities are likely to have a notable bearing on the 'quality' of the egg, which has been exploited in fish through transcriptomics (Ma et al. 2012; Lanes et al. 2013; Chapman et al. 2014) and proteomics studies (Rime et al. 2004; Crespel et al. 2008; Papakostas et al. 2010; Castets et al. 2012). The first study to evaluate both the transcriptome and the proteome of essentially the same sample of fully grown ovarian zebrafish follicles concluded that both omics approaches yielded very different information (Knoll-Gellida et al. 2006). Moreover, the transcriptome was not predictive of protein abundance in the proteome, highlighting the need for both proteome and transcriptome analyses. Presently, transcriptomics studies on ovulated eggs from oviparous vertebrates seem to be more numerous than comparable proteomics studies, warranting proteomics analyses to gain novel insights into egg quality during very early embryonic development.

A global proteomic approach aims at the identification, and often quantification, of entire protein complements in various biological systems and thus provides insights into the relationships of proteins within relevant processes such as metabolic and regulatory pathways (Cerda et al. 2008). Furthermore, as a global profiling approach, proteomics is complementary to (rather than repeating) mRNA expression profiling at a different level (c.f., Knoll-Gellida et al. 2006). Often mRNA levels do not fully represent the processes researchers aim to investigate. Discrepancies between corresponding mRNA and protein levels are apparent and may depend on various regulatory mechanisms such as translational regulation, protein stability and protein transport processes (Anderson and Anderson 1998; Martyniuk and Denslow 2009; Schwanhäusser et al. 2011).

Proteomic methods have only recently been applied to questions related to embryogenesis, toxicology, biomedicine, aquaculture, environmental effects and food safety in fish (Reviewed in Forné et al. 2010; Sanchez et al. 2011; Rodrigues et al. 2012; Lokman and Symonds 2014). Most of these studies illustrate the power of proteomic methods in describing the relevant processes that influence development and the differences between normal embryos and those that have been exposed to toxic agents. These studies were undertaken mostly in zebrafish (*Danio rerio*) and salmonids (reviewed in Forné et al. 2010). To date, most studies have used 2D gel electrophoresis-based approaches (Forné et al. 2010). However, gel-free protein quantification by mass spectrometry using stable isotope labelling techniques, such as iTRAQ, allows *simultaneous* analysis of up to eight samples that have been labelled differentially with unique isobaric tags. The iTRAQ approach facilitates not only the differential profiling of relative protein abundances but also simultaneous global protein identification and therefore provides information of regulated and nonregulated proteins. Such a global picture of protein expression is usually not achieved by gel-based approaches. (Zieske 2006). Indeed, iTRAQ was recently viewed as a sensitive technique that can be employed very successfully even when full genome sequence data are lacking (Martyniuk et al. 2012).

Of the studies mentioned above, only a few are directly related to oocyte and egg quality, and these are all gel-based studies. Differences were identified in the abundance of various proteins, albeit without multiple statistical testing, in oocytes (Yoshikuni et al. 2003), coelomic fluid (Rime et al. 2004), fertilized eggs (Crespel et al. 2008) and larvae (Castets et al. 2012), and the suitability of proteomics in identifying protein targets that are involved in processes affecting egg and larval quality in fish could be demonstrated. However, due to the complexity of this biological system, more comprehensive approaches are required to gain deeper insights into the underlying mechanisms.

In hapuku (*Polyprion oxygeneios*) aquaculture, fertilized eggs show variable quality with some batches containing mostly poorly developing embryos, as assessed by blastomere morphology scoring (Kohn and Symonds 2012). Elucidating the biochemical processes that lead to poor development could benefit the reproductive success of this species in captivity and may be a step forward in addressing the common constraint of poor egg quality in finfish aquaculture.

The main aims of the present study were to (1) describe the proteome of early-stage embryos of hapuku with regard to protein functions and biochemical pathways and to (2) correlate the abundance of

these proteins with blastomere symmetry to find potential target proteins/pathways that are associated with egg quality.

Materials and Methods

Sampling

Hapuku broodstock were reared at the National Institute of Water and Atmospheric Research (NIWA) aquaculture facility in Bream Bay, Ruakaka, New Zealand. Hapuku are serial batch spawners, and each spawning female produces multiple batches of eggs during the spawning season in spring. Wild-caught broodstock of both sexes, most of which had spawned in previous years, were housed in communal tanks and allowed to spawn without intervention. In 2009, fertilized eggs were collected from two broodstock tanks (11 females and 7 males in one tank; 6 females and 9 males in another) and in 2011 from one broodstock tank (9 females and 8 males). As parentage analyses were not conducted, the contribution of the specific individuals to the batches analysed is not known. The protocols for fish rearing and sampling of fertilized eggs were described in Kohn and Symonds (2012).

Eight batches of fertilized eggs (at the 8-cell embryo stage) that exhibited differences in blastomere morphology scores (Kohn and Symonds 2012) were selected during the 2009 ($n = 3$ batches, 30 embryos/batch) and 2011 ($n = 5$ batches, $\sim$100 embryos/batch) spawning seasons. The embryos were photographed, and the blastomere symmetry was given a quality score of between '0', being the *best quality* with perfect symmetry, and '4' representing the *worst quality*, for completely asymmetric blastomere patterning—in doing so, we obtained egg batches from hapuku (Fig. 1) that ranged between 'moderately good' (score = 1.50) to 'bad' (3.50); we refer to these batches in italics text (e.g. Table 1) as good (G), moderate (M) and bad (X), followed by the quality score (e.g. *M2.45* represents a batch of moderate quality with a score of 2.45). All methods used followed those detailed in Kohn and Symonds (2012), except that only symmetry was scored as a measure of egg quality, as validated previously (Kohn and Symonds 2012). The collected eggs were snap-frozen in liquid nitrogen and stored at −70 °C prior to protein extraction.

Protein extraction

Approximately 500 mg of 8-cell-stage embryos from each batch were placed in liquid nitrogen and promptly crushed using a pestle (BIO PLAS, USA) and homogenized in ×4 (v/w) phosphate buffer (10 mM, pH 7.4) and ×1 (v/w) protease inhibitor (×10, Complete Mini, Roche, Germany). Samples were then homogenized further using an electric homogenizer (IKA T25B Ultra TURRAX, Malaysia) and finally sonicated (Ultra Tip, Wave Energy Systems, USA) for 3 × 30 s. Homogenized samples were centrifuged at 10,600*g* (25 min, 4 °C), and the supernatant, excluding the floating lipid fraction, was retained for iTRAQ analysis (see below). Protein contents of the supernatants were estimated using a bicinchoninic acid assay (Sigma-Aldrich, http://www.sigmaaldrich.com) on a plate reader (FLUOstar Omega, BMG Labtech, Germany).

iTRAQ analysis

The 8-plex iTRAQ method (Applied Biosystems, USA) used in the present study utilized eight different stable isotope-labelled isobaric tags that modify peptides (after proteins were trypsin-digested) at the N terminus and the lysine side chain, as well as the tyrosine side chain at a lower frequency. After differential labelling (50 µg protein/sample), all eight samples were pooled and analysed by liquid chromatography-coupled tandem mass spectrometry. Due to the isobaric nature of the tag, the same peptides from different samples behave as one species in liquid chromatography, ionization and mass measurement. Upon peptide fragmentation by collision-induced dissociation (CID)-based ion activation, however, tag-specific reporter fragment ions are released and recorded in the fragment ion spectrum. The relative intensities of the eight different reporter fragment ions are used for quantification and other peptide-specific fragment ions for identification (Ross et al. 2004). Accordingly, the iTRAQ approach only reveals relative protein quantities unless the absolute amount of proteins is known for one of the samples. In this study, we focused our attention on soluble proteins from 8-cell-stage embryos. Although the iTRAQ approach is suitable for the quantification of membrane proteins, our protein extraction was biased towards water-soluble proteins. In future studies, the screen for

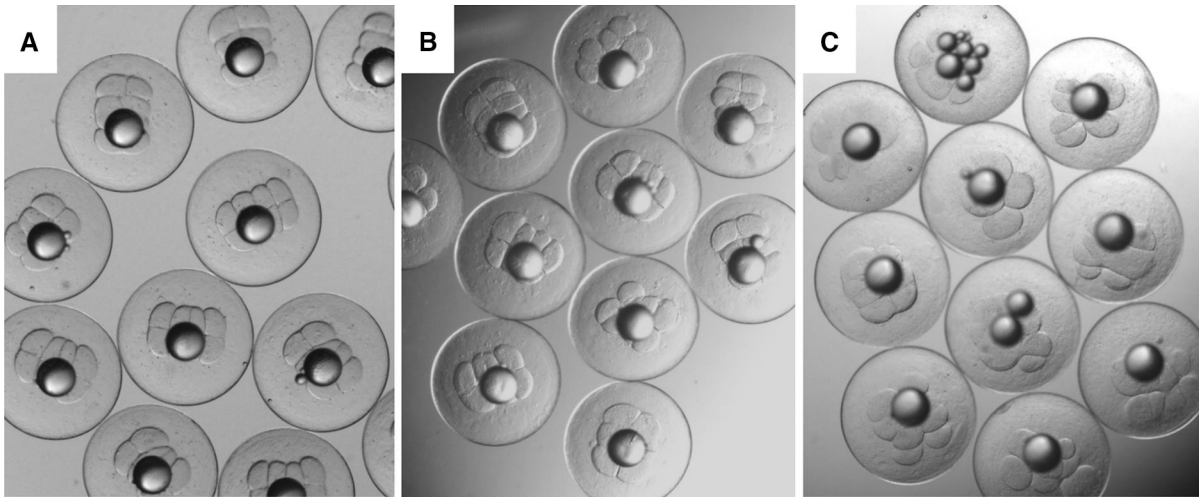


Fig. 1 Batches of fertilized eggs from hapuku, *P. oxygeneios*, displaying blastomeres with different degrees of symmetry. Eggs were scored as moderately good (**a**; quality score = *G1.77*), moderate (**b**; quality score = *M2.45*) or bad (**c**; quality score = *X3.50*). For details on egg quality scores, see text

Table 1 Blastomere quality scores ('quality') of the eight batches of eggs from hapuku (*P. oxygeneios*) used for iTRAQ proteomic analysis

Batch ID[a]	113	114	115	116	117	118	119	121
n	30	30	100	96	96	30	99	97
Score[b]	1.50 ± 0.24	1.70 ± 0.19	1.77 ± 0.08	2.44 ± 0.09	2.45 ± 0.06	2.93 ± 0.16	3.31 ± 0.08	3.50 ± 0.06
Quality[c]	G	G	G	M	M	X	X	X

SE standard error

[a] Different batches of eggs are identified on the basis of their isobaric iTRAQ tag. All peak intensities are standardized against sample 113

[b] Quality scores of *n* fertilized eggs are expressed as means ± standard errors

[c] Egg quality was moderately good (G), moderate (M) or bad (X)

potential target proteins may be extended to membrane protein fractions.

Filter-aided sample preparation (FASP) and iTRAQ labelling

All samples were processed using a filter-aided sample preparation (FASP, Wiśniewski et al. 2009) procedure that was modified for compatibility with downstream iTRAQ labelling. In brief, supernatants from each homogenized sample were loaded into 0.5 ml ultrafiltration units (10 kDa cut-off) and buffer-exchanged into 8 M urea in 0.25 M triethylammonium bicarbonate (TEAB) by at least three consecutive centrifugation steps, each reducing the volume down to 20 µl. Samples were then reduced and alkylated on-filter using the reagents from the iTRAQ kit and buffer-exchanged to 0.5 M TEAB by two consecutive centrifugation steps. An aliquot of each sample was taken for protein measurement using the Bradford method. Proteins were digested with trypsin (1 µg trypsin/20 µg protein) on-filter over night at 37 °C. After the first digestion, an additional aliquot of trypsin was added to each sample and samples were incubated for another 4 h. The entire filter content was then transferred to a new tube, and peptides were labelled with the iTRAQ reagents (8-plex kit) according to the manufacturer's instructions. Samples were pooled following the protein measurement.

Peptide fractionation and LC–MS/MS analysis

To increase the number of peptide identifications and quantification counts per protein in the complex protein extracts of 8-cell-stage embryos, we implemented a two-dimensional peptide prefractionation using off-line peptide isoelectric focusing followed by uHPLC inline coupled to tandem mass spectrometry, a procedure that is compatible with iTRAQ quantification (Chenau et al. 2008). Therefore, peptides of the pooled iTRAQ-labelled sample were subjected to off-gel isoelectric focusing using an Agilent OFFGEL Fractionator (Agilent Technologies) and separated into 24 fractions along a pH gradient from 4 to 10 according to the manufacturer's instructions. Fractions were concentrated using a centrifugal vacuum concentrator and analysed in duplicate by nanoflow liquid chromatography-coupled LTQ Orbitrap tandem mass spectrometry (LC–Orbitrap MS/MS). One half of each concentrated fraction was loaded onto a trap column and separated on an in-house emitter tip column (75 um ID PicoTip fused silica tubing (New Objectives, USA) packed with C-18 material on a length of 12 cm) through an acetonitrile gradient from 5 to 35 % in aqueous 0.2 % formic acid for 45 min followed by an increase to 99 % acetonitrile for 10 min. The Orbitrap analyser was operated in full MS mode at a resolution (full width at half maximum—FWHM) of 60,000 at m/z 400 followed by eight data-dependent MS/MS scans. The first four MS/MS scans were performed as CID scans on the strongest four precursors for peptide identification followed by four high-energy collision-induced dissociation (HCD) scans of the same four precursors for reporter ion quantification. CID fragment ions were detected in the LTQ ion trap and HCD fragment ions in the Orbitrap analyser at a resolution FWHM of 15,000 at m/z 400. Dynamic exclusion was enabled, allowing two repeated fragmentation analyses on the same precursor during a 3-min time window.

Data analysis

We aimed to identify associations between protein species abundance and egg quality and therefore used regression analysis in order to evaluate how well protein abundance predicts egg quality. Statistical analyses were conducted using SPSS (version 19, IBM®) and R software (R development Core Team 2012). Raw spectra were processed through the Proteome Discoverer software (Thermo Scientific) and then submitted to the Mascot (Matrix Science LTD, UK) and SEQUEST (Eng et al. 1994) search engines for protein identification. Spectra were researched against the Teleostei subset of the NCBI nonredundant amino acid sequence database and its decoy version with reversed sequence entries to estimate the false discovery rate (FDR). The Mascot and SEQUEST search settings were identical and allowed for full tryptic peptides only with two missed cleavage sites and the following variable modifications: iTRAQ labels (N terminus, K, Y), oxidation (M), methylthio (C) and deamidation (N, Q). The score threshold for identification was adjusted to achieve an estimated FDR of <1 %. Only proteins that were identified by two or more peptide hits were accepted. The identified proteins were assigned to one of nine groups according to their function/pathway using the free software KEGG (http://www.genome.jp/kegg/pathway.html). Relative reporter ion intensities from the iTRAQ analysis were standardized against the sample labelled as '113' (*G1.50*; Table 1). Thus, intensity values were expressed as a ratio. A mixed-effects model analysis implemented in the R package 'lme4' (Bates et al. 2011) was used to test for relationships between the relative abundances of the proteins and egg batch quality within each of the nine functional groups of proteins. More specifically, for each protein group, we ran separate random slope mixed models with the abundance as the response (dependent) variable, the egg quality as a fixed effect, the protein identity and the sample identity as random intercepts and the egg quality as a random slope for each protein. A simple linear regression analysis was then performed for each protein separately to identify trends in relative abundance of a given protein in relation to egg quality (with the abundance as the response and the egg quality as a predictor). Performing multiple tests can lead to the increased probability of reporting false-positive results (i.e. significant associations that appear by chance), and in such situation, significance levels need to be adjusted accordingly to the number of tests performed (Manly et al. 2004). A false discovery rate (FDR) correction was, therefore, applied to the p values from the mixed-model analysis and to the p values from the simple

regression analysis, using the R package 'fdrtool' (Strimmer 2008) in order to control the number of false discoveries in multiple tests.

Results

Mascot/SEQUEST matched the suite of sequenced oligopeptides to 121 proteins with unique NCBI accession ID numbers; of these, 27 were vitellogenins (Vtgs) and 23 were 'unknown' or 'unnamed'. These 23 proteins were identified manually by submitting their unique peptide sequences to NCBI and selecting the first protein hit associated with a teleost. All of the 98 identified proteins were assigned to one of nine groups (amino acid metabolism antioxidants/housekeeping, ATP metabolism, glycolysis, heat-shock proteins, Krebs cycle, other metabolism-related proteins, purine metabolism and pentose phosphate pathway) based on their function and/or relevance to a biochemical pathway. A tenth group included proteins that did not share functionality or pathway (Table 2).

The mixed-effects model fitted for each protein group showed a statistically significant relationship between decreasing relative protein abundance and decreasing egg quality for the heat-shock protein group ($p = 0.046$). No statistically significant differences were identified for any of the other groups (all $p > 0.05$). When FDR correction was applied to the mixed-model results, all p values became statistically nonsignificant.

The simple regression analysis performed for each of the 121 proteins separately revealed seven proteins that exhibited statistically significant associations (at $p < 0.05$) between egg quality and relative protein abundance (Fig. 2; for results from the other proteins, see Table 2). All of these proteins (vitellogenin, heat-shock protein-70, 78 kDa glucose-regulated protein, elongation factor-2, GTP-binding nuclear protein Ran, iduronate 2-sulfatase and 6-phosphogluconate dehydrogenase) demonstrated decreased relative abundance with decreasing egg quality. The strongest trend (slope estimate -0.239 ± 0.086 SE, $R^2_{\text{adj}} = 0.492$, Fig. 2) was observed for the enzyme iduronate 2-sulfatase (Ids) for which the abundance ratios ranged between 1.48 and 0.66 from the highest (*G1.50, G1.70, G1.77*) to lowest quality egg batches (*X2.93*, *X3.31*, *X3.50*), respectively (Table 2). When FDR correction was applied to the regression results, the trends observed in the seven proteins listed above were no longer statistically significant; thus, using FDR, the likelihood of obtaining individual tests with the observed probability values when conducting 121 such tests was $p \geq 0.59$.

On subjecting proteins and their abundances to cluster analysis, several of the heat-shock proteins grouped together. However, the heat map did not clearly associate these proteins, or any other cluster, with egg quality (Fig. 3).

Discussion

The present study investigated the proteome of early-stage fertilized eggs of hapuku and the relationship between egg quality and the relative abundance of identified proteins. The functions and/or biochemical pathways corresponding to the proteins identified in this study depict common cellular processes that were reported in earlier studies that investigated protein contents in fish eggs and embryos (e.g. Papakostas et al. 2010; Castets et al. 2012). Unlike the above-mentioned studies, however, which used gel-based methods, the present study utilized the gel-free iTRAQ method. The use of iTRAQ allowed for simultaneous analysis of eight different samples, one of which was used as the reference sample.

We found that the relative abundance of heat-shock proteins (as a group) decreased with decreasing egg quality. Heat-shock proteins (HSP) in general, and Hsp70 specifically, are highly expressed, not only in response to stress, but also during normal development of fish, for example, in zebrafish (reviewed in Basu et al. 2002). Heat-shock proteins were also abundant in normal oocytes of gilthead seabream (*Sparus aurata*, Ziv et al. 2008). Castets et al. (2012) found several heat-shock proteins (including Hsp70) to have reduced relative abundances in poor quality eggs, whereas Chapman et al. (2014) identified *hsp90* transcripts among a subset of 233 mRNA species whose abundance had predictive value for egg quality in striped bass. A comparable general trend was found in the present study, in which the heat-shock protein group, split between two clusters in the heat map, exhibited lower relative abundance in lower quality eggs. Furthermore, two of the seven proteins that exhibited a trend in relative abundance associated with egg quality belong to the HSP70 family [Hsp70 and

Table 2 Proteins identified from hapuku (*P. oxygeneios*) 8-cell-stage embryo homogenates using proteomic iTRAQ analysis, presented by functional group

Gi: protein ID	Protein description	Group	Slope	SE	*t* value	*p* value
124020700	Betaine-homocysteine *S*-methyltransferase [*Paralichthys olivaceus*]	AAM	0.037	0.036	1.044	0.337
213512192	Protein arginine *N*-methyltransferase 1 [*Salmo salar*]	AAM	−0.007	0.045	−0.191	0.855
303225727	Betaine-homocysteine methyltransferase [*Plecoglossus altivelis*]	AAM	−0.070	0.054	−1.352	0.225
348545056	PREDICTED: asparagine synthetase [glutamine-hydrolysing]-like [*Oreochromis niloticus*]	AAM	−0.045	0.039	−1.121	0.305
323650178	Pterin-4-alpha-carbinolamine dehydratase [*Perca flavescens*]	AAM	−0.004	0.070	−0.049	0.962
47208664 (410932927)	PREDICTED: asparagine-tRNA ligase, cytoplasmic-like [*Takifugu rubripes*]	AAM	−0.024	0.100	−0.248	0.812
47218064 (348519833)	PREDICTED: D-3-phosphoglycerate dehydrogenase-like [*Oreochromis niloticus*]	AAM	0.071	0.220	0.312	0.765
154183805	Hibadhb [*Haplochromis burtoni*]	AH	−0.001	0.050	−0.028	0.979
190410767	Rho-class glutathione-*S*-transferase [*Paralichthys olivaceus*]	AH	0.035	0.131	0.269	0.797
198285617	Vertebrate peroxiredoxin 3 [*Salmo salar*]	AH	0.057	0.114	0.496	0.637
209733330	Peroxiredoxin-4 [*Salmo salar*]	AH	0.015	0.062	0.251	0.810
220679252	Glutathione-*S*-transferase M [*Danio rerio*]	AH	−0.025	0.012	−2.048	0.087
224613274	Disulphide-isomerase precursor [*Salmo salar*]	AH	−0.013	0.120	−0.110	0.916
225706354	Lactoylglutathione lyase [*Osmerus mordax*]	AH	−0.024	0.092	−0.264	0.800
229366432	Peroxiredoxin-1 [*Anoplopoma fimbria*]	AH	0.027	0.031	0.852	0.427
315143064	Copper zinc superoxide dismutase [*Lates calcarifer*]	AH	−0.101	0.054	−1.883	0.109
333827867	Mitochondrial peroxiredoxin 3 [*Sparus aurata*]	AH	0.041	0.111	0.369	0.725
333827869	Mitochondrial peroxiredoxin 5 [*Sparus aurata*]	AH	0.003	0.048	0.071	0.946
348500116	PREDICTED: lactoylglutathione lyase-like [*Oreochromis niloticus*]	AH	−0.002	0.053	−0.030	0.977
57908848	Cu/Zn superoxide dismutase [*Trematomus bernacchii*]	AH	−0.090	0.055	−1.628	0.155
86370988	Ubiquitin C [*Ictalurus punctatus*]	AH	−0.032	0.045	−0.699	0.511
47224642 (157152707)	Mn superoxide dismutase [*Takifugu obscurus*]	AH	0.041	0.062	0.664	0.532
194500331	Nucleoside diphosphate kinase [*Sparus aurata*]	ATP	0.014	0.031	0.457	0.664
30142116	Nucleoside diphosphate kinase [*Oreochromis mossambicus*]	ATP	0.030	0.021	1.414	0.207
122890758 (432961110)	PREDICTED: V-type proton ATPase catalytic subunit A-like [*Oryzias latipes*]	ATP	−0.059	0.036	−1.620	0.156
198285477	ATP synthase H+ transporting mitochondrial F1 complex beta [*Salmo salar*]	ATP	0.042	0.037	1.144	0.296
213512628	ATP synthase subunit alpha, mitochondrial [*Salmo salar*]	ATP	0.046	0.030	1.565	0.169
27545193	Brain creatine kinase b [*Danio rerio*]	ATP	0.009	0.042	0.216	0.836
317768001	Creatine kinase b-type [*Ictalurus punctatus*]	ATP	−0.001	0.029	−0.050	0.962
348520642	PREDICTED: creatine kinase, testis isozyme-like [*Oreochromis niloticus*]	ATP	−0.005	0.043	−0.103	0.921
119943230	Glyceraldehyde 3-phosphate dehydrogenase [*Misgurnus anguillicaudatus*]	GLY	0.085	0.068	1.248	0.259
224587654	Pyruvate kinase muscle isozyme [*Salmo salar*]	GLY	−0.020	0.038	−0.527	0.617
229366696	Alpha-enolase [*Anoplopoma fimbria*]	GLY	0.098	0.043	2.304	0.061

Table 2 continued

Gi: protein ID	Protein description	Group	Slope	SE	*t* value	*p* value
291195949	Aldolase A [*Thunnus albacares*]	GLY	0.035	0.067	0.526	0.618
317119973	Glyceraldehyde-3-phosphate dehydrogenase [*Coregonus clupeaformis*]	GLY	0.154	0.111	1.395	0.212
348514662	PREDICTED: alpha-enolase-like isoform 2 [*Oreochromis niloticus*]	GLY	0.093	0.051	1.846	0.115
348526250	PREDICTED: triosephosphate isomerase B [*Oreochromis niloticus*]	GLY	0.008	0.065	0.119	0.910
74096037	Pyruvate kinase [*Takifugu rubripes*]	GLY	0.003	0.044	−0.066	0.950
110226518	Heat-shock cognate 71 [*Paralichthys olivaceus*]	HSP	−0.065	0.051	−1.285	0.246
157679184	**HSP70 protein [*Poecilia reticulata*]**	**HSP**	−0.078	0.021	−3.632	**0.011**
209153200	60 kDa heat-shock protein, mitochondrial precursor [*Salmo salar*]	HSP	−0.071	0.048	−1.482	0.189
348544003	PREDICTED: stress-70 protein, mitochondrial-like [*Oreochromis niloticus*]	HSP	0.035	0.040	0.860	0.423
3513540	Heat-shock protein 70 [*Paralichthys olivaceus*]	HSP	−0.074	0.050	−1.481	0.189
348515387	**PREDICTED: 78 kDa glucose-regulated protein-like [*Oreochromis niloticus*]**	**HSP**	−0.072	0.029	−2.474	**0.048**
47225582 (303225725)	Heat-shock protein 70 kDa [*Plecoglossus altivelis*]	HSP	−0.054	0.047	−1.146	0.296
14583133	Malate dehydrogenase [*Sphyraena idiastes*]	KRC	−0.027	0.048	−0.565	0.592
225708458	Malate dehydrogenase, cytoplasmic [*Osmerus mordax*]	KRC	0.028	0.082	0.336	0.749
335060453	Citrate synthase [*Lates calcarifer*]	KRC	0.081	0.056	1.447	0.198
348544879	PREDICTED: malate dehydrogenase, cytoplasmic-like, partial [*Oreochromis niloticus*]	KRC	0.136	0.108	1.260	0.255
224587009	**Elongation factor 2 [*Salmo salar*]**	**NG**	−0.190	0.048	−3.927	**0.008**
229367044	**GTP-binding nuclear protein Ran [*Anoplopoma fimbria*]**	**NG**	−0.099	0.033	−2.973	**0.025**
239596162	Elongation factor 1-alpha [*Danio rerio*]	NG	−0.239	0.272	−0.881	0.412
317419952	Glycyl-tRNA synthetase [*Dicentrarchus labrax*]	NG	−0.111	0.081	−1.362	0.222
102579822	Glucose-regulated protein 75 [*Dentex dentex*]	NG	−0.029	0.059	−0.499	0.635
1351868	RecName: full = actin, alpha skeletal muscle; AltName: full = alpha-actin-1; flags: precursor	NG	0.053	0.055	0.960	0.374
148806551	Beta actin [*Liza aurata*]	NG	0.057	0.055	1.025	0.345
213512082	Syncoilin, intermediate filament protein [*Salmo salar*]	NG	−0.080	0.069	−1.161	0.290
225706824	Epididymal secretory protein E1 precursor [*Osmerus mordax*]	NG	−0.080	0.078	−1.015	0.349
229366360	Cofilin-2 [*Anoplopoma fimbria*]	NG	−0.045	0.042	−1.071	0.326
308321274	Peptidyl-prolyl *cis–trans* isomerase [*Ictalurus furcatus*]	NG	0.089	0.036	2.457	0.049
315570434	Complement component c3 [*Epinephelus coioides*]	NG	−0.125	0.092	−1.353	0.225
317418812	Melanocyte proliferating gene 1 [*Dicentrarchus labrax*]	NG	−0.165	0.120	−1.379	0.217
348510759	PREDICTED: glycogen phosphorylase, liver form-like isoform 3 [*Oreochromis niloticus*]	NG	−0.068	0.075	−0.916	0.395
348528665	**PREDICTED: iduronate 2-sulfatase-like [*Oreochromis niloticus*]**	**NG**	−0.239	0.086	−2.787	**0.032**
348532526	PREDICTED: 14-3-3 protein epsilon-like [*Oreochromis niloticus*]	NG	0.003	0.093	0.031	0.977
47196915 (410927798)	PREDICTED: histone H4-like [*Takifugu rubripes*]	NG	−0.051	0.086	−0.593	0.575

Table 2 continued

Gi: protein ID	Protein description	Group	Slope	SE	*t* value	*p* value
47210712 (432943845)	PREDICTED: T-complex protein 1 subunit beta-like [*Oryzias latipes*]	NG	0.011	0.148	0.074	0.943
47220173 (348528308)	PREDICTED: ES1 protein homolog, mitochondrial-like [*Oreochromis niloticus*]	NG	−0.024	0.092	−0.256	0.807
47225287 (229366360)	Cofilin-2 [*Anoplopoma fimbria*]	NG	−0.037	0.040	−0.929	0.389
47228615 (410899695)	PREDICTED: ubiquitin-like modifier-activating enzyme 1-like [*Takifugu rubripes*]	NG	−0.016	0.063	−0.261	0.803
48526097 (46811257)	Alpha-1-antitrypsin [*Pseudopleuronectes americanus*]	NG	−0.49	0.104	−0.466	0.658
348521480	PREDICTED: ubiquitin-like modifier-activating enzyme 1-like [*Oreochromis niloticus*]	NG	0.034	0.030	1.140	0.298
225794853	Acetyl-Coenzyme A acyltransferase [*Perca flavescens*]	OTM	0.052	0.127	0.413	0.694
334362378	Acetyl-Coenzyme A acyltransferase [*Epinephelus coioides*]	OTM	0.053	0.148	0.360	0.731
348527276	PREDICTED: putative aminopeptidase W07G4.4-like [*Oreochromis niloticus*]	OTM	−0.069	0.039	−1.771	0.127
348529566	PREDICTED: *S*-methyl-5′-thioadenosine phosphorylase-like [*Oreochromis niloticus*]	OTM	0.004	0.084	0.048	0.963
47209002 (410924035)	PREDICTED: medium-chain specific acyl-CoA dehydrogenase, mitochondrial-like [*Takifugu rubripes*]	OTM	0.158	0.216	0.734	0.491
47214599 (348511918)	PREDICTED: alcohol dehydrogenase class-3-like [*Oreochromis niloticus*]	OTM	−0.083	0.083	−0.992	0.359
47221528 (432960854)	PREDICTED: fructose-1,6-bisphosphatase 1-like isoform 2 [*Oryzias latipes*]	OTM	−0.099	0.139	−0.712	0.503
47222337 (21715873)	glutamate dehydrogenase [*Oncorhynchus mykiss*]	OTM	−0.156	0.185	−0.845	0.431
47225904 (432860335)	PREDICTED: peptidyl-prolyl *cis–trans* isomerase B-like [*Oryzias latipes*]	OTM	0.170	0.133	1.283	0.247
229365718	Inorganic pyrophosphatase [*Anoplopoma fimbria*]	OTM	0.079	0.082	0.965	0.372
215272373	Adenosylhomocysteinase [*Salmo salar*]	PM	0.022	0.024	0.914	0.396
223647846	Bifunctional purine biosynthesis protein PURH [*Salmo salar*]	PM	0.041	0.065	0.624	0.556
348521812	PREDICTED: adenosylhomocysteinase B-like [*Oreochromis niloticus*]	PM	0.022	0.032	0.676	0.524
47214847 (40363541)	Adenosylhomocysteinase [*Danio rerio*]	PM	0.053	0.046	1.149	0.294
47219328 (410920856)	PREDICTED: multifunctional protein ADE2-like [*Takifugu rubripes*]	PM	−0.067	0.096	−0.695	0.513
47230601 (410912454)	PREDICTED: bifunctional purine biosynthesis protein PURH-like [*Takifugu rubripes*]	PM	0.017	0.047	0.361	0.730
328677195	PREDICTED: adenylate kinase 2, mitochondrial-like [*Oreochromis niloticus*]	PM	0.013	0.059	0.228	0.828
223648108	6-Phosphogluconate dehydrogenase, decarboxylating [*Salmo salar*]	PPP	−0.129	0.099	−1.313	0.237
225706208	Transaldolase [*Osmerus mordax*]	PPP	0.000	0.045	0.007	0.995
348514848	**PREDICTED: 6-phosphogluconate dehydrogenase, decarboxylating [*Oreochromis niloticus*]**	**PPP**	−0.121	0.040	−3.029	**0.023**
348521484	PREDICTED: transketolase [*Oreochromis niloticus*]	PPP	0.004	0.051	0.075	0.943
348526660	PREDICTED: transaldolase-like [*Oreochromis niloticus*]	PPP	−0.027	0.052	−0.528	0.616
148299212	Vitellogenin [*Hippoglossus hippoglossus*]	VIT	0.023	0.010	2.364	0.056

Table 2 continued

Gi: protein ID	Protein description	Group	Slope	SE	*t* value	*p* value
148299214	Vitellogenin [*Hippoglossus hippoglossus*]	VIT	0.015	0.014	1.046	0.336
149850246	Vitellogenin A [*Mugil cephalus*]	VIT	0.003	0.012	0.229	0.826
149850248	Vitellogenin B [*Mugil cephalus*]	VIT	−0.015	0.024	−0.623	0.556
149850252	Phosvitinless vitellogenin [*Mugil cephalus*]	VIT	−0.018	0.042	−0.420	0.689
157278415	Vitellogenin II precursor [*Oryzias latipes*]	VIT	−0.015	0.022	−0.696	0.512
215397755	Major vitellogenin isoform 1 [*Clupea harengus*]	VIT	0.006	0.012	0.470	0.655
217314658	Vitellogenin Ab1 [*Centrolabrus exoletus*]	VIT	0.015	0.006	2.426	0.051
217314666	Vitellogenin Ab2 [*Labrus mixtus*]	VIT	0.007	0.013	0.553	0.600
225794846	Vitellogenin C [*Perca flavescens*]	VIT	−0.071	0.065	−1.088	0.318
260159577	**Vitellogenin [*Thunnus thynnus*]**	**VIT**	−0.103	0.041	−2.498	**0.047**
290793106	Vitellogenin B [*Thunnus thynnus*]	VIT	0.014	0.014	1.016	0.349
290793108	Vitellogenin C [*Thunnus thynnus*]	VIT	−0.015	0.045	−0.340	0.745
309242177	Vitellogenin [*Trematomus bernacchii*]	VIT	0.015	0.015	1.028	0.344
3123011	Vitellogenin precursor	VIT	−0.010	0.015	−0.668	0.529
326375569	Vitellogenin Ab [*Morone saxatilis*]	VIT	−0.009	0.020	−0.455	0.665
326375571	Vitellogenin C [*Morone saxatilis*]	VIT	−0.022	0.044	−0.516	0.624
346990909	Vitellogenin [*Gobiomorphus cotidianus*]	VIT	−0.341	0.216	−1.578	0.166
56785787	Vitellogenin [*Epinephelus coioides*]	VIT	0.011	0.017	0.666	0.530
62241078	Vitellogenin [*Verasper moseri*]	VIT	0.016	0.016	0.988	0.361
62241080	Vitellogenin [*Verasper moseri*]	VIT	0.020	0.012	1.627	0.155
62241084	Vitellogenin [Gambusia affinis]	VIT	0.014	0.016	0.917	0.395
71011626	Vitellogenin B [*Morone americana*]	VIT	−0.007	0.017	−0.399	0.703
71011912	vitellogenin C [Morone americana]	VIT	−0.015	0.044	−0.347	0.740
74149754	Vitellogenin [*Pagrus major*]	VIT	−0.017	0.013	−1.310	0.238
74149756	Vitellogenin [*Pagrus major*]	VIT	−0.010	0.018	−0.587	0.579
74149758	Phosvitinless vitellogenin [*Pagrus major*]	VIT	−0.025	0.044	−0.575	0.586

The relationship between the abundance of each protein (relative signal intensity) in the different samples ($n = 8$) and their embryo quality scores (based on blastomere symmetry) were determined using linear regression

Proteins are divided into major groups according to their function/biochemical pathway. 'protein ID' refers to the protein identification number (GI number) on NCBI. Numbers in brackets indicate the first manually identified protein hit that corresponded to a teleost (see text for details). The regression model was fitted separately for each protein, and *p* values shown are not adjusted for the false discovery rate. *SE* standard error. Protein functional groups are coded as: *AAM* amino acid metabolism, *AH* antioxidants/housekeeping, *ATP* ATP metabolism, *GLY* glycolysis, *HSP* heat-shock proteins, *KRC* Krebs cycle, *NG* nongrouped proteins, *OTM* other metabolism-related proteins, *PM* purine metabolism, *PPP* pentose phosphate pathway, *VIT* vitellogenins. The seven proteins that exhibited a statistically significant trend (prior to FDR) are highlighted in bold font

78 kDa glucose-regulated protein (Grp78)]. Grp78 functions as a chaperone involved in translocation and protein folding (Haas 1994) and has been previously identified in fish liver proteome (Malécot et al. 2009). Taken together, these results reinforce the existence of a link between the abundance of heat-shock proteins and egg quality in hapuku and, quite possibly, teleost fish in general.

Significantly, a negative trend (decrease in relative abundance with decreasing quality) was found in the present study for the enzyme iduronate 2-sulfatase (Ids). Activity of this enzyme during the early stages (64-cell stage) of zebrafish development was previously detected and peaked at the gastrula stage (Moro et al. 2010), indicating that Ids may be especially important during the early stages of embryo development; indeed, knock-down of the gene encoding this enzyme was strongly associated with severe head and trunk deformities in developing zebrafish larvae (Moro et al. 2010). These findings are of particular interest in the light of the results from the present study, since poor quality hapuku eggs were

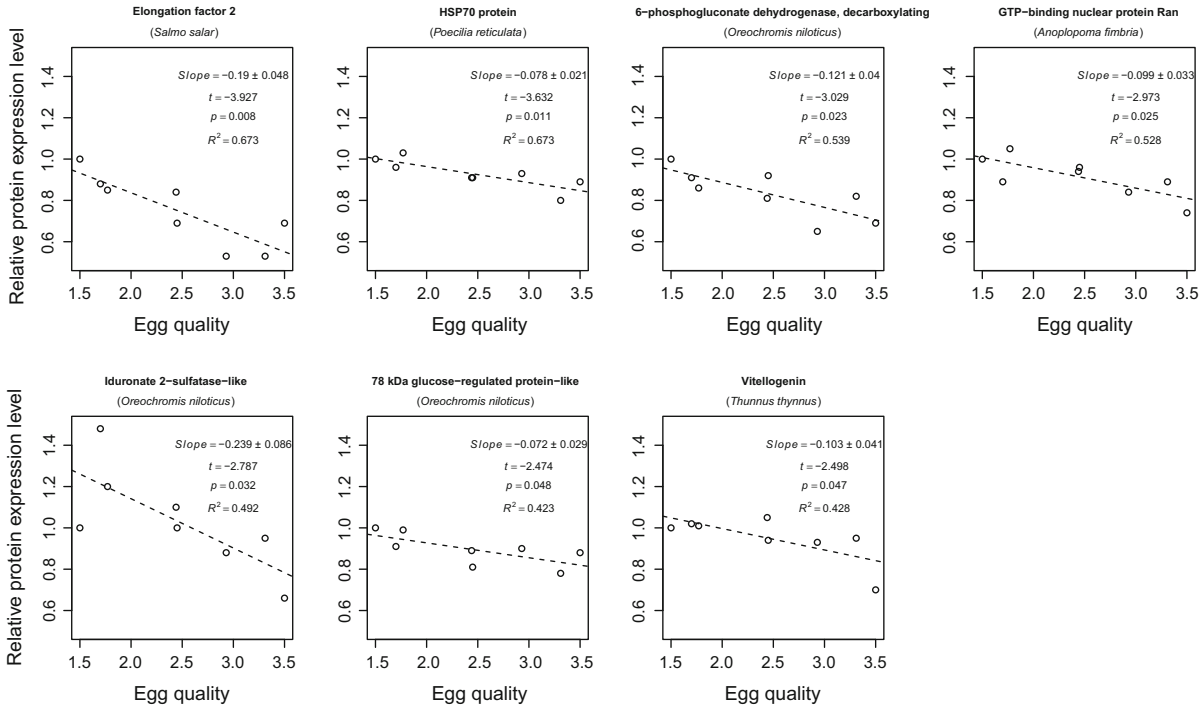


Fig. 2 Relationships between blastomere symmetry score and relative abundances of a set of proteins from hapuku, *P. oxygeneios*. Relative abundances of proteins in eight batches of eggs were determined using iTRAQ, and abundances were correlated with the egg blastomere symmetry scores ('egg quality'); only significant relationships (prior to FDR correction) are shown

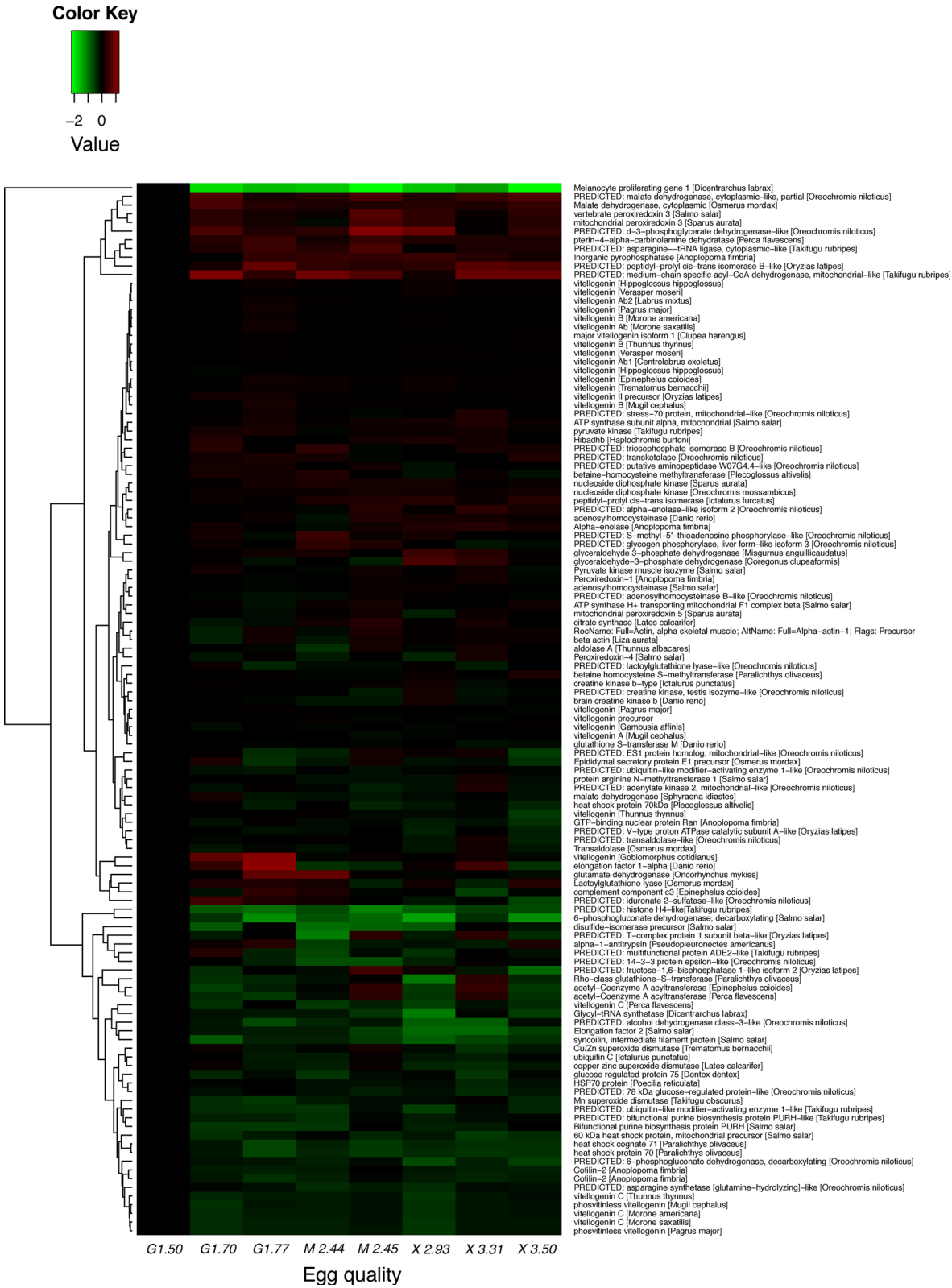
characterized by severe blastomere asymmetry (Kohn and Symonds 2012), which correlated with low abundances of Ids. As *ids* transcripts have been found in unfertilized zebrafish eggs (Moro et al. 2010), this mRNA species is likely to be of maternal origin and, as such, contributes to the 'legacy' of a female to its offspring (Gandolfi and Gandolfi 2001); accordingly, the production of eggs low in *ids* transcript or Ids protein abundance may well reflect a 'poor legacy' by a female, and early identification and purging of such females from the broodstock tank could improve hatchery production efficiency.

Notwithstanding the identification of a single Vtg peptide that showed a significant positive correlation between quality and relative Vtg abundance, the overall relative abundance of identified Vtgs showed no significant relationship with egg quality.

The yolk proteins that are derived from the cleavage of Vtgs were shown to comprise more than 90 % of total egg protein in zebrafish (*D. rerio*) embryos (Link et al. 2006). Because of their abundance, yolk proteins can mask the presence of less abundant proteins that may be important to embryo development (Link et al. 2006). Attempts to reduce the amount of Vtg in the samples, using ammonium sulphate precipitation and immunoprecipitation in the present study (results not shown), did not yield the desired results. Thus, similar to previous studies which investigated egg quality in fish using proteomics (Yoshikuni et al. 2003; Rime et al. 2004, Crespel et al. 2008, Castets et al. 2012), the analysis in the present study was carried out with yolk proteins included in the tested samples.

Ziv et al. (2008) found several proteins related to translation control and protein synthesis in maturing oocytes of zebrafish and gilthead seabream. In the present study, the relative abundance of elongation factor-2, a ubiquitous protein that is critical for protein synthesis (Robinson et al. 1974), was lower with decreasing egg quality. A similar trend was seen for the GTP-binding nuclear protein Ran, which is involved in protein import into the nucleus and is known to affect DNA replication, cell cycle and preservation of nuclear structure (Moore and Blobel 1993).

The citric acid cycle, glycolysis and pentose phosphate pathways are fundamental to biological


Color Key
−2 0
Value
Melanocyte proliferating gene 1 [Dicentrarchus labrax]
PREDICTED: malate dehydrogenase, cytoplasmic–like, partial [Oreochromis niloticus]
Malate dehydrogenase, cytoplasmic [Osmerus mordax]
vertebrate peroxiredoxin 3 [Salmo salar]
mitochondrial peroxiredoxin 3 [Sparus aurata]
PREDICTED: d–3–phosphoglycerate dehydrogenase–like [Oreochromis niloticus]
pterin–4–alpha–carbinolamine dehydratase [Perca flavescens]
PREDICTED: asparagine–tRNA ligase, cytoplasmic–like [Takifugu rubripes]
Inorganic pyrophosphatase [Anoplopoma fimbria]
PREDICTED: peptidyl–prolyl cis–trans isomerase B–like [Oryzias latipes]
PREDICTED: medium–chain specific acyl–CoA dehydrogenase, mitochondrial–like [Takifugu rubripes]
vitellogenin [Hippoglossus hippoglossus]
vitellogenin [Verasper moseri]
vitellogenin Ab2 [Labrus mixtus]
vitellogenin [Pagrus major]
vitellogenin B [Morone americana]
vitellogenin Ab [Morone saxatilis]
major vitellogenin isoform 1 [Clupea harengus]
vitellogenin B [Thunnus thynnus]
vitellogenin [Verasper moseri]
vitellogenin Ab1 [Centrolabrus exoletus]
vitellogenin [Hippoglossus hippoglossus]
vitellogenin [Epinephelus coioides]
vitellogenin [Trematomus bernacchii]
vitellogenin II precursor [Oryzias latipes]
vitellogenin B [Mugil cephalus]
PREDICTED: stress–70 protein, mitochondrial–like [Oreochromis niloticus]
ATP synthase subunit alpha, mitochondrial [Salmo salar]
pyruvate kinase [Takifugu rubripes]
Hibadhb [Haplochromis burtoni]
PREDICTED: triosephosphate isomerase B [Oreochromis niloticus]
PREDICTED: transketolase [Oreochromis niloticus]
PREDICTED: putative aminopeptidase W07G4.4–like [Oreochromis niloticus]
betaine–homocysteine methyltransferase [Pleocoglossus altivelis]
nucleoside diphosphate kinase [Sparus aurata]
nucleoside diphosphate kinase [Oreochromis mossambicus]
peptidyl–prolyl cis–trans isomerase [Ictalurus furcatus]
PREDICTED: alpha–enolase–like isoform 2 [Oreochromis niloticus]
adenosylhomocysteinase [Danio rerio]
Alpha–enolase [Anoplopoma fimbria]
PREDICTED: S–methyl–5'–thioadenosine phosphorylase–like [Oreochromis niloticus]
PREDICTED: glycogen phosphorylase, liver form–like isoform 3 [Oreochromis niloticus]
glyceraldehyde 3–phosphate dehydrogenase [Misgurnus anguillicaudatus]
glyceraldehyde–3–phosphate dehydrogenase [Coregonus clupeaformis]
Pyruvate kinase muscle isozyme [Salmo salar]
Peroxiredoxin–1 [Anoplopoma fimbria]
adenosylhomocysteinase [Salmo salar]
PREDICTED: adenosylhomocysteinase B–like [Oreochromis niloticus]
ATP synthase H+ transporting mitochondrial F1 complex beta [Salmo salar]
mitochondrial peroxiredoxin 5 [Sparus aurata]
citrate synthase [Lates calcarifer]
RecName: Full=Actin, alpha skeletal muscle; AltName: Full=Alpha–actin–1; Flags: Precursor
beta actin [Liza aurata]
aldolase A [Thunnus albacares]
Peroxiredoxin–4 [Salmo salar]
PREDICTED: lactoylglutathione lyase–like [Oreochromis niloticus]
betaine homocysteine S–methyltransferase [Paralichthys olivaceus]
creatine kinase b–type [Ictalurus punctatus]
PREDICTED: creatine kinase, testis isozyme–like [Oreochromis niloticus]
brain creatine kinase b [Danio rerio]
vitellogenin [Pagrus major]
vitellogenin precursor
vitellogenin [Gambusia affinis]
vitellogenin A [Mugil cephalus]
glutathione S–transferase M [Danio rerio]
PREDICTED: ES1 protein homolog, mitochondrial–like [Oreochromis niloticus]
Epididymal secretory protein E1 precursor [Osmerus mordax]
PREDICTED: ubiquitin–like modifier–activating enzyme 1–like [Oreochromis niloticus]
protein arginine N–methyltransferase 1 [Salmo salar]
PREDICTED: adenylate kinase 2, mitochondrial–like [Oreochromis niloticus]
malate dehydrogenase [Sphyraena idiastes]
heat shock protein 70kDa [Plecoglossus altivelis]
vitellogenin [Thunnus thynnus]
GTP–binding nuclear protein Ran [Anoplopoma fimbria]
PREDICTED: V–type proton ATPase catalytic subunit A–like [Oryzias latipes]
PREDICTED: transaldolase–like [Oreochromis niloticus]
Transaldolase [Osmerus mordax]
vitellogenin [Gobiomorphus cotidianus]
elongation factor 1–alpha [Danio rerio]
glutamate dehydrogenase [Oncorhynchus mykiss]
Lactoylglutathione lyase [Osmerus mordax]
complement component c3 [Epinephelus coioides]
PREDICTED: iduronate 2–sulfatase–like [Oreochromis niloticus]
PREDICTED: histone H4–like[Takifugu rubripes]
6–phosphogluconate dehydrogenase, decarboxylating [Salmo salar]
disulfide–isomerase precursor [Salmo salar]
PREDICTED: T–complex protein 1 subunit beta–like [Oryzias latipes]
alpha–1–antitrypsin [Pseudopleuronectes americanus]
PREDICTED: multifunctional protein ADE2–like [Takifugu rubripes]
PREDICTED: 14–3–3 protein epsilon–like [Oreochromis niloticus]
PREDICTED: fructose–1,6–bisphosphatase 1–like isoform 2 [Oryzias latipes]
Rho–class glutathione–S–transferase [Paralichthys olivaceus]
acetyl–Coenzyme A acyltransferase [Epinephelus coioides]
acetyl–Coenzyme A acyltransferase [Perca flavescens]
vitellogenin C [Perca flavescens]
Glycyl–tRNA synthetase [Dicentrarchus labrax]
PREDICTED: alcohol dehydrogenase class–3–like [Oreochromis niloticus]
Elongation factor 2 [Salmo salar]
syncollin, intermediate filament protein [Salmo salar]
Cu/Zn superoxide dismutase [Trematomus bernacchii]
ubiquitin C [Ictalurus punctatus]
copper zinc superoxide dismutase [Lates calcarifer]
glucose regulated protein 75 [Dentex dentex]
HSP70 protein [Poecilia reticulata]
PREDICTED: 78 kDa glucose–regulated protein–like [Oreochromis niloticus]
Mn superoxide dismutase [Takifugu obscurus]
PREDICTED: ubiquitin–like modifier–activating enzyme 1–like [Takifugu rubripes]
PREDICTED: bifunctional purine biosynthesis protein PURH–like [Takifugu rubripes]
Bifunctional purine biosynthesis protein PURH [Salmo salar]
60 kDa heat shock protein, mitochondrial precursor [Salmo salar]
heat shock cognate 71 [Paralichthys olivaceus]
heat shock protein 70 [Paralichthys olivaceus]
PREDICTED: 6–phosphogluconate dehydrogenase, decarboxylating [Oreochromis niloticus]
Cofilin–2 [Anoplopoma fimbria]
Cofilin–2 [Anoplopoma fimbria]
PREDICTED: asparagine synthetase [glutamine–hydrolyzing]–like [Oreochromis niloticus]
vitellogenin C [Thunnus thynnus]
phosvitinless vitellogenin [Mugil cephalus]
vitellogenin C [Morone americana]
vitellogenin C [Morone saxatilis]
phosvitinless vitellogenin [Pagrus major]
G1.50 G1.70 G1.77 M 2.44 M 2.45 X 2.93 X 3.31 X 3.50
Egg quality

◀**Fig. 3** Heat map, representing the hierarchical clustering of peptide abundance data ($n = 121$) in early embryos (8-cell stage) of different quality ('0' is best, '4' is worst) in hapuku (*P. oxygeneios*). The abundance of each peptide within a sample was normalized over that of the sample isobarically labelled with a tag of 113 Da (*G1.50*; peptides in this sample all received a relative abundance score of $^{2}\log(1) = 0$, represented by the *black colour)*. Relative peptide abundances for the remaining 7 samples are indicated on the colour scale above the heat map, such that a score of −1 (*green*) or +1 (*red*) represents a half-fold or twofold increase in relative peptide abundance, respectively. For details on egg quality scores, see text. (Color figure online)

systems and function in essential carbon metabolic processes and metabolism. Aconitase hydratase, an enzyme involved in the citric acid cycle, was found to be less abundant in lower quality eggs in Eurasian perch (Castets et al. 2012). In the present study, the relative abundance of 6-phosphogluconate dehydrogenase, which is an enzyme in the pentose phosphate pathway, also exhibited a negative trend, decreasing in relative abundance from the high- to low-quality groups. Similarly, transcript abundances of a suite of genes associated with metabolism-related pathways in eggs from wild Atlantic cod brood stock ('high quality') differed from those in farmed brood stock (Lanes et al. 2013), reinforcing the perhaps obvious notion that metabolic 'state' of the egg is likely to be important for its quality.

The likelihood of false positives (type I error) when examining a large data set is high, particularly when the differences between samples are small or associations are weak. Indeed, when FDR correction was applied to our results, these uncovered associations were no longer statistically significant. This may have been influenced by two factors. Firstly, the differences between the egg quality groups may not have been extreme enough for this type of analysis. We note that the highest quality batch score in the present study corresponded to a 'moderate' quality batch as described in Kohn and Symonds (2012). Therefore, the lack of strictly 'good' quality batches has made it more difficult to detect differences between batches with differing quality scores, which is also evident from the clustering analysis shown in the heat map. Secondly, recent reviews that have critically examined the efficiency of the iTRAQ method in quantifying proteins (Christoforou and Lilley 2012; Evans et al. 2012) have argued that iTRAQ analysis may underestimate the fold change in protein abundance. It has been shown that ratios of reporter ions, especially of low-abundance peptides, are biased towards '1' based on the co-isolation and co-fragmentation of peptides from unregulated high-abundance proteins which have very similar *m*/*z* values to the targeted low-abundance peptides (Ow et al. 2009, Karp et al. 2010, and Mahoney et al. 2011). This results in mixed reporter ion ratios within the fragment ion spectrum. Thus, iTRAQ might have a limited sensitivity for the quantification of weak fold changes in low-abundance proteins. Furthermore, the contamination of specific peptide fragment spectra depends on the sample complexity and the sample matrix with which the peptides co-elute. Consequently, not every peptide fragment ion spectrum of the same protein shows contaminated reporter ion ratios which results in high coefficients of variation in low-abundance proteins. The aim of the here-presented iTRAQ study was to perform a global screen for potential target proteins or pathways that are associated with egg quality. Such targets, however, require further confirmation using orthogonal and more accurate methods which in general have a limited throughput and are not suitable for global quantitative profiling. Regardless, proteins with similar functions to those identified here have been implicated in other studies (which, incidentally, have not implemented FDR correction) to be important for embryo development. Therefore, the trends observed here can help elucidate the importance of these proteins with regard to the normal development of early-stage embryos.

Whilst our analyses identified a number of candidate proteins that correlated with egg quality, it is important to note that individual candidates are likely to have limited predictive value of egg quality, as illustrated by Chapman et al. (2014). These authors concluded from their analysis of transcriptome data on striped bass (*Morone saxatilis*) that individual candidate transcripts may only explain ~1 % of the variation in egg quality. They further contended that egg quality clearly is too complex a trait to be linked to a single candidate, but that as (relatively) few as 250 genes could explain as much as 90 % of the variation in egg quality. Therefore, strong associations between egg quality and candidate gene expression (mRNA, protein) may well prove useful as low-cost and quick turnover biomarkers in future.

In conclusion, our study reported on the proteome of hapuku embryos at the early 8-cell stage of

development. Generically, heat-shock proteins decreased in abundance with egg quality. Regression analysis further identified seven proteins that exhibited a decreasing trend in relative abundance in relation to egg quality—of particular interest is the association between the decreasing relative abundance of Ids that was previously linked to impaired fish embryo development and decreasing egg quality. We contend that Ids and heat-shock proteins are obvious candidate biomarkers for egg quality in hapuku.

Acknowledgments We thank the technical staff at the Centre for Protein Research, University of Otago, for their assistance in performing the iTRAQ analysis. We also thank the technical staff at NIWA Aquaculture Park in Bream Bay, Ruakaka, for their assistance with sample collection and Mr Ken Miller, University of Otago, for his assistance with tidying up some of the figures. We thank Dr Alvin Setiawan for providing useful comments to the manuscript draft.

References

Anderson NL, Anderson NG (1998) Proteome and proteomics: new technologies, new concepts and new words. Electrophoresis 19:1853–1861

Basu N, Todgham AE, Ackerman PA, Bibeau MR, Nakano K, Schulte PM, Iwama GK (2002) Heat shock protein genes and their functional significance in fish. Gene 295:173–183

Bates D, Maechler M, Bolker B (2011). lme4: linear mixed-effects models R package, version 0.999375-42. http://CRAN.R-project.org/package=lme4. Accessed 2 Dec 2013

Castets M-D, Schaerlinger B, Silvestre F, Gardeur J-N, Dieu M, Corbier C, Kestemont P, Fontaine P (2012) Combined analysis of *Perca fluviatilis* reproductive performance and oocyte proteomic profile. Theriogenology 778:432–442

Cerda J, Bobe J, Babin PJ, Admon A, Lubzens E (2008) Functional genomics and proteomic approaches for the study of gamete formation and viability in farmed finfish. Rev Fish Sci 16:56–72

Chapman RW, Reading BJ, Sullivan CG (2014) Ovary transcriptome profiling via artificial intelligence reveals a transcriptomic fingerprint predicting egg quality in striped bass, *Morone saxatilis*. PLoS ONE 9:e96818

Chenau J, Michelland S, Sidibe J, Seve M (2008) Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling. Proteome Sci 6:9

Christoforou AL, Lilley KS (2012) Isobaric tagging approaches in quantitative proteomics: the ups and downs. Anal Bioanal Chem 404:1029–1037

Crespel A, Rime H, Fraboulet E, Bobe J, Fauvel C (2008) Egg quality in domesticated and wild seabass (*Dicentrarchus labrax*): a proteomic analysis. Cybium 32:205

Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5:967–989

Evans C, Noirel J, Ow SY, Salim M, Pereire-Medrano AG, Couto N, Pandhal J, Smith D, Pham TK, Karunakaran E, Zou X, Biggs CA, Wright PC (2012) An insight into iTRAQ: where do we stand now? Anal Bioanal Chem 404:1011–1027

Forné I, Abián J, Cerda J (2010) Fish proteome analysis: model organisms and non-sequenced species. Proteomics 10:858–872

Gandolfi TALB, Gandolfi F (2001) The maternal legacy to the embryo: cytoplasmic components and their effects on early development. Theriogenology 55:1255–1276

Haas IG (1994) BiP (GRP78), an essential hsp 70 resident protein in the endoplasmic reticulum. Experientia 50:1012–1020

Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS (2010) Addressing accuracy and precision issues in iTRAQ quantitation. Mol Cell Proteomics 9:1885–1897

Knoll-Gellida A, Andre M, Gattegno T, Forgue J, Admon A, Babin PJ (2006) Molecular phenotype of zebrafish ovarian follicle by serial analysis of gene expression and proteomic profiling, and comparison with the transcriptomes of other animals. BMC Genom 7:46

Kohn YY, Symonds JE (2012) Evaluation of egg quality parameters as predictors of hatching success and early larval survival in hapuku (*Polyprion oxygeneios*). Aquaculture 342–343:42–47

Lanes CFC, Bizuayehu TT, de Oliveira Fernandes JM, Kiron V, Babiak I (2013) Transcriptome of Atlantic cod (*Gadus morhua* L.) early embryos from farmed and wild broodstocks. Mar Biotechnol 15:677–694

Li L, Lu X, Dean J (2013) The maternal to zygotic transition in mammals. Mol Aspects Med 34:919–938

Link V, Shevchenko A, Heisenberg C-P (2006) Proteomics of early zebrafish embryos. BMC Dev Biol 6:1

Lokman PM, Symonds JE (2014) Molecular and biochemical tricks of the research trade: ~ omics approaches in finfish aquaculture. N Z J Mar Freshwater Res 48:492–505

Ma H, Hostuttler M, Hairong W, Rexroad CE III, Yao JB (2012) Characterization of the rainbow trout egg microRNA transcriptome. PLoS ONE 7:e39649

Mahoney DW, Therneau TM, Heppelmann CJ, Higgins L, Benson LM, Zenka RM, Jagtap P, Nelsestuen GL, Bergen HR, Oberg AL (2011) Relative quantification: characterization of bias, variability and fold changes in mass spectrometry data from iTRAQ-labeled peptides. J Proteome Res 10:4325–4333

Malécot M, Mezhound K, Marie A, Praseuth D, Puiseux-Dao S, Edery M (2009) Proteomic study of mictocystin-LR on organelle and membrane proteins in medaka fish liver. Aquat Toxicol 94:153–161

Manly KF, Nettleton D, Gene Hwang JT (2004) Genomics, prior probability, and statistical tests of multiple hypotheses. Genome Res 14:997–1001

Martyniuk CJ, Denslow ND (2009) Towards functional genomics in fish using quantitative proteomics. Gen Comp Endocrinol 16:135–141

Martyniuk CJ, Alvarez S, Denslow ND (2012) DIGE and iTRAQ as biomarker discovery tools in aquatic toxicology. Ecotoxicol Environ Saf 76:3–10

Moore MS, Blobel G (1993) The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. Nature 365:661–663

Moro E, Tomanin R, Frisco A, Modena N, Tiso N, Scarpa M, Argenton F (2010) A novel functional role of iduronate-2-sulfatase in zebrafish early development. Matrix Biol 29:43–50

Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC (2009) iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". J Proteome Res 8:5347–5355

Palace VP, Werner J (2006) Vitamins A and E in the maternal diet influence egg quality and early life stage development in fish: a review. Scientia Marina 70(Suppl 2):41–57

Papakostas S, Vøllestad LA, Primmer CR, Leder EH (2010) Proteomic profiling of early life stages of European grayling (*Thymallus thymallus*). J Proteome Res 9:4790–4800

R Development Core Team (2012) R foundation for statistical computing. R Development Core Team, Vienna

Rime H, Guitton N, Pineau C, Bonnet E, Bobe J, Jalabert B (2004) Post-ovulatory ageing and egg quality: a proteomic analysis of rainbow trout coelomic fluid. Reprod Biol Endocrinol 2:26–35

Robinson EA, Henriksen O, Maxwell ES (1974) Elongation factor 2, amino acid sequence at the site of adenosine diphosphate ribosylation. J Biol Chem 249:5088–5093

Rodrigues PM, Silva TS, Dias J, Jessen F (2012) Proteomics in aquaculture: applications and trends. J Proteomics 75:4325–4345

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169

Salze G, Tocher DR, Roy WJ, Robertson DR (2005) Egg quality determinants in cod (*Gadus morhua* L.): egg performance and lipids in eggs from farmed and wild broodstock. Aquac Res 36:1488–1499

Sanchez BC, Ralston-Hooper K, Sepúlveda MS (2011) Review of recent proteomic applications in aquatic toxicology. Environ Toxicol Chem 30:274–282

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473:337–342

Shields RJ, Brown NP, Bromage NR (1997) Blastomere morphology as a predictive measure of fish egg viability. Aquaculture 155:1–12

Strimmer K (2008) fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics 24:1461–1462

Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. Nat Methods 6:359–362

Yoshikuni M, Sagegami R, Nagahama Y (2003) Proteome analysis: a new approach to identify key proteins involved in acquisition of maturational competence and oocyte maturation of medaka oocytes. Fish Physiol Biochem 28:379–380

Zieske LR (2006) A perspective on the use of iTRAQ^{TM} reagent technology for protein complex and profiling studies. J Exp Bot 57:1501–1508

Ziv T, Gattegno T, Chapovetsky V, Wolf H, Barnea E, Lubzens E, Admon A (2008) Comparative proteomics of the developing fish (zebrafish and gilthead seabream) oocytes. Comp Biochem Physiol 3D:12–35