



**HAL**  
open science

## Neurodevelopment vs. the immune system: Complementary contributions of maternally-inherited gene transcripts and proteins to successful embryonic development in fish

Daniel Źarski, Aurelie Le Cam, Thomas Frohlich, Miwako Kösters, Christophe Klopp, Joanna Nynca, Slawomir Ciesielski, Beata Sarosiek, Katarzyna Dryl, Jerome Montfort, et al.

### ► To cite this version:

Daniel Źarski, Aurelie Le Cam, Thomas Frohlich, Miwako Kösters, Christophe Klopp, et al.. Neurodevelopment vs. the immune system: Complementary contributions of maternally-inherited gene transcripts and proteins to successful embryonic development in fish. *Genomics*, 2021, 113 (6), pp.3811-3826. 10.1016/j.ygeno.2021.09.003 . hal-03342588v2

**HAL Id: hal-03342588**

**<https://hal.inrae.fr/hal-03342588v2>**

Submitted on 5 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Neurodevelopment vs. the immune system: Complementary contributions of maternally-inherited gene transcripts and proteins to successful embryonic development in fish

Daniel Zarski <sup>a</sup>, Aurelie Le Cam <sup>b</sup>, Thomas Frohlich <sup>c</sup>, Miwako Kösters <sup>c</sup>, Christophe Klopp <sup>d</sup>, Joanna Nynca <sup>a</sup>, Sławomir Ciesielski <sup>e</sup>, Beata Sarosiek <sup>a</sup>, Katarzyna Dryl <sup>a</sup>, Jerome Montfort <sup>b</sup>, Jarosław Król <sup>f</sup>, Pascal Fontaine <sup>g</sup>, Andrzej Ciereszko <sup>a</sup>, Julien Bobe <sup>b</sup>.

**a** : Department of Gamete and Embryo Biology, Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Tuwima str. 10, 10-748 Olsztyn, Poland

**b** : INRAE, UR1037 Fish Physiology and Genomics, F-35000 Rennes, France

**c** : Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig Maximilian University of Munich, 81377 Munich, Germany

**d** : Sigene, UR875, INRAE, 31326 Castanet-Tolosan, France

**e** : Department of Environmental Biotechnology, University of Warmia and Mazury in Olsztyn, Poland

**f** : Department of Salmonid Research, The Stanislaw Sakowicz Inland Fisheries Institute, Olsztyn, Poland

**g** : University of Lorraine, INRAE, UR AFPA, F-54000 Nancy, France

*Published in Genomics. Available online: <https://doi.org/10.1016/j.ygeno.2021.09.003>*

## Keywords

Egg quality, pikeperch, reproduction, transcriptomics, proteomics, embryo lethality.

## Abstract

The aim of this study was to investigate the respective contribution of maternally-inherited mRNAs and proteins to egg molecular cargo and to its developmental competence in fish using pikeperch as a model. Our study provides novel insights into the understanding of type-specific roles of maternally-inherited molecules in fish. Here we show, for the first time, that transcripts and proteins have distinct, yet complementary, functions in the egg of teleost fish. Maternally-inherited mRNAs would shape embryo neurodevelopment, while maternally-inherited proteins would rather be responsible for protecting the embryo against pathogens. Additionally, we observed that processes directly preceding ovulation may considerably affect the reproductive success by modifying expression level of genes crucial for proper embryonic development, being novel fish egg quality markers (e.g., *smarca4* or *h3f3a*). These results are of major importance for understanding the influence of external factors on reproductive fitness in both captive and wild-type fish species.

## 1. Introduction

---

The developmental competence of metazoans ova relies on maternally derived mRNAs (messenger ribonucleic acids) and proteins that drive cellular division until zygotic genome activation (ZGA) [1–3]. Further developmental processes in fish, including cellular differentiation, are followed by reprogramming processes upon ZGA, during which maternally inherited mRNAs undergo degradation and are replaced by zygotic transcripts [4]. In this way, new proteins taking part in embryogenesis are also being produced. However, mRNA degradation upon ZGA is most likely a prolonged process during which selective degradation (targeted to specific mRNAs) occurs [5,6]. Consequently, it has been found that

many mRNAs of maternal origin participate in embryonic development well beyond ZGA, consequently representing a potential nongenetic inheritance pathway [7]; this also probably applies to maternally derived proteins. It is therefore clear that both sets of molecules – transcripts and proteins – of maternal origin are important factors in embryonic development. The lacking or very poor concordance between protein and transcript levels in mammalian oocytes during early development [8] and in *Xenopus laevis* eggs and early embryos [9] suggests that maternally derived proteins and mRNAs play different roles during embryogenesis. This hypothesis is additionally supported by a study on zebrafish

(*Danio rerio*), in which different functions of transcripts and proteins were suggested in the embryos at 24 h post fertilization [10]. However, this study was performed on embryos at a late stage of embryonic development (during advanced organogenesis, long after ZGA), representing molecular consequences stemming from maternal and paternal factors as well as their interactions. Therefore, there is still a lack of consideration of the specific roles of the two types of molecules of purely maternal origin in unfertilized ova of vertebrates, including fishes. This knowledge is of the highest importance for understanding and identifying the processes and mechanisms responsible for the reproductive capacity of teleost fishes.

In finfishes, oogenesis is the process of oocyte growth, during which proteins and lipids are deposited into the oocyte [11,12], followed by final oocyte maturation (FOM), during which reorganization of cellular structures occurs [13–15]. This latter process, considered translationally quiescent, is accompanied by transcription within the ovaries, and genomic maturation [12,16] leads to the deposition of maternally derived molecular cargo in the ovulated, fully mature egg. Upon ovulation, eggs remain transcriptionally and translationally silent or significantly repressed until ZGA [17,18]. Among the egg constituents, many research efforts have focused on fats and the major fraction of proteins (such as vitellogenins), which serve as nutrients for developing embryos and larvae (as constituents of the yolk sac) [12,19,20]. However, in recent years, increasing attention has been focused on transcripts and the remaining proteins that contribute to embryonic development. There are numerous studies showing that either transcriptomic [1,17,21–24] or proteomic profiles [25,26] may be used for predicting egg quality and indicating their importance in the developmental competence of eggs. A comparative transcriptomic-proteomic analysis of fully grown ovarian follicles (before ovulation) of zebrafish (*Danio rerio*) identified only a small number (approximately 60) of proteins [27]. Another study characterized the proteome and transcriptome of zebrafish embryos at a late developmental stage (i.e., 24 h post fertilization), linking both types of molecules to crucial developmental processes [10]. However, these studies did not consider the developmental competence of the eggs analyzed and referred to only a single model species. Data on integrated transcriptomic-proteomic analysis from other ecologically and commercially relevant species are still missing. Consequently, a large knowledge gap related to the characterization of the proteome and transcriptome in fully competent ovulated eggs still exists. Additionally, the link between transcriptomic-proteomic profiles and egg developmental competence is not available for teleosts. Such information could allow us to identify the processes for which the two types of molecules are responsible and would be indispensable in understanding their role in developing fish embryos. This may allow the identification of factors determining egg quality, and more generally factors involved in reproductive biology, in finfishes.

Egg quality in finfishes has serious implications for wild and captive stocks. It has been identified as a crucial factor determining the effectiveness of natural recruitment and commercial production [12,28,29]. Therefore, egg quality – defined as the ability of the oocyte to be fertilized and develop into normal embryos [30] – is attracting much attention from scientists. This allowed us to identify basic extrinsic (such as environmental or nutritional factors) and intrinsic (e.g., in vivo

aging) determinants of egg quality in finfishes [22,29,31]. However, despite the great progress that has been made, the processes involved in the determination of egg quality are still poorly understood. It should be emphasized that this knowledge is of great importance for understanding the reproductive biology of finfishes, the management of wild stocks and the aquaculture industry [12,29].

Most of the studies considering egg quality at the molecular level have focused on farmed or model fishes, where a similar approach of categorizing particular samples as either high or low quality has been implemented. This has included the collection of subsamples for molecular analysis immediately after egg collection and monitoring the developmental competence (by determining the embryonic survival rate at various stages of development) of an equivalent subsample during parallel incubation. This approach, although state-of-the-art [29], does not exclude any groups of eggs, a considerable disadvantage. The eggs collected from domesticated fish are of various qualities (including those showing symptoms of overripening and/or intraovarian aging processes), and therefore it is possible that molecular analysis may be performed in cells that are suffering from a lack of basic structural integrity and/or undergoing internal disintegration (also at the molecular level) [24,32–34]. The inclusion of eggs with considerably altered fertilization capacity (which is an important component of egg quality definition [30]), creates the risk of masking many important molecular relationships, limiting the further development of our knowledge. These approaches may also lead to the mixing of unfertilizable eggs with eggs having altered molecular machinery leading to early embryonic lethality – two distinct phenomena. Initial preselection was considered in a study by Zarski et al. [21], where eggs of seabass (*Dicentrarchus labrax*) characterized by apparent abnormalities (as practiced in commercial hatcheries) were discarded from analysis. Additionally, in a recent study on rainbow trout (*Oncorhynchus mykiss*), eggs characterized by overmaturation were excluded from the analysis [17]. These approaches provide new insights into the molecular mechanisms conditioning egg quality, highlighting the good potential of preselecting eggs for molecular analysis. However, in the study of Zarski et al. [21], the strategy undertaken (based on evaluating embryonic survival rate at the 4-cell stage, i.e., 3 h post fertilization) led to the analysis of groups of eggs featuring high variation in quality, which was another limiting factor in such studies. Therefore, a more careful preselection approach based on precisely chosen quality-related indices provides an opportunity to expand our knowledge and would allow us to focus on the molecular mechanisms leading to early lethality (and not related to a fundamental lack of fertilizability), which have been barely considered to date [17].

The aim of this study was to characterize the transcriptomic and proteomic profiles of high-quality eggs using pikeperch (*Sander lucio-perca*) as a model, the top predator of Holarctic freshwater bodies with high commercial interest [35,36], and to perform a comparative functional gene set enrichment analysis of the two types of molecules. Next, a comparative analysis of the transcriptomes and proteomes obtained from pikeperch eggs characterized by high and low quality – following a specifically developed preselection procedure – was performed to thoroughly characterize the respective contribution of maternally-inherited mRNAs and maternally-inherited proteins to egg developmental competence.

## 2. Material and Methods

### 2.1. Ethics statement

All procedures requiring animal handling (such as reproductive procedures) were performed in compliance with European and national regulations on animal welfare and ethics in experimentation on animals. Whenever appropriate, the procedures were approved by the Lorraine Ethics Committee for Animal Experimentation (CELMEA; APAFIS-2016022913149909).

### 2.2. General study design

The study comprised of two separate objectives:

1. Objective no. 1: aimed to explore the integrated transcriptomic-proteomic profile of high-quality eggs;
2. Objective no. 2: aimed at (i) comparative transcriptomic-proteomic analysis of eggs characterized by either high or low quality and (ii) identification of molecular markers of egg quality.

Eggs were collected and developmental competence (quality) was evaluated with the same methods in realization of both objectives (described below). Transcriptomic and proteomic analyses were identical for both objectives. Real-time qPCR validation of candidate genes was performed only for objective no. 2. In silico and functional gene set enrichment analyses were performed separately for the two scientific objectives undertaken, as specified below.

### 2.3. Broodstock management and reproductive procedures

Pikeperch broodstock (Asialor fish farm, Pierrevillers, France) grown in recirculating aquaculture system (RAS) with fully controlled environmental conditions was used for the study. The RAS was equipped with UV sterilizers, biological filtration and a drum filter to ensure appropriate solids removal. The oxygen levels were kept at >80% saturation, ammonia in the range between 0.2 and 1.1 mg L<sup>-1</sup>, and nitrite between 0.05 and 0.70 mg L<sup>-1</sup>. The temperature (within a range of 8–23 °C, f0.5 °C) and photoperiod (f1 min) were controlled automatically. The fish had already spawned 3 times before the study was conducted, meaning that they were identified as fully 'functional' spawners. To promote the annual gonadal cycle, the fish were exposed to annual fluctuations in the photothermal program described by Fontaine et al. [37] as further modified by Zarski et al. [38] so that the fish were exposed to a wintering period (with temperature below 10 °C) for 4 months. Fish were fed throughout the year at a rate of 0.2–1% biomass, depending on the temperature and apparent satiation, with compound feed (50% protein, 11% fat, 10% moisture, 1.55% crude fiber, 1.35% phosphorus, 9.5% ash, and 17.9% nitrogen-free extract; Le Gouessant, France) [38].

Controlled reproduction of the fish was performed according to state-of-the-art protocols. In total, 60 males and 60 females were reproduced following the methods described by Zarski et al. [38]. For induction of ovulation and spermiation, human chorionic gonadotropin (hCG; Chorulon, Intervet, France) was

injected into each fish intraperitoneally at doses of 250 and 500 IU kg<sup>-1</sup> for males and females, respectively. At the time of hormonal treatment, the maturation stage of each fish was scored with the use of catheterization (as described by Zarski et al. [14]). All the fish were identified to be in maturation stage I (representing postvitellogenin oocytes entering the FOM process) out of 6 distinct stages [14]. Next, the fish were periodically monitored for ovulation by gentle pressure of the abdomen. When a particular female was found to have ovulated, the eggs were collected into individual sealable, dry, plastic containers and kept at temperatures between 8 and 10 °C for less than 30 min prior to further procedures. Sperm were collected into dry disposable syringes at between 5 and 7 days following hormonal stimulation (as recommended by Zarski et al. [39]) and were stored at 4 °C prior to use (not longer than 15 min), as is practiced in commercial production.

Each collected egg batch was aliquoted into 3 portions. Portion 1 was subportioned into individually labeled cryotubes and immediately snap frozen in liquid nitrogen (later stored at -80 °C prior to proteomic and transcriptomic analysis). The second portion was subjected to a preselection procedure (as described further). Portion 3 was subjected to fertilization to assess the egg quality indices. Such a strategy allowed the independent handling of the eggs used further for observation of the embryonic development success (fertilization and incubation) and being not disturbed by any other procedures (such as the ones applied during the pre-selection procedures).

During the controlled reproduction, the photoperiod was set to 14 h of light per day (provided with neon tube lights), and the temperature was maintained at 12 °C from the day of injection until the end of the experiment. Before each manipulation (catheterization, ovulation control, gamete collection, etc.), fish were anesthetized with MS-222 (150 mg L<sup>-1</sup>; Sigma-Aldrich) [40].

### 2.4. Preselection procedure

The preselection procedure is illustrated in Fig. 1. The procedure included placement of the eggs on a plastic Petri dish (to prevent excessive surface adhesiveness of the eggs – a natural feature of pike-perch eggs) and their evaluation under the stereomicroscope for fragmentation of lipid droplets (an indicator of lower quality eggs in percids; [41]) or any other aberrations (symptoms of aging or lack of structural integrity). For the next step, only eggs of normal appearance were used. Next, the eggs were activated with hatchery water, and cortical reaction intensity was evaluated at 3–5 min post water activation, which is a specific quality indicator of pikeperch eggs [42]. The eggs exhibiting at least 90% cortical reaction were further observed for up to 30 min following activation. This was done to determine whether the eggs were able to form blastodiscs at the animal pole (indicating their fertilization capacity – for details, see Zarski et al. [42]), as only the eggs exhibiting blastodisc formation are capable of cellular cleavages following fertilization. The eggs exhibiting a lack of animal poles formation, internal damage or any other irregularities were not considered in further steps of the study (Fig. 1). This preselection procedure allowed us to discard the eggs exhibiting apparent symptoms of lower quality and consider only those eggs with a normal appearance and the

potential to be fertilized – an important component of egg quality definition [30].

Consequently, eggs characterized by early lethality were considered in our study what allowed to shed new light on the molecular processes that condition the developmental competence of the eggs. It should be highlighted that the observations of the cortical reaction and animal pole formation are not affected by the presence or absence of sperm [42] making this strategy fully feasible.

## 2.5. Gamete management and fertilization

To assess the egg quality indices, the eggs were subjected to controlled fertilization according to the procedure described by Roche et al. [43]. Briefly, 5 mL of hatchery water was poured over individually labeled glass Petri dishes (30 mm in diameter). Next, a sample of eggs (approximately 100 eggs at one time) was placed onto the Petri dish simultaneously with 50 µl of a sperm sample (pooled sperm obtained from three males each time). Next, the content of the Petri dish was hand-shaken for approximately 10 s to disperse the semen and promote fertilization. Later, the eggs were evenly dispersed across the Petri dish (by gentle hand-mixing) so that they could sediment and adhere to the bottom of the dish. After 5–10 min, each Petri dish was washed carefully with hatchery water to remove excess sperm and other residues. The eggs were further incubated while adhered to these Petri dishes in an individual 500 mL beaker. The water temperature during incubation was 12 °C. Two Petri dishes (duplicates) were established for each egg sample from each female.

For fertilization, freshly collected sperm samples exhibiting at least 80% cell motility (verified under a light microscope at  $\times 400$  magnification for each sample as described by Cejko et al. [44]) were always used. The sperm was pooled shortly before use for fertilization, as recommended by Schaefer et al. [45]. Using pooled sperm reduced the possible negative male effect on embryonic development, as it was previously reported for percid fishes [46].

### 2.5.1. Evaluation of egg quality indices and designation of samples for molecular analysis

During the incubation period, four egg-quality-related parameters were recorded: fertilization rate (FR) 8 h post fertilization, survival rate (SR) of embryos at 72 h post fertilization (both FR and SR were determined by direct counting of live and dead eggs under a stereomicroscope), hatching rate (HR) of larvae and deformity rate (DR) of hatched larvae. Determination of HR was conducted for 5 days (between the 9th and 14th days post fertilization), during which hatched larvae were removed each day from the incubators and counted manually. Next, the larvae were anesthetized (MS-222, 150 mg L<sup>-1</sup>), and each specimen was individually scored for the occurrence of larval deformity as described previously [47,48]. DR was calculated as follows:  $DR = 100 \times \text{number of larvae showing deformity} / \text{total number of larvae hatched}$ . The analysis was performed with GraphPad Prism 9 software (GraphPad Software, San Diego, CA, USA).

For characterization of the transcriptomic and proteomic profiles, 8 egg batches were characterized by high egg quality

(FR and SR > 85%, HR > 60%, and DR < 15%; [48,49]). For the comparative analysis of high- and low-quality eggs, the five samples characterized by the highest egg quality (FR and SR > 80%, HR > 75%) and the five samples characterized by the lowest egg quality (FR > 80%, SR < 40%, HR < 35%) among the preselected samples were used.

## 2.6. RNA extraction

Total RNA was extracted from 50 mg of eggs (approx. 50 eggs) using a Total RNA Mini kit (Cat. No. 031–25; A&A Biotechnology, Gdynia, Poland) following the manufacturer's protocol. Next, the genomic DNA from the samples was removed using a Clean-Up RNA Concentrator (Cat. No. 039-100C; A&A Biotechnology, Gdynia, Poland) according to the manufacturer's protocol. The concentration of RNA obtained was measured with a NanoDrop 2000 instrument (Thermo Fisher Scientific, Wilmington, DE, USA). The quality of the RNA was verified with a 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA). All the samples had RIN (RNA integrity number) values higher than 9.5. The total RNA obtained was aliquoted into two portions and stored at -80 °C prior to further analysis. One portion was used for transcriptomic profiling, and the other portion was used for validation of candidate genes with quantitative real-time polymerase chain reaction (qPCR).

## 2.7. Transcriptomic profiling

Transcriptomic profiling of the samples was performed with a pikeperch-specific microarray (8 × 60 k, Agilent), which was designed and successfully validated previously (for details, see Zarski et al. [48]). The microarray design can be accessed via the Gene Expression Omnibus [50] platform under accession number GPL27937. For analysis, the samples were randomly distributed on the microarray slides. Labeling and hybridization of the samples on the microarrays was performed according to the 'One-Color Microarray-Based Gene Expression Analysis (Low-Input QuickAmp Labeling)' protocol from the manufacturer (Agilent). Briefly, 150 ng of total RNA from each sample was amplified and labeled using Cy3-CTP. The yield (>1.65 µg cRNA) and specific activity (>9 pmol of Cy3 per µg of cRNA) of the Cy3-cRNA produced were checked with a NanoDrop instrument. From the Cy3-cRNA preparation, 1.65 µg was fragmented and hybridized on a subarray. Hybridization was carried out for 17 h at 65 °C in a rotating hybridization oven. Next, the array was washed and scanned with an Agilent Scanner (Agilent DNA Microarray Scanner, Agilent Technologies, Massy, France) with standard parameters for a 8 × 60 K gene expression oligoarray (3 µm and 20 bits). The Agilent Feature Extraction software (10.7.1.1) was used to extract the data following the appropriate GE protocol, and the data were further subjected to statistical analysis with GeneSpring GX software (Agilent Technologies, Santa Clara, CA, USA). The gene expression data were scale normalized and log(2) transformed prior to the statistical analysis. Raw data from the comparative analysis of eggs of high and low quality can be accessed via the NCBI Gene Expression Omnibus [50] under the GSE167376 accession number [the data will be made publicly accessible following publication; reviewers can access the data by using the following token: erkbmaaybjkjob].

For the transcriptomic profile analysis, only genes found to be expressed in at least 75% of the samples were used (the gold standard in microarray analysis). Differentially expressed genes (DEGs) between the groups representing high (HQ; n = 5) and low (LQ, n = 5) egg quality were identified with the use of Gene Spring GX software. The differences between the groups were analyzed by unpaired t-test according to the criteria of a minimum twofold change in expression and a p-value <0.01. In this analysis, the Benjamini-Hochberg correction was applied. Average linkage clustering analysis (Gene Cluster 3.0) was performed for the differentially abundant genes (unsupervised linkage).

## 2.8. Mass spectrometry-based analysis of egg proteins

Eggs were homogenized for 10 s using a homogenizer (Art-Micra D8, Micra GmbH, Heitersheim, Germany) and by centrifugation at a speed of 23,500 rpm through QIAshredder devices (Qiagen, Hilden, Germany). Samples were then diluted 1:1 with 8 M urea/0.4 M NH<sub>4</sub>HCO<sub>3</sub>, and the protein concentrations were determined with a Bradford assay [51]. Prior to digestion, proteins were reduced for 30 min at 37 °C using dithioerythritol at a final concentration of 5 mM and carbamidomethylated with iodoacetamide (final concentration 15 mM) for 30 min at room temperature. Samples of 10 µg total protein were digested for 4 h using 100 ng LysC (FUJIFILM Wako Pure Chemicals, Osaka, Japan). After dilution of the samples with water to 1 M urea, a second overnight digestion step using 200 ng modified porcine trypsin (Promega, Madison, WI, USA) was performed at 37 °C. One microgram of peptide was injected into an Ultimate 3000 RSLC chromatography system connected to a Q Exactive HF-X mass spectrometer (Thermo Scientific, Waltham, MA, USA). Samples were transferred to a PepMap 100 C18 trap column (100 µm × 2 cm, 5 µM particles, Thermo Scientific) at a flow rate of 5 µl/min using mobile phase A (0.1% formic acid and 1% acetonitrile in water). For separation, an EASY-Spray column (Pep-Map RSLC C18, 75 µm × 50 cm, 2 µm particles, Thermo Scientific) at a flow rate of 250 nl/min was used. For the chromatography method, a two-step gradient from 3% mobile phase B (0.1% formic acid in acetonitrile) to 25% B in 160 min and from 25% to 40% B in 10 min was used. Spectra were acquired in the data-dependent mode with a maximum of 15 MS/MS spectra per survey scan.

For data analysis and label-free protein quantification, MaxQuant (v.1.6.1.0) [52] and the Sander luciopeca sequences in the NCBI database were used. During the analysis of the proteomic profile of high-quality eggs, only the proteins identified in at least 50% of the samples (i.e., in ≥4 out of 8 samples) were considered abundant and considered in the functional analysis (criteria suggested by Li et al. [53]). Statistical evaluation of the data from comparative analysis of high- and low-quality eggs was performed with Perseus V1.5.3.2. Between-group comparisons were performed with a t-test (followed by permutation-based FDR calculation at a significance level of 5%). The mass spectrometry proteomics data have been deposited in the Proteo-meXchange Consortium via the PRIDE [54] partner repository with the dataset identifier PXD023229 [the data will become freely accessible following publication; reviewers can access the data by logging in to PRIDE <https://www.ebi.ac.uk/pride/> with the following details:

Username: [reviewer\\_pxd023229@ebi.ac.uk](mailto:reviewer_pxd023229@ebi.ac.uk); Password: MEXVDYmg].

## 2.9. In silico, Gene Ontology and functional analysis of transcriptomic and proteomic data obtained during realization of objective no. 1

For both transcriptome and proteome data, protein RefSeq accession numbers for each transcript and protein were obtained. Next, the RefSeq identifiers were indexed against 21,471 human proteins in Swiss-Prot (downloaded from the UniProt database [[uniprot.org](http://uniprot.org)] on 12 October 2020) with Diamond v0.9.22, followed by alignment of the sequences using BLASTP with an e-value <10 e<sup>-5</sup>. The resulting file was filtered to keep only the best match for each protein. This enabled us to retrieve gene names and UniProt accession numbers for successfully aligned proteins, which were further used to perform Gene Ontology analysis.

Gene Ontology (GO) and KEGG pathway analyses were performed using the ShinyGO online platform [55] with biological processes considered as targets. Two separate analyses were performed, as presented in Supplementary File 1.

The first GO analysis (Supplementary File 1a) aimed at comparative analysis of the full transcriptome and proteome. To this end, each data set (transcriptome and proteome) was evaluated following the same analytical approach. First, the higher level GO categories were obtained (100 categories for each data set) and compared to each other for the occurrence of common and distinct GO categories. Next, enrichment analysis was performed to obtain 500 significantly enriched (FDR < 0.05) terms for each of the data sets. The obtained terms were further subjected to network analysis based on hierarchical clustering (related GO terms were linked together only when they shared overlap of at least 50% of genes). This allowed us to identify clusters (at least four different ontological terms joined together) characterized further with additional enrichment analysis of the gene list for each cluster, which enabled us to identify biological processes characteristic of each cluster.

The second GO analysis (Supplementary File 1b) was followed by the identification of transcriptome-specific, proteome-specific and conspecific subsets of genes, as revealed by a Venn diagram [56]. Each of the subsets was later analyzed with the ShinyGO platform in three steps:

- Step 1: Enrichment analysis (identification of the 30 most enriched processes with FDR < 0.05) was performed;
- Step 2: Hierarchical clustering was performed to identify separate clusters (sharing at least 30% of genes), with the condition that each cluster include at least three ontological terms;
- Step 3: Additional enrichment analysis of the genes constituting separate clusters was performed to find the 3 most enriched terms representative of each cluster.



Additionally, for each subset of genes, the 5 most enriched KEGG (<https://www.kegg.jp/>) pathways were identified.

## **2.10. In silico, Gene Ontology and functional analysis of DEGs obtained from realization of objective no. 2**

For each DEG, a human homolog was obtained (as described in the previous section). Next, hierarchical clustering-based networking of the 30 most enriched (FDR < 0.05) biological processes was performed with the ShinyGO platform. GO terms were connected to each other only when they shared at least 50% of their genes. In this way, four different clusters connecting at least 3 different GO terms were identified. Next, the genes forming each cluster were once again subjected to GO enrichment analysis, enabling the identification of the 3 most enriched (FDR < 0.05) GO terms characterizing the analyzed DEGs. Additionally, all the DEGs were mapped to clusters of biological processes that were significantly enriched in the transcriptomic and proteomic profiles of high-quality eggs to determine their relevance in whole processes specific to either the transcriptome or proteome.

## **2.11. Designation of candidate gene markers of egg quality in pikeperch during the realization of objective no. 2**

To identify the most suitable candidate gene markers to enable pikeperch egg quality prediction, zebrafish gene identifiers for DEGs (identified during comparative analysis of high and low egg quality) were retrieved using ShinyGO. Next, the expression pattern (provided as the number of transcripts recorded per million reads; TPM) during zebrafish embryonic development for each gene was obtained from the Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) based on the data published by White et al. [57]. Next, the data were grouped into 5 developmental periods (pre-ZGA, blastula/gastrula, somitogenesis, prim stage and 2–5 days post fertilization) identified by White et al. [57] as representative crucial developmental phases. For these periods, the mean expression value was calculated. Furthermore, the data were log(2) transformed and arranged in ascending order based on the pre-ZGA phase as the phase during which the maternal transcriptome plays a major role. The data were visualized as a heatmap with a color scale representing the normalized

expression level using GraphPad Prism 9. As the most suitable candidate gene markers, 20 genes characterized with the highest expression during the pre-ZGA phase in zebrafish were considered and taken for qPCR validation.

## **2.12. qPCR validation of candidate gene markers of egg quality in pikeperch during the realization of objective no. 2**

Total RNA obtained from eggs characterized by either high or low egg quality was subjected to reverse transcription (1.7 ng of total RNA) using a RevertAid First Strand cDNA Synthesis Kit (Cat. No. K1622, Thermo Fisher Scientific) according to the manufacturer's protocol. The protocol involved oligo(dT)18 primers as well as an optional incubation step (5 min at 65 °C) aiming at the removal of secondary structures.

Real-time qPCR was performed for each of the selected candidate genes using a Vii7 (Applied Biosystems) thermocycler. For each qPCR (reaction volume 20 µL), 10 ng cDNA template was used along with DyNAmo HS SYBR Green qPCR Master Mix (Cat. No. F410XL, Thermo Fisher Scientific) and 0.5 µM forward and reverse primers (sequences of all the primers are provided in Supplementary File 2), which were designed with the Primer3Plus online platform [58]. The following cycling conditions were applied: enzyme activation for 10 min at 95 °C followed by 40 cycles of denaturation at 95 °C for 15 s and annealing and elongation at 60 °C for 1 min. After each amplification, melting curve analysis was performed to verify the amplification specificity and compare it with the predicted melting curve verified with uMELT [59]. In the analysis for each gene, a standard curve was calculated using a series of 6 twofold dilutions to determine reaction efficiency (reaction efficiencies between 85 and 110% were considered acceptable). Relative expression for each gene was normalized as the geometric mean of expression values recorded for 4 reference genes, which were chosen on the basis of their stable expression levels and close-to-mean expression values in the microarray analysis [21]. Each reaction for qPCR validation was performed in duplicate. The data were compared between the groups (low and high egg quality) using a t-test (GraphPad Prism 9). Differences between groups were considered significant at  $p < 0.05$ .

## **3. Results**

### **3.1. Scientific objective no. 1: transcriptomic-proteomic characterization of eggs**

#### **3.1.1. Transcriptomic profile of high-quality eggs**

Transcriptomic analysis revealed 10,238 expressed genes encoding unique proteins, including 9465 proteins for which human homologs could be identified, encoded by 9233 genes, which were considered in further analysis. The 4 most highly expressed genes, based on the expression level in the microarray, were *hdgf2*, *cldn4*, galactose-specific lectin

*nattectin-like* (for which no human homolog was found) and *mt-nd5*. All the identified genes are presented in Supplementary File 3.

#### **3.1.2. Proteomic profile of high-quality eggs**

The proteomic analysis allowed us to identify 1450 proteins (Supplementary File 4), of which 806 could be detected in at least 50% of the samples and were considered abundant. The most abundant proteins were fish-specific vitellogenins and galactose-specific lectin *nattectin-like* proteins. For 790 of these 806 abundant proteins, human homologs could be found, which corresponded to 699 unique proteins taken for further Gene Ontology analysis. The most abundant proteins

for which human homologs were found were zona pellucida sperm-binding proteins 3 and 4 (encoded by zp3 and zp4, respectively). All the proteins considered abundant are presented in Supplementary File 5.

### **3.1.3. Functional analysis of transcriptome and proteome datasets**

The transcriptome and proteome datasets were each grouped into 100 higher level Gene Ontology biological process categories (Supplementary File 6). Among 94 conspecific categories, 10 were found to cover various processes linked to reproduction, 6 were related to development, and 12 were related to the immune system (Fig. 2). The remaining categories covered many important but very general biological processes related to, among others, methylation, cell function and maintenance, regulatory processes, protein processing and cell responses to various factors. Among the 6 proteome-specific terms, the acrosome reaction process included only two proteins: zona pellucida 3 and 4, which were among the most abundant proteins. It should be highlighted that of the 6 transcriptome-specific higher level Gene Ontology terms identified, three were found to comprise genes involved in various processes affecting behavioral features (i.e., multiorganism behavior, regulation of behavior and intraspecies interaction between organisms) (Supplementary File 6) indicating neurogenesis to be transcriptomic feature of the eggs.

Enrichment analysis yielded the 500 most enriched GO terms for each dataset (i.e., transcriptome and proteome) (Supplementary File 7). Among the transcriptomes, the most enriched (with  $FDR < 3.08e-09$ ) processes were clustered into several processes, with neurogenesis being highly represented and specific only for transcriptomic profile (Fig. 3).

The most enriched processes ( $FDR < 2.12e-05$ ) identified on the basis of the proteome were clustered into six distinguishable functional categories. From among all these terms (presented on Fig. 4) immune response should be highlighted as the one being specific only to the proteomic profile of pikeperch eggs.

### **3.1.4. Functional analysis of genes specific to or shared between the transcriptome and proteome**

Among all the genes and proteins identified to be present in the high-quality eggs of pikeperch, 509 were found to be conspecific to both data sets. Among all the transcripts identified, 8724 did not correspond to proteins in the proteome, and among all the proteins identified, 190 were not encoded by genes identified in the transcriptome (Fig. 5).

The set of transcriptome-specific genes could be organized into 5 Gene Ontology clusters (sharing at least 30% of genes) of the 30 most enriched biological processes (as shown on Fig. 6a), with neurogenesis being again transcriptome-specific ontological term. The involvement of transcriptomic cargo in neurogenesis has additionally been confirmed by KEGG pathway enrichment analysis which identified axon guidance among the most enriched ones (Fig. 6d). Detailed Gene

Ontology analysis of the groups of genes constituting each cluster confirmed their relevance to the processes in which they were grouped (Supplementary File 8).

Gene Ontology analysis of the proteins identified in pikeperch eggs revealed 4 different clusters comprising 28 out of the 30 most enriched biological processes (Fig. 6b). The data set analyzed here contained hits related to the ribosome and proteasome among the top 5 most enriched KEGG pathways. In addition, it should be highlighted that the proteome-specific gene set contained proteins related to metabolic pathways (similar to transcriptome-specific pathways) as well as protein modification activities, though oxidative phosphorylation appeared to be the primary pathway enriched (Fig. 6d). Furthermore, detailed Gene Ontology analysis of proteins constituting particular clusters indicated that the first two clusters were actually mostly related to translational processes. The proteins constituting the third cluster, comprising 13 Gene Ontology terms, were found to be related to various catabolic processes, including catabolism of RNAs and proteins. Proteins identified in the fourth and final cluster were found to be involved in the biosynthesis of organonitrogen compounds, with the biosynthesis of amides presumed to be primarily targeted in this process. KEGG pathway analysis of the genes constituting different clusters showed that the proteome-specific gene set was mostly related to the ribosome pathway, regardless of the cluster analyzed. In addition, it should be highlighted that the proteasome was among the most enriched KEGG pathways (Supplementary File 8).

Functional analysis of the gene set shared between the transcriptome and proteome allowed us to distinguish four clusters in which 27 out of 30 most enriched Gene Ontology terms were grouped (Fig. 6c). The genes in this gene set were found to be involved in similar KEGG pathways to those identified for the proteome-specific pathway. This includes ribosome, proteasome and metabolic pathways accompanied by carbon metabolism, among others (Fig. 6d). Specific functional analysis of the gene set constituting each of the Gene Ontology clusters revealed similarities with the analysis of the proteome-specific set of genes highlighting processes being under the control of two sets of molecules analyzed, at the same time. This includes translation process (with two clusters of genes involved in this process), catabolic processes and organonitrogen compound biosynthetic processes. KEGG analysis revealed that in each cluster, the most enriched pathway was ribosomes, followed by carbon metabolism (in three clusters), proteasomes (in two clusters) and others, including RNA transport and metabolic pathways (Supplementary File 8).

## **3.2. Scientific objective no. 2: transcriptomic-proteomic profiling of egg quality**

### **3.2.1. Evaluation of egg quality**

Among the eggs exhibiting formation of the blastodiscs at the animal pole within 1 h post fertilization, two groups could be distinguished, i.e., eggs of high and low quality. Significant differences in embryonic survival could be detected at 72 h post fertilization (after ZGA). In the high-quality group, over 70% of larvae hatched, whereas in the low-quality group, less than 20% hatched. In both groups, less than 10% of hatched



larvae were found to exhibit developmental deformities (Fig. 7).

### 3.2.2. Transcriptomic analysis of high- and low-quality eggs

Within the comparative analysis of the transcriptomes of high- and low-quality eggs, successful hybridization was recorded for 17,243 probes. Statistical analysis revealed 85 differentially expressed genes (DEGs) (FDR < 0.01 with at least a 2-fold change) (Fig. 8a, Supplementary File 9). Only 2 DEGs were found to be downregulated in high-quality eggs (*ctnn2b*, *msn*), and the remaining 83 DEGs were down-regulated in low-quality eggs. Gene Ontology analysis revealed four different clusters (with at least 3 terms and sharing at least 50% of genes) that were found to be related to RNA splicing (two clusters), chromosome organization and cell junction organization. The most enriched term – organelle organization, associated with 31 DEGs – was not connected to any cluster (Fig. 8b). It should be highlighted that 33 DEGs were found to be so involved in neurogenesis (as shown o Fig. 3) suggesting that genes involved in this process were also considerably affected by early lethality in pikeperch.

## 4. Discussion

---

### 4.1. Scientific objective no. 1

#### 4.1.1. . Molecular profile of high-quality eggs

Our study provides novel insights into the understanding of the role of maternally-derived molecular cargo in fish. The two data sets contributed similarly to the highest level GO profile (Fig. 2). Both groups were associated with many processes that are critical for reproduction, development and the immune system. The same processes, while being very general, were already reported to be important components of the developmental competence of embryos in finfishes [17,21,60]. However, one unique feature of the transcriptome in this analysis that sheds light on the regulation of behavior should be mentioned. In a recent study, Zarski et al. (2020) indicated that genes involved in neurogenesis are highly modulated by domestication levels in the eggs of pikeperch. This suggests that the predetermination of nervous system development by maternally-inherited mRNAs should receive more attention. In particular, maternally-inherited mRNAs would shape embryo neurodevelopment and possibly the future behavior of the fish [61]. In contrast, in regard to proteomic-specific high level GO categories, the acrosome reaction category attracts attention, as it is well known that spermatozoa of teleost fishes have no acrosome [62]. This category included zona pellucida proteins, which are important components of egg envelopes that play various roles during oogenesis, fertilization and embryonic development [63]. It is worth mentioning that various forms of zona pellucida proteins were identified in the pikeperch egg proteome (Supplementary File 4), confirming their importance in the cells analyzed.

Further analyses show, for the first time, that transcripts and proteins have distinct, yet complementary, functions in the egg of teleost fish. The clustering analysis of most enriched GO terms indicates certain unique functionalities for each of the two data sets analyzed. It should be highlighted that among

By analyzing the expression levels of DEGs during the embryonic development of zebrafish, it was found that the expression levels of only 34 DEGs increased after ZGA. The remaining 47 DEGs were characterized by lower expression following ZGA (Fig. 9). qPCR validation of the 20 DEGs characterized by the highest expression before ZGA in zebrafish confirmed a significant difference (t-test,  $p < 0.05$ ) for 85% of them (Fig. 10). Interestingly, 14 of them were also found to take part in most represented biological processes characterizing transcriptomic portrait of high-quality eggs in pikeperch (as shown on Fig. 3).

### 3.2.3. Proteomic analysis of high- and low-quality eggs

Proteomic analysis of eggs of different quality enabled the identification of 946 proteins. Statistical analysis did not reveal any proteins with significantly different abundance between high- and low-quality eggs (FDR > 0.05).

the specific GO terms, those involved in neurogenesis were characteristic of the transcriptome (Fig. 3, Supplementary File 7). On the other hand, high representation of molecules responsible for the immune response could be identified as characteristically enriched in the proteome and was not found in the top 500 enriched processes for the transcriptome data set (Supplementary File 4). Obviously, immune response genes are important mRNAs identified in eggs of percids [64] and in pikeperch (Supplementary File 3). However, given that translation is quiescent, it seems to be more likely that the proteins already present in the egg play a main role in the defense against pathogens in eggs and early embryos. Considering that several important immune response translation proteins are among the most abundant proteins (next to vitellogenins and zona pellucida proteins), including galactose-specific lectin natectin-like proteins [65], it is clear that defense against pathogens is among the priorities for fish eggs. It was previously suggested that immune response-dependent protein profile in eggs can be modulated by the rearing environment in which females are kept [49]. Therefore, it should be highlighted that health and welfare status of broodfish before FOM, during which the majority of proteins are deposited [12], may affect the capacity of the eggs and developing embryo to tackle pathogens. However, the role of the transcriptome in defense mechanisms should not be neglected, because genes encoding immune response proteins are also present in the egg at very high levels, including the galactose-specific lectin nat-tectin protein (third most expressed gene). This indicates that this protein may play a crucial role throughout embryogenesis as a first line of immune defense. However, further, more dedicated study of this issue is required to draw valid conclusions.

#### **4.1.2. Role of transcriptome-specific, protein-specific and conspecific molecules**

Molecular profiling revealed characteristic features of the transcriptome-specific subset of genes. The mRNA contained in the oocyte was found to be considerably enriched in processes related to transcription, protein phosphorylation, biogenesis of cellular compounds and localization and transport of the proteins. This highlights that maternally-inherited transcriptomic cargo plays a crucial role in the intracellular 'management' of the proteins produced but is also responsible for gene expression (features also recorded at the level of the whole transcriptome; see Fig. 3). However, the most unique characteristics were associated with the genes involved in neurogenesis; this was additionally confirmed by KEGG pathway enrichment analysis, which identified axon guidance as among the most enriched pathways (for details see Supplementary File 8). It should be emphasized that axon guidance was already reported to be modified by the domestication process in pikeperch eggs [48]. Therefore, our study clearly indicates that development of the nervous system is a characteristic 'maternal legacy' in the ovulated egg and possibly in the developing embryo as well. This again brings attention to behavioral traits as factors conditioning adaptability of the fish to particular habitats and/or culture conditions.

Functional analysis of proteome-specific and conspecific subsets of data for the two types of molecules indicates the high importance of translation under the control of both types of molecules. This is in accordance with previous reports indicating that translation is among the crucial processes conditioning developmental competence in fishes [1,21], similar to the case in other animals [18]. In addition, clustering revealed the importance of protein and nucleic acid catabolism as well as organonitrogen compound biosynthesis, for which the cell is prepared at both the transcriptomic and proteomic levels. Interestingly, the KEGG pathway enrichment analysis highlights, along with metabolism or oxidative phosphorylation, two major elements: the ribosome and the proteasome (see also Supplementary File 8). This additionally confirms the very great importance of intracellular biosynthesis of proteins and their degradation, suggesting extensive remodeling within the eggs. Zarski et al. [21] reported that genes involved in ubiquitination (intracellular protein degradation) are crucial for the developing embryo. Additionally, Chapman et al. [23] reported that the expression level of ubiquitin-related genes in ovaries was related to developmental competence of eggs in striped bass (*Morone saxatilis*). Our study confirms that the potential for translation and degradation of proteins and nucleic acids is among the most important features of ovulated fish eggs.

### **4.2. Scientific objective no. 2**

#### **4.2.1. Molecular profiling of egg quality**

Egg quality recorded during the controlled reproduction of pike-perch is usually highly variable [35,38]. In fish, such variability is derived from both unfertilizable eggs (e.g., overripened eggs) and eggs that are competent for fertilization but die during the early stages of embryonic development [17] due to either maternal or paternal causes. However, in our

study the latter has been limited by using pooled sperm from several males what has been reported to compromise negative paternal effect on embryonic development in percids [46]. Therefore, maternally derived early lethality, still barely studied in fish, could be considered in our study thanks to the specific preselection procedure applied (Fig. 1). This allowed us to find, for the first time in percids, that such early embryonic lethality significantly contributes to the overall egg quality. However, it should be noted that the eggs exhibiting high early embryo lethality did not contribute to an observed increase in the incidence of larval deformity, which is a growing and serious concern in percids [47]. Therefore, the eggs preselected for molecular analysis in our study were characterized by high fertilization capacity but different developmental potential following fertilization. Thus, we were able to focus on the molecular profile being predictive of altered 'machinery' leading to developmental failure at early stages, which was done for the first time with such complex molecular profiling.

Among the two types of molecules investigated in our study, only the transcriptomic profile was found to be predictive of egg quality, whereas proteomic profiling did not reveal any significant differences between high- and low-quality eggs. This is in contrast to other studies revealing differentially abundant proteins in eggs of different quality. For example, in Eurasian perch [26] and in zebrafish [25], respectively, tens and hundreds of differentially abundant egg-quality-dependent proteins were detected. Such discrepancies between the published studies and the data obtained in our study may arise for various reasons. In the case of Eurasian perch reported by Castets et al. [26], it can be presumed that the sampling strategy, without any preselection, could affect the overall proteomic profile observed. However, in the case of zebrafish, almost all the eggs analyzed developed to the approximately 256-cell stage, indicating that embryonic development failed after successful fertilization. However, the main reason for the discrepancies observed could be associated with the different proteomic approaches that were used. In the study of Yilmaz et al. [25], proteomic profiling was performed after initial molecular weight-based fractionation of the proteins, which increased the number of quantified proteins. However, in the case of zebrafish, highly abundant proteins (such as vitellogenins, zona pellucida proteins or ribosomal proteins) were also found to be differentially abundant between high- and low-quality eggs. Therefore, it could be assumed that our straightforward LC-MS/MS approach without prefractionation allowed us to identify egg-quality-related proteins and should detect protein-dependent processes affecting egg quality. Of course, the discrepancies in overall findings between the studies may also be explained by species-specific factors and egg characteristics. Species specificities in the molecular profile of eggs have also been evidenced at the transcriptome level in a comparative analysis of cod and trout eggs [66]. Nevertheless, it is clear that we did not find strong egg quality-dependent proteome alterations in pikeperch, which could also be related to the applied preselection procedure. This may suggest that early embryo lethality in this species is determined after the generation of the protein reservoir in the egg (ending before the start of the FOM process; [12]), which highlights the fact that processes occurring during the FOM are strong modulators of egg quality. This finding is in agreement with a recent study suggesting that the processes directly preceding ovulation, and not those occurring during oogenesis, are the main modulators of egg quality [17].

However, to draw final conclusions, more detailed research is still needed.

#### 4.2.2. Candidate markers of egg quality

Among 85 DEGs identified to be predictive of early embryo lethality in our study, only 2 (*ctnnd2* and *msn*) were downregulated in high-quality eggs, whereas the remaining genes were downregulated in low-quality eggs (Fig. 8). The identified DEGs were found to be enriched in organelle organization, RNA splicing, chromosome organization and cell junction organization. Interestingly, none of these ontology terms have been previously reported to be clearly related to egg quality in teleosts. However, when mapping these DEGs to the processes enriched in the whole transcriptome (see Fig. 3) and proteome (Fig. 4), it becomes apparent that this set of genes plays multiple roles in the cell, including in protein modification, transcription, and translation and in cellular signaling, which were previously reported to have egg-quality-modulatory effects in other species [17,21,24,60,67]. Interestingly, a high number of DEGs (33) were also found to play a role in neurogenesis, representing the first time that egg quality markers have been linked with this important biological process. These observations highlight the added value of the combined ontological analysis (typical enrichment analysis and global overview of the whole transcriptome) implemented here, as it provides new insights into the roles of DEGs in the developmental competence of pikeperch eggs and in our study draws attention to the multifunctionality of the gene set. Moreover, it highlights particular genes as modulators of developmental competence rather than involved solely in the processes for which this gene set is enriched. From this point of view, our study provides novel insights into the processes determining egg quality.

A comparison of published data with the set of DEGs identified in the present study indicates the uniqueness of the data obtained since none of the candidate genes obtained for pikeperch was reported to be quality-dependent in various fish species [1,17,21,22,68,69]. Therefore, we examined at the expression profile of all the DEGs during embryogenesis in zebrafish based on data published by White et al. [57]. This allowed us to find that not all the genes were expressed at the zygote stage in zebrafish, despite being relatively well expressed in pikeperch eggs (see Supplementary File 3). This includes *slc4a2*, *slc35d1*, *mapk12*, *msn*, *nmrk2* and *slc6a7* (Fig. 8). This analysis reveals that the maternal transcriptomic cargo contained in the eggs may vary among species. A similar conclusion was drawn by Myers et al. [70], who reported that genes constituting quality markers in other species (as reported by Sullivan et al. [67]) were also poorly

expressed or not at all expressed in ovulated eggs of hybrid catfish until neurulation. Interspecies differences in transcript abundance in eggs between cod and rainbow trout were also reported by Wargelius et al. [66]. This pattern highlights that the interspecies comparison of the molecular markers related to egg quality may suffer from very large biases. It also allows us to suggest that the molecular mechanisms driving early embryogenesis may also vary across species. However, this needs to be studied in greater detail, with similar analytical workflows for data collection (laboratory techniques, bioinformatics, etc.) to draw valid conclusions. Nevertheless, our study is the first to provide a set of molecular markers of early embryo lethality in pikeperch, which can be a valuable tool in further studies on the developmental competence of eggs in this important fish species.

The selection of candidate genes based on expression patterns during zebrafish embryonic development led to the identification of 20 genes that are crucial for embryonic development (Fig. 9). Importantly, 17 of them (85%) were successfully validated with qPCR and successfully mapped to the processes characterizing transcriptomic portrait of high quality eggs (see Fig. 3). Therefore, the current study provides a set of solid candidate quality markers crucial for development in fish and other animals. For example, *h3f3a* was found to be an essential maternal factor for successful oocyte reprogramming in mice [71], and *snrpc* was previously reported to play a crucial role during amphibian and fish development [72]; *cldn7* was responsible for maintaining blastocyst development in pigs [73], whereas *ctnna1* was reported to be important for the development of the heart during mammalian embryogenesis [74]. Moreover, *ctnnd1* expression, which is dependent on the level of luteinizing hormone, was found to be crucial during maturation in human oocytes [75], suggesting that this gene could be a valuable marker of not only egg quality but also FOM progression. Another gene, *smarca4* (also known as *brg1*), was reported to regulate ZGA in the mouse [76], which makes it a highly promising candidate gene for monitoring early embryonic lethality stemming from failure during ZGA in fishes. Several identified genes are important in mammals for processes related to the progression of oogenesis (*sub1*; [77]), oocyte development (*sumo2*; [78]) and meiosis (*sepin2*; [79]), further highlighting the possibility of using these markers in developmental studies in fishes. In the case of the remaining positively validated transcripts, such as *u2af115*, *cyca*, *brd3*, *ivns1abp*, *cnih1*, *morf41* and *ptms*, their involvement in early embryogenesis remains to be explored in fishes and other animals, as they likely constitute a valuable set of candidate genes.

## 5. Conclusion

---

Our study provides novel insight into the role of maternally-inherited molecular cargo in finfishes. The approach taken in our study sheds light on the importance of the transcriptome in the development of the nervous system, which confirms the role of neurogenesis-related mRNAs as very important nongenetic heritable factors [7,80]. On the other hand, proteomic analysis highlights the crucial and specific role of proteins in the immune response in ovulated eggs. Additionally, integrated transcriptomic-proteomic analysis draws attention to the galactose-specific lectin *nattectin* gene

and protein as the frontline defense molecule for the egg and – most likely – the developing embryo.

The molecular analysis of egg developmental competence emphasizes postvitellogenic processes (FOM and ovulation) as those compromising the transcriptomic profile but not affecting proteomic cargo. This brings attention to the need for careful reconsideration of the pre-spawning conditions that the fish are exposed to and reproductive protocols (previously reported to affect egg quality [38] as well as transcriptomic profile [22]) as very strong modulators of egg quality and their

molecular structure. Although the mechanisms driving these alterations and the consequences stemming from the differential abundance of the transcripts remain to be explored, the candidate quality markers provided in this study are the first to be identified in pikeperch and represent valuable resources for further studies on the reproductive biology of this and other fish species. All of these results are of major

importance for understanding the influence of external factors on reproductive fitness in both captive and wild-type fish species.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ygeno.2021.09.003>.

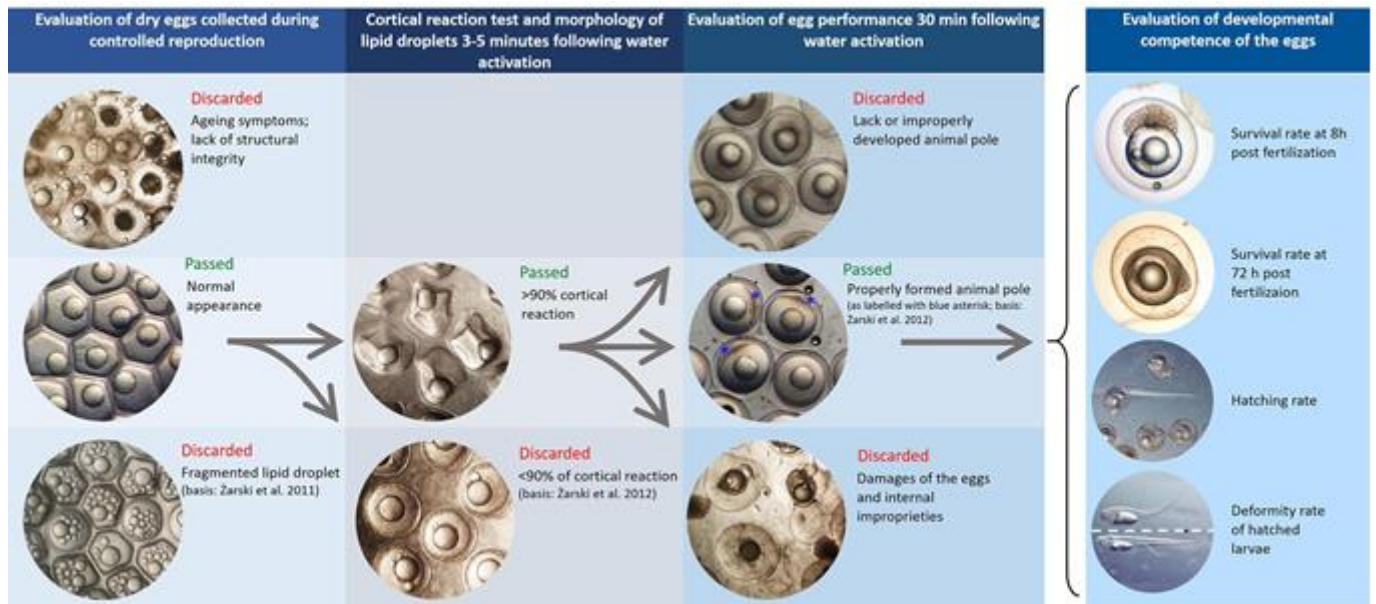
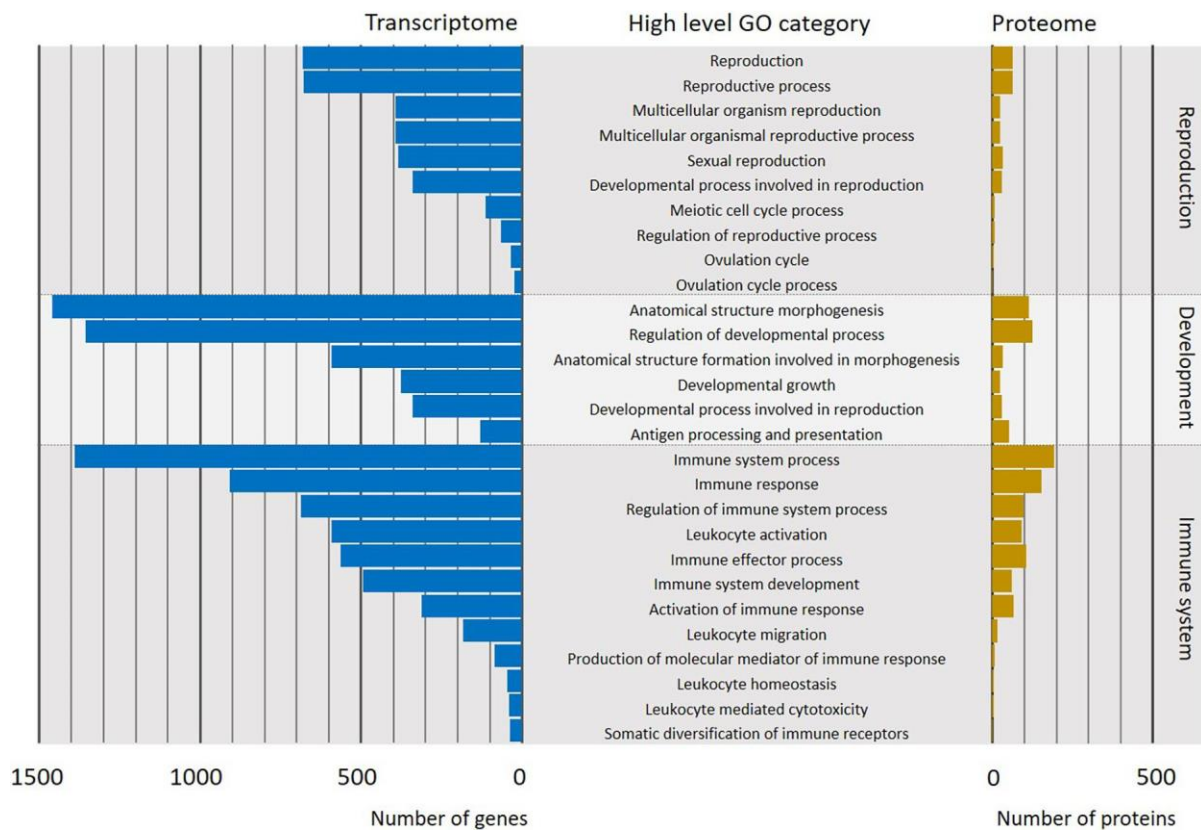
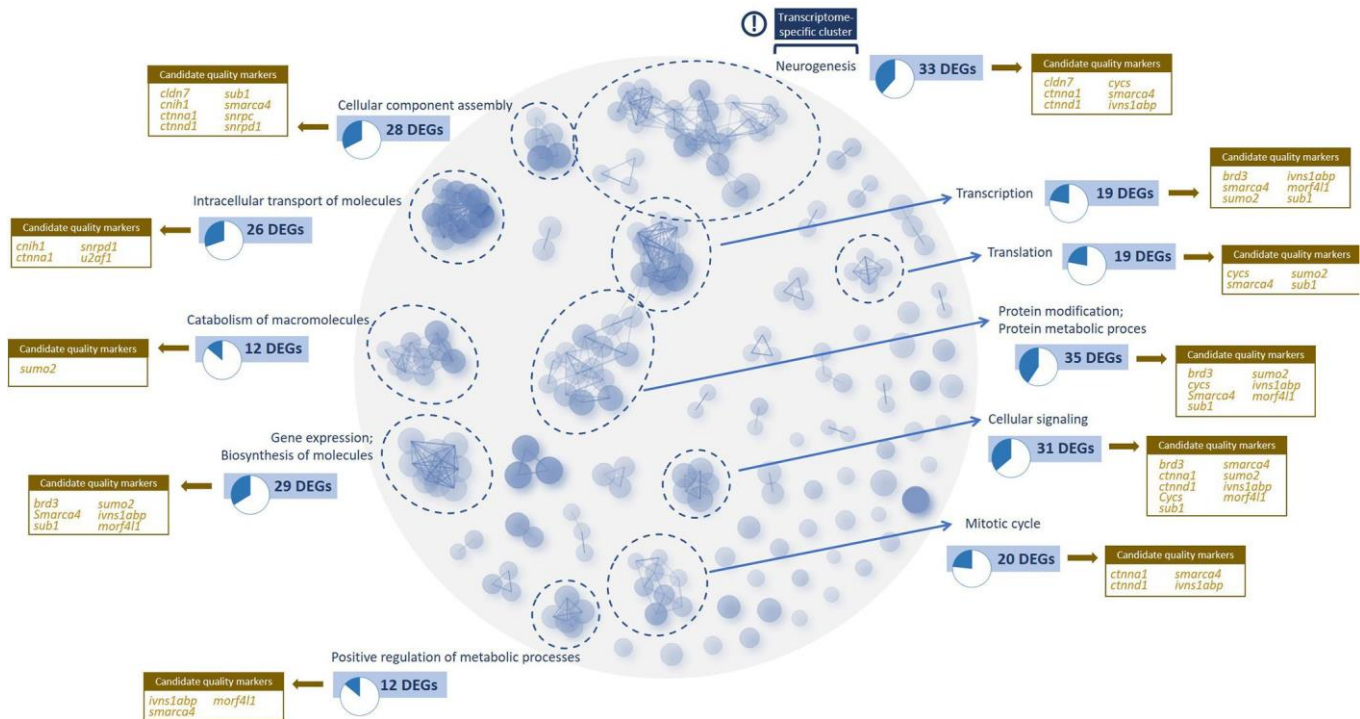


Figure 1. Illustration of the preselection strategy of the pikeperch eggs for the comparative transcriptome and proteome analysis in relation to the developmental competence of the eggs.



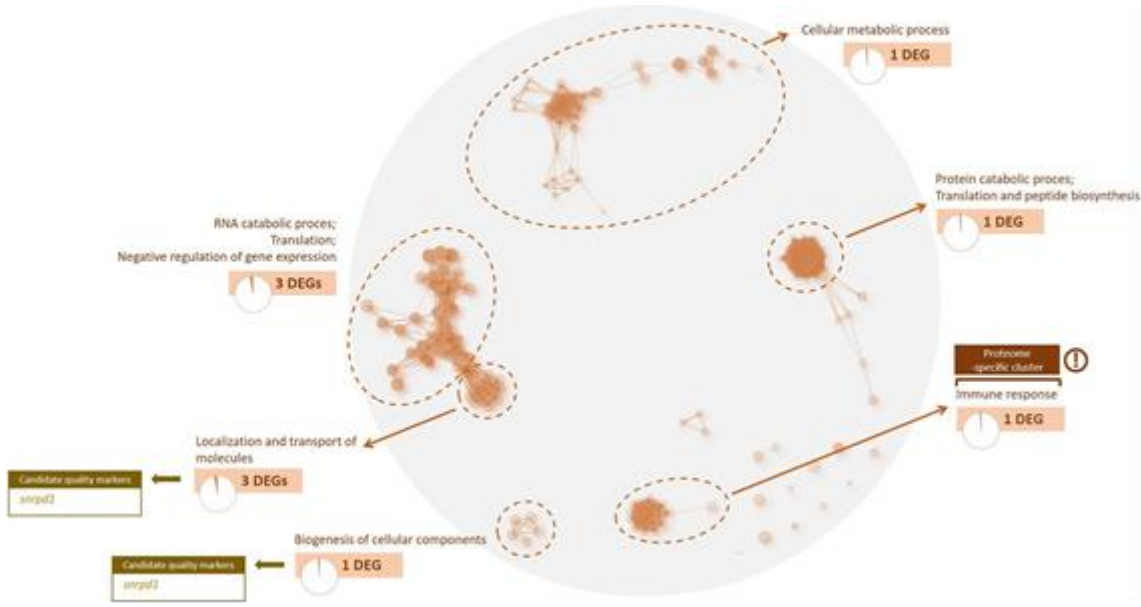
**Figure 2.** Number of genes and proteins associated with higher level Gene Ontology (GO) categories related to reproduction, development or immune system identified in the transcriptome and proteome of high-quality pikeperch eggs (n = 8). For further details, see also Supplementary File 6.





**Figure 3. Clustering of the 500 most enriched biological processes (BPs) obtained during the functional enrichment analysis (FDR < 0.05) of all the genes identified in the transcriptome obtained from high-quality pikeperch eggs (n = 8).**

For further details, see also Supplementary File 7. Circled clusters are those comprising at least four different BPs, for which additional functional analysis was performed to identify the main BP. Additionally, the number of differentially expressed genes (DEGs) identified between high- and low-quality eggs (for details see also Supplementary File 9) mapped to each of the main BPs were visualized. Process marked with encircled exclamation mark indicates transcriptome-specific cluster. In the boxes aligned to particular cluster successfully validated with qPCR candidate egg quality markers are shown.



**Figure 4. Clustering of the 500 most enriched biological processes (BPs) obtained during the functional enrichment analysis (FDR < 0.05) of all the proteins identified in the proteome obtained from high-quality pikeperch eggs (n = 8). For further details, see also Supplementary File 7.**

Circled clusters are those comprising at least four different BPs, for which additional functional analysis was performed to identify the main BP. Additionally, the number of differentially expressed genes (DEGs) identified between high- and low-quality eggs (for further details, see also Supplementary File 9) mapped to each of the main BPs were visualized. Process marked with encircled exclamation mark indicates transcriptome-specific cluster. In the boxes aligned to particular cluster successfully validated with qPCR candidate egg quality markers are shown.

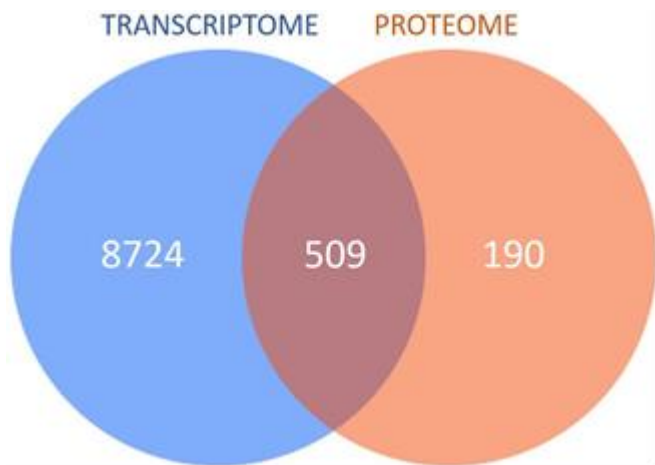


Figure 5. Venn diagram showing the number of transcriptome-specific, proteome-specific and conspecific genes and associated proteins identified during the transcriptomic and proteomic profiling of high-quality pikeperch eggs (n = 8).

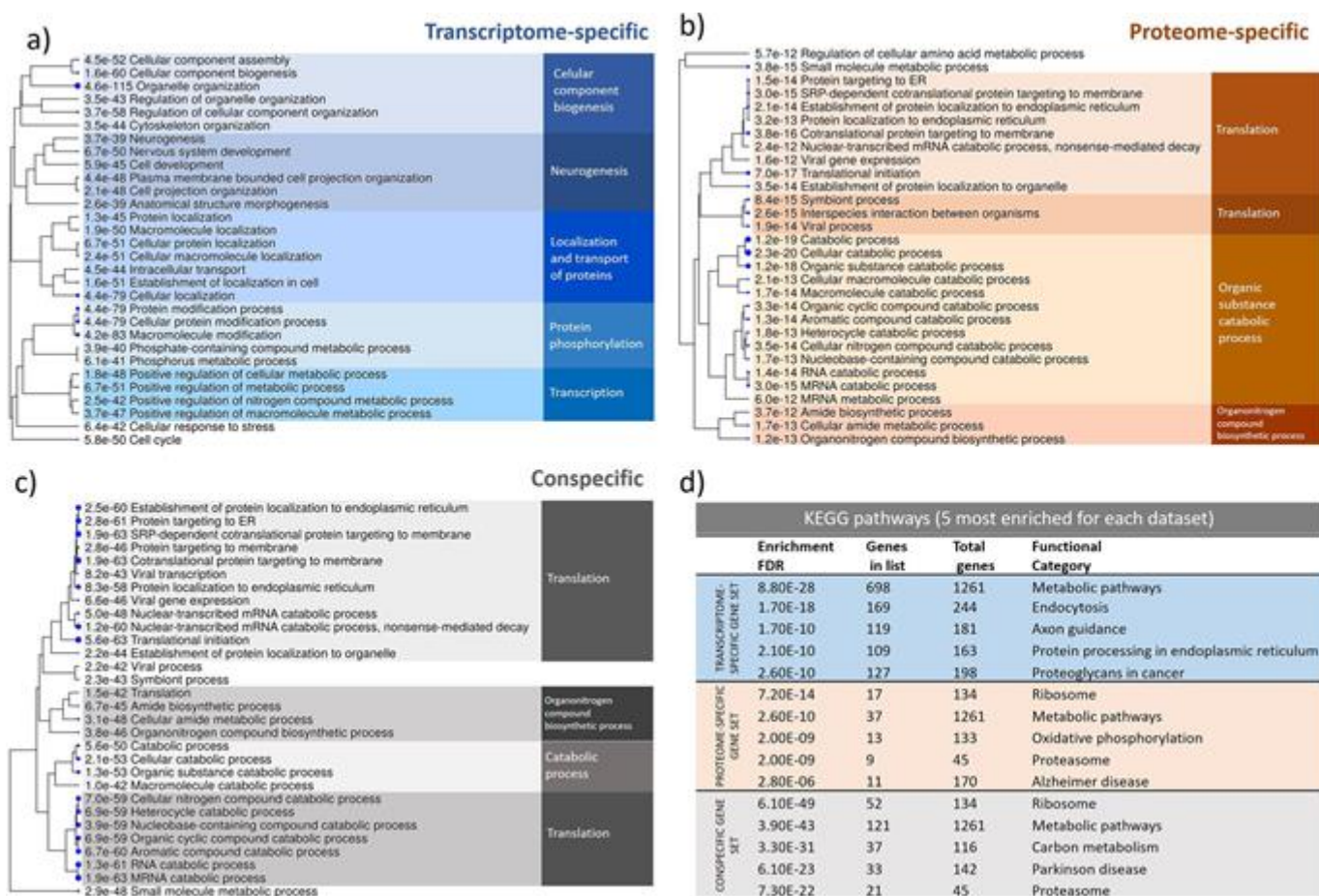
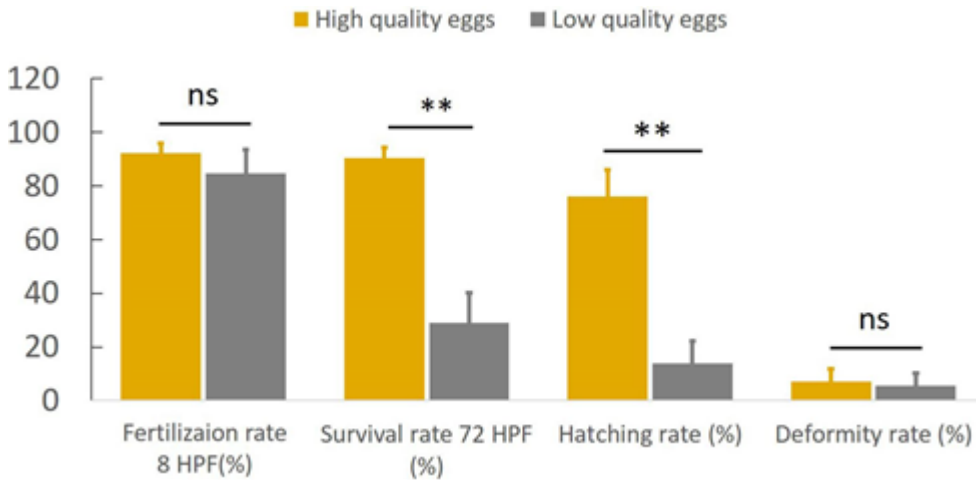
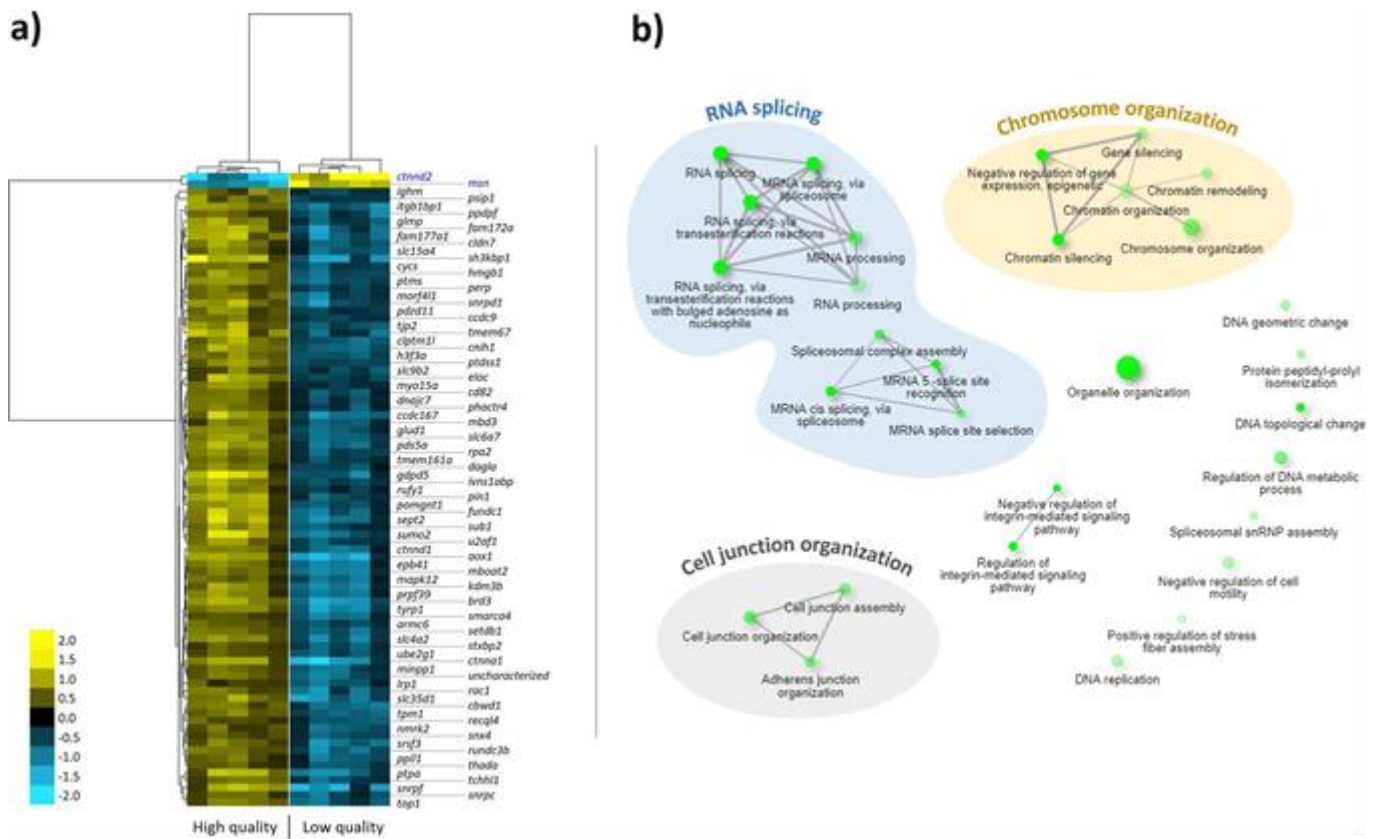


Figure 6. Clustering analysis of the 30 most enriched (FDR < 0.05) biological processes (BPs) for (a) transcriptome-specific, (b) proteome-specific and (c) conspecific sets of genes and associated proteins identified during transcriptomic and proteomic profiling of high-quality pikeperch eggs (n = 8). Entities constituting each cluster (grouped whenever the terms shared at least 30% of the genes/proteins) were again subjected to enrichment analysis to identify the main BP (for further details, see also Supplementary File 8). Additionally, the five most enriched KEGG pathways for each subset of data are presented in panel 'd'



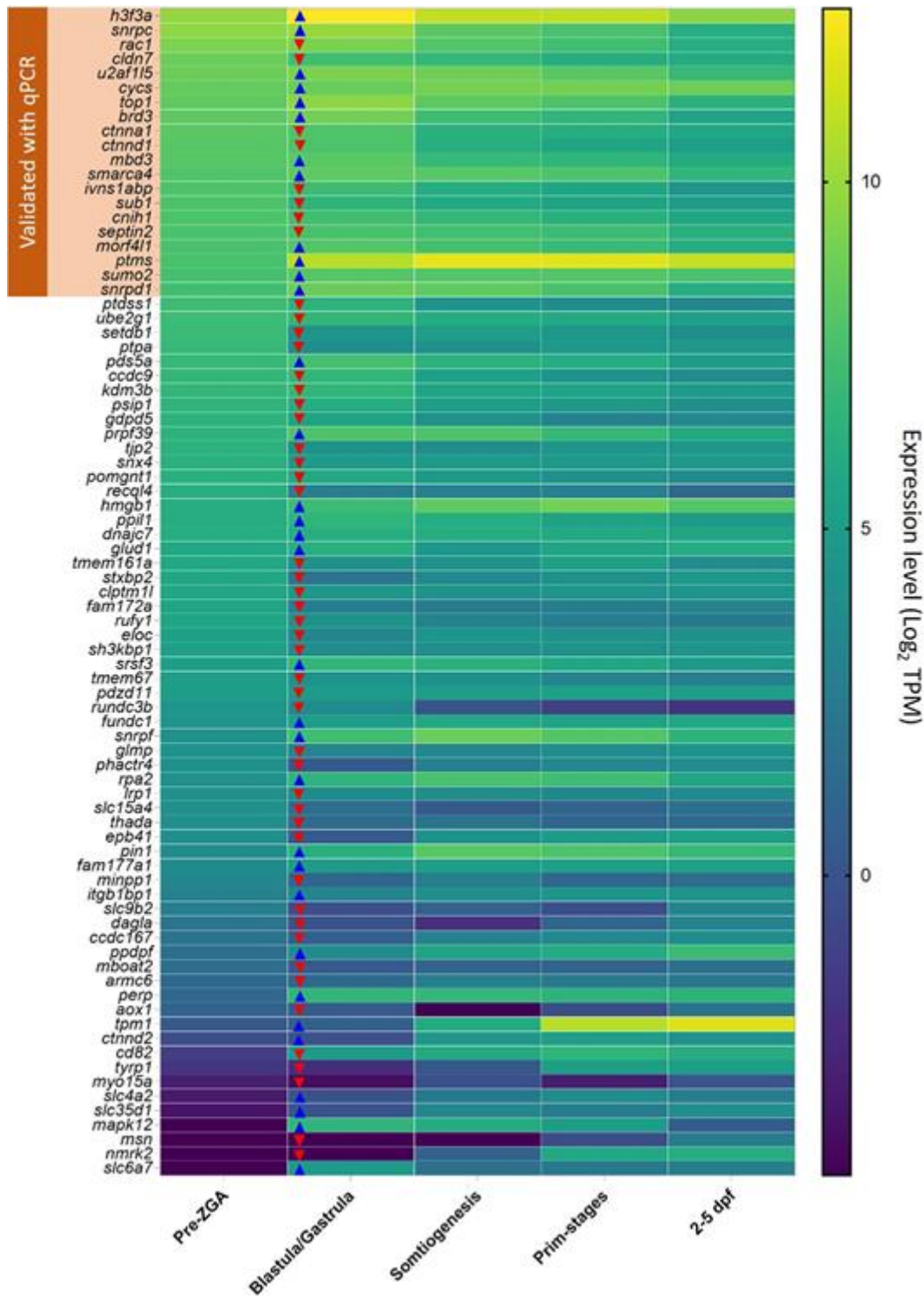
**Figure 7. . Results of evaluation of developmental competence of the preselected eggs assigned to high and low egg quality groups (n = 5 for each group). ‘HPF’ stands for ‘hours post fertilization’; ‘ns’ stands for ‘nonsignificant’ (highlighting that no significant differences were detected;  $p > 0.05$ ). Double asterisks indicate that the data between the groups were significantly different ( $p < 0.01$ ).**



**Figure 8. Panel ‘a’: Unsupervised average linkage clustering of 85 differentially expressed genes (DEGs) between high- and low-quality pikeperch eggs (n = 5 for each group).**

Each row represents the same gene, whereas each column represents an RNA sample obtained from a batch of freshly ovulated eggs. The expression level for each gene is presented using a color intensity scale, where yellow and blue represent over- and under-expression, respectively. Black represents the median gene abundance. Panel ‘b’: Clustering analysis of the 30 most enriched ontological terms (darker nodes are more significantly enriched gene sets; larger nodes represent larger gene sets; thicker edges represent more overlapping genes). Genes in clusters highlighted with different colors were additionally verified to be involved in RNA splicing, chromosome organization and cell junction organization in a separate analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)





**Figure 9. Heat map presenting the expression levels of differentially expressed genes identified during comparative transcriptomic analysis of high- and low-quality pikeperch eggs during zebrafish embryonic and larval development (based on data from White et al. 2017).**

Genes are arranged from top to bottom in descending order of the expression level based on the data provided in the first column (pre-ZGA; period before zygotic genome activation [ZGA]). Blue (up) and red (down) arrowheads provided on the blastula/gastrula stage column indicate up- and downregulation of particular genes following ZGA, respectively. TPM – number of transcripts per million reads, dpf – days post fertilization (2–5 dpf refers to the post hatching period). The first 20 genes (exhibiting the highest expression in the pre-ZGA period in zebrafish), highlighted and labeled in the upper-left corner, were subjected to qPCR validation of their expression level in pikeperch eggs (of various quality). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

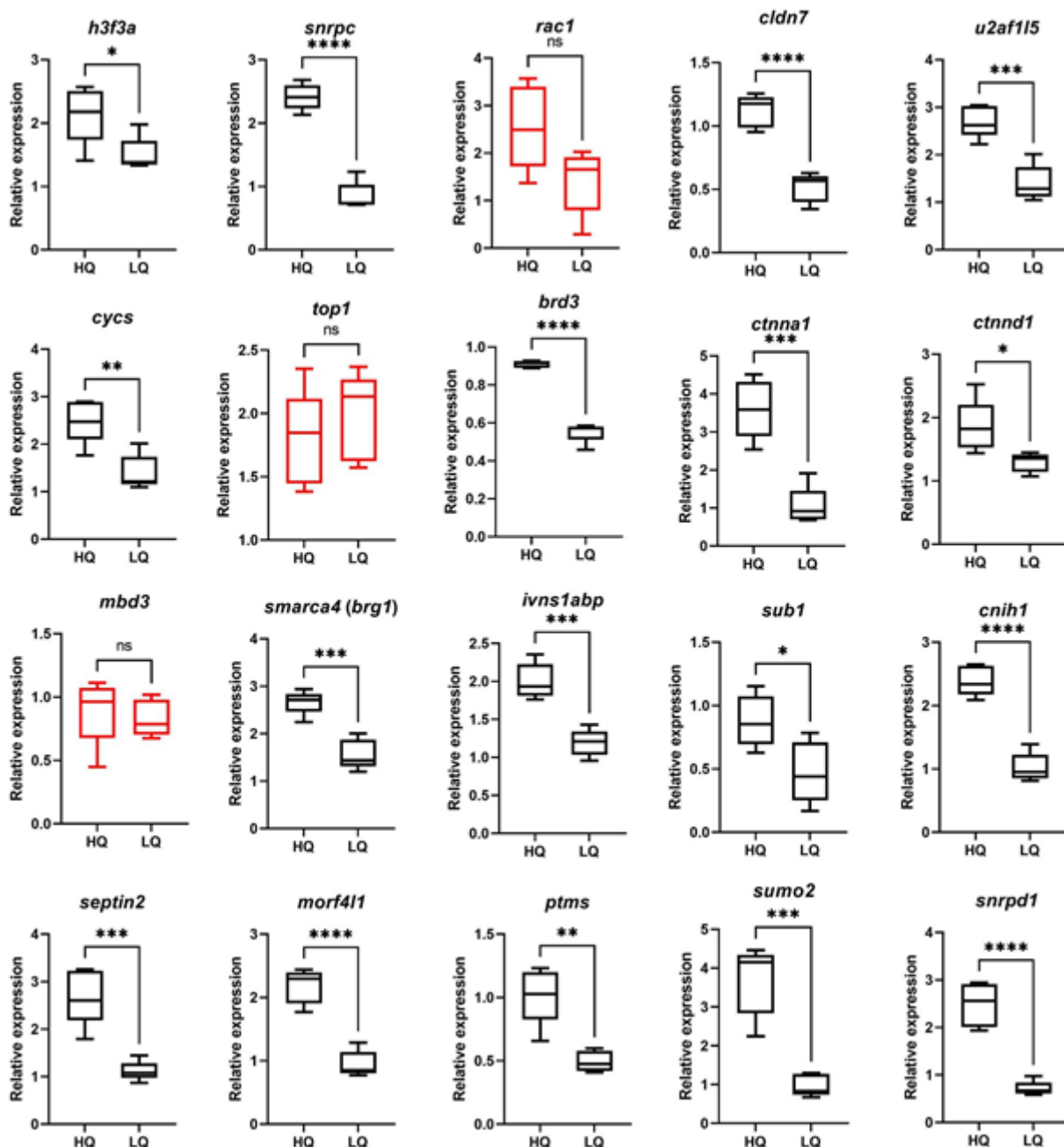


Figure 10. Results of validation of candidate genes (differentially expressed between high and low egg quality in pikeperch; n = 5 for each group) with qPCR. Boxes represent the upper and lower quartiles, with the median value represented by the horizontal bar inside the box. Whiskers represent the minimum and maximum values recorded. The results of statistical analysis are presented as follows: ns – nonsignificant, \* – p < 0.05, \*\* – p < 0.01, \*\*\* – p < 0.001, \*\*\*\* – p < 0.0001.

## 6. References

- [1] C.T. Cheung, T.V. Nguyen, A. le Cam, A. Patinote, L. Journot, C. Reynes, J. Bobe, What makes a bad egg? Egg transcriptome reveals dysregulation of translational machinery and novel fertility genes important for fertilization, *BMC Genomics* 20 (2019) 1–14, <https://doi.org/10.1186/s12864-019-5930-8>.
- [2] W. Tadros, H.D. Lipshitz, The maternal-to-zygotic transition: a play in two acts, *Development* 136 (2009) 3033–3042, <https://doi.org/10.1242/dev.033183>.
- [3] K. Zhang, G.W. Smith, Maternal control of early embryogenesis in mammals, *Reprod. Fertil. Dev.* 27 (2015) 880–896, <https://doi.org/10.1071/RD14441>.
- [4] F. Ren, Q. Lin, G. Gong, X. Du, H. Dan, W. Qin, R. Miao, Y. Xiong, R. Xiao, X. Li, J.-F. Gui, J. Mei, Igf2bp3 maintains maternal RNA stability and ensures early embryo development in zebrafish, *Commun. Biol.* 3 (2020) 94, <https://doi.org/10.1038/s42003-020-0827-2>.
- [5] Z. Alizadeh, S.-I. Kageyama, F. Aoki, Degradation of maternal mRNA in mouse embryos: selective degradation of specific mRNAs after fertilization, *Mol. Reprod. Dev.* 72 (2005) 281–290, <https://doi.org/10.1002/mrd.20340>.
- [6] D.S. Wagner, R. Dosch, K.A. Mintzer, G. Runke, A.P. Wiemelt, M.C. Mullins, Maternal control of vertebrate development before the Midblastula transition, *Dev. Cell* 6 (2004) 771–780, <https://doi.org/10.1016/j.devcel.2004.05.002>.
- [7] I. Adrian-Kalchauer, J.-C.C. Waiser, M. Schwaiger, P. Burkhardt-Holm, RNA sequencing of early round goby embryos reveals that maternal experiences can shape the maternal RNA contribution in a wild vertebrate, *BMC Evol. Biol.* 18 (2018) 34, <https://doi.org/10.1186/s12862-018-1132-2>.
- [8] S. Israel, M. Ernst, O.E. Psathaki, H.C.A. Drexler, E. Casser, Y. Suzuki, W. Makalowski, M. Boiani, G. Fuellen, L. Taher, An integrated genome-wide multi-omics analysis of gene expression dynamics in the preimplantation mouse embryo, *Sci. Rep.* 9 (2019) 1–15, <https://doi.org/10.1038/s41598-019-49817-3>.
- [9] A.H. Smits, R.G.H. Lindeboom, M. Perino, S.J. van Heeringen, G.J.C. Veenstra, M. Vermeulen, Global absolute quantification reveals tight regulation of protein expression in single Xenopus eggs, *Nucleic Acids Res.* 42 (2014) 9880–9891, <https://doi.org/10.1093/nar/gku661>.
- [10] A. Alli Shaik, S. Wee, R.H.X. Li, Z. Li, T.J. Carney, S. Mathavan, J. Gunaratne, Functional mapping of the zebrafish early embryo proteome and transcriptome, *J. Proteome Res.* 13 (2014) 5536–5550, <https://doi.org/10.1021/pr5005136>.
- [11] E. Lubzens, G. Young, J. Bobe, J. Cerdá, A. Oogenesis in teleosts: how fish eggs are formed, *Gen. Comp. Endocrinol.* 165 (2010) 367–389, <https://doi.org/10.1016/j.ygcen.2009.05.022>.
- [12] B.J. Reading, L.K. Andersen, Y.W. Ryu, Y. Mushiobira, T. Todo, N. Hiramatsu, Oogenesis and egg quality in finfish: yolk formation and other factors influencing female fertility, *Fishes* 3 (2018) 1–28, <https://doi.org/10.3390/fishes3040045>.
- [13] Y. Nagahama, M. Yamashita, Regulation of oocyte maturation in fish, *Develop. Growth Differ.* 50 (2008) S195–S219, <https://doi.org/10.1111/j.1440-169X.2008.01019.x>.
- [14] D. Zarski, D. Kucharczyk, K. Targońska, K. Palińska, K. Kupren, P. Fontaine, P. Kestemont, A new classification of pre-ovulatory oocyte maturation stages in pikeperch, *Sander lucioperca* (L.), and its application during artificial reproduction, *Aquac. Res.* 43 (2012) 713–721, <https://doi.org/10.1111/j.1365-2109.2011.02879.x>.
- [15] D. Zarski, S. Krejszef, A. Horváth, Z. Bokor, K. Palińska, K. Szentés, J. Łuczynska, K. Targońska, K. Kupren, B. Urbányi, D. Kucharczyk, Dynamics of composition and morphology in oocytes of Eurasian perch, *Perca fluviatilis* L., during induced spawning, *Aquaculture* 364–365 (2012) 103–110, <https://doi.org/10.1016/j.aquaculture.2012.07.030>.
- [16] L. Kleppe, R.B. Edvardsen, T. Furmanek, G.L. Taranger, A. Wargelius, Global transcriptome analysis identifies regulated transcripts and pathways activated during oogenesis and early embryogenesis in Atlantic cod, *Mol. Reprod. Dev.* 81 (2014) 619–635, <https://doi.org/10.1002/mrd.22328>.
- [17] H. Ma, K. Martin, D. Dixon, A.G. Hernandez, G.M. Weber, Transcriptome analysis of egg viability in rainbow trout, *Oncorhynchus mykiss*, *BMC Genomics* 20 (2019) 1–15, <https://doi.org/10.1186/s12864-019-5690-5>.
- [18] C.L. Winata, V. Korzh, The translational regulation of maternal mRNAs in time and space, *FEBS Lett.* 592 (2018) 3007–3023, <https://doi.org/10.1002/1873-3468.13183>.
- [19] N. Hiramatsu, T. Todo, C. v. Sullivan, J. Schilling, B.J. Reading, T. Matsubara, Y.-W. Ryu, H. Mizuta, W. Luo, O. Nishimiya, M. Wu, Y. Mushiobira, O. Yilmaz, A. Hara, Ovarian yolk formation in fishes: molecular mechanisms underlying formation of lipid droplets and vitellogenin-derived yolk proteins, *Gen. Comp. Endocrinol.* 221 (2015) 1–7, <https://doi.org/10.1016/j.ygcen.2015.01.025>.
- [20] B.J. Reading, C.V. Sullivan, J. Schilling, Vitellogenesis in fishes, in: *Reference Module in Life Sciences*, Elsevier, 2017, pp. 635–646, <https://doi.org/10.1016/B978-0-12-809633-8.03076-4>.
- [21] D. Zarski, T. Nguyen, A. le Cam, J. Montfort, G. Dutto, M.O.M.O. Vidal, C. Fauvel, J. Bobe, Transcriptomic profiling of egg quality in sea bass (*Dicentrarchus labrax*). Sheds light on genes involved in ubiquitination and translation, *Mar. Biotechnol.* 19 (2017) 102–115, <https://doi.org/10.1007/s10126-017-9732-1>.
- [22] E. Bonnet, A. Fostier, J. Bobe, Microarray-based analysis of fish egg quality after natural or controlled ovulation, *BMC Genomics* 8 (2007) 55, <https://doi.org/10.1186/1471-2164-8-55>.
- [23] R.W. Chapman, B.J. Reading, C. v. Sullivan, Ovary transcriptome profiling via artificial intelligence reveals a transcriptomic fingerprint predicting egg quality in striped bass, *Morone saxatilis*, *PLoS One* 9 (2014) e96818, <https://doi.org/10.1371/journal.pone.0096818>.
- [24] T.T. Bizuayehu, M. Mommens, A.Y.M. Sundaram, A.K.S. Dhanasiri, I. Babiak, Postovulatory maternal transcriptome in Atlantic salmon and its relation to developmental potential of embryos, *BMC Genomics* 20 (2019) 315, <https://doi.org/10.1186/s12864-019-5667-4>.
- [25] O. Yilmaz, A. Patinote, T.V. Nguyen, E. Com, R. Lavigne, C. Pineau, C.v. Sullivan, J. Bobe, Scrambled eggs: proteomic portraits and novel biomarkers of egg quality in zebrafish (*Danio rerio*), *PLoS ONE* 12 (2017) 1–24, <https://doi.org/10.1371/journal.pone.0188084>.
- [26] M.-D.D. Castets, B. Schaefer, F. Silvestre, J.-N.N. Gardeur, M. Dieu, C. Corbier, P. Kestemont, P. Fontaine, Combined analysis of *Perca fluviatilis* reproductive performance and oocyte proteomic profile, *Theriogenology* 78 (2012), <https://doi.org/10.1016/j.theriogenology.2012.02.023>, 432–442, 442.e1–13.
- [27] A. Knoll-Gellida, M. André, T. Gattegno, J. Forgue, A. Admon, P.J. Babin, Molecular phenotype of zebrafish ovarian follicle by serial analysis of gene expression and proteomic profiling, and comparison with the transcriptomes of other animals, *BMC Genomics* 7 (2006) 46, <https://doi.org/10.1186/1471-2164-7-46>.
- [28] S. Brooks, C.R. Tyler, J.P. Sumpter, Egg quality in fish: what makes a good egg? *Rev. Fish Biol. Fish.* 7 (1997) 387–416, <https://doi.org/10.1023/A:1018400130692>.
- [29] J. Bobe, Egg quality in fish: present and future challenges, *Animal Frontiers* 5 (2015) 66–72, <https://doi.org/10.2527/af.2015-0010>.
- [30] J. Bobe, C. Labbe, Egg and sperm quality in fish, *Gen. Comp. Endocrinol.* 165 (2010) 535–548, <https://doi.org/10.1016/j.ygcen.2009.02.011>.
- [31] F. Soares, M.T. Dinis, F. Lahnsteiner, L. Ribeiro, Egg quality determination in teleost fish, in: *Methods in Reproductive Aquaculture*, CRC Press, 2008, pp. 149–180, <https://doi.org/10.1201/9780849380549.ch4>.
- [32] S. Aegerter, B. Jalabert, J. Bobe, Large scale real-time PCR analysis of mRNA abundance in rainbow trout eggs in relationship with egg quality and postovulatory ageing, *Mol. Reprod. Dev.* 72 (2005) 377–385, <https://doi.org/10.1002/mrd.20361>.
- [33] H. Rime, N. Guitton, C. Pineau, E. Bonnet, J. Bobe, B. Jalabert, Post-ovulatory Ageing and egg quality: a proteomic analysis of rainbow trout coelomic fluid, *Reprod. Biol. Endocrinol.* 2 (2004) 26, <https://doi.org/10.1186/1477-7827-2-26>.
- [34] A.M. Samarin, A.M. Samarin, T. Policar, Cellular and molecular changes associated with fish oocyte ageing, *Rev. Aquac.* 11 (2019) 619–630, <https://doi.org/10.1111/raq.12249>.
- [35] T. Policar, F.J.F.J. Schaefer, E. Panana, S. Meyer, S. Teerlinck, D. Toner, D. Zarski, Recent progress in European percid fish culture production technology—tackling bottlenecks, *Aquac. Int.* 27 (2019) 1151–1174, <https://doi.org/10.1007/s10499-019-00433-y>.
- [36] J. Lappalainen, H. Dorner, K. Wysujack, Reproduction biology of pikeperch (*Sander lucioperca* (L.)) - a review, *Ecol. Freshw. Fish* (2003).
- [37] P. Fontaine, N. Wang, B. Hermelink, Broodstock management and control of the reproductive cycle, in: P. Kestemont, K. Dąbrowski, R.C. Summerfelt (Eds.), *Biology and Culture of Percid Fishes*, Springer Netherlands, Dordrecht, 2015, pp. 103–122, [https://doi.org/10.1007/978-94-017-7227-3\\_3](https://doi.org/10.1007/978-94-017-7227-3_3).
- [38] D. Zarski, P. Fontaine, J. Roche, M. Alix, M. Blecha, C. Broquard, J. Król, S. Milla Time of response to hormonal treatment but not the type of a spawning agent affects the reproductive effectiveness in domesticated pikeperch, *Sander lucioperca*, *Aquaculture* 503 (2019) 527–536, <https://doi.org/10.1016/j.aquaculture.2019.01.042>.
- [39] D. Zarski, I. ben Ammar, G. Bernáth, S. Baekelandt, Z. Bokor, K. Palińska-Zarska, P. Fontaine, A. Horváth, P. Kestemont, S.N.M. Mandiki, Repeated hormonal induction of spermiation affects the stress but not the immune response in pikeperch (*Sander lucioperca*), *Fish Shellfish Immunol.* 101 (2020) 143–151, <https://doi.org/10.1016/j.fsi.2020.03.057>.
- [40] J. Kristan, A. Stara, M. Polgesek, A. Drasovean, J. Kolarova, J. Priborsky, M. Blecha, P. Svacina, T. Policar, J. Velisek, Efficacy of different anaesthetics for pikeperch (*Sander lucioperca* L.) in relation to water temperature, in: *Neuroendocrinology Letters*, Maghira and Maas Publications, 2014, pp. 81–85. <http://www.scopus.com/inward/record.url?eid=s2-2-0-84928988842&partnerID=tZOTx3y1>.
- [41] D. Zarski, K. Palińska, K. Targońska, Z. Bokor, L. Kotrik, S. Krejszef, K. Kupren, A. Horváth, B. Urbányi, D. Kucharczyk, Oocyte quality indicators in Eurasian perch, *Perca fluviatilis* L., during reproduction under controlled conditions, *Aquaculture* 313 (2011) 84–91, <https://doi.org/10.1016/j.aquaculture.2011.01.032>.
- [42] D. Zarski, S. Krejszef, K. Palińska, K. Targońska, K. Kupren, P. Fontaine, P. Kestemont, D. Kucharczyk, Cortical reaction as an egg quality indicator in artificial reproduction of pikeperch, *Sander lucioperca*, *Reprod. Fertil. Dev.* 24 (2012) 843, <https://doi.org/10.1071/RD11264>.
- [43] J. Roche, D. Zarski, A. Khendek, I. ben Ammar, C. Broquard, A. Depp, Y. Ledoré, T. Policar, P. Fontaine, S. Milla, D1, but not D2, dopamine receptor regulates steroid levels during the final stages of pikeperch gametogenesis, *Animal* 12 (2018) 2587–2597, <https://doi.org/10.1017/S1751731118000824>.
- [44] B.I. Cejko, R.K. Kowalski, D. Kucharczyk, K. Targońska, S. Krejszef, D. Zarski, J. Glogowski, Influence of the length of time after hormonal stimulation on selected parameters of milk of ide *Leuciscus idus* L., *Aquac. Res.* 41 (2010) 804–813, <https://doi.org/10.1111/j.1365-2109.2009.02357.x>.
- [45] F.J.J. Schaefer, J.L.L. Overton, J. Bossuyt, D. Zarski, W. Kloas, S. Wuertz, Management of pikeperch *Sander lucioperca* (Linnaeus, 1758) sperm quality after stripping, *J. Appl. Ichthyol.* 32 (2016) 1099–1106, <https://doi.org/10.1111/jai.13144>.
- [46] D. Kucharczyk, A. Szczerbowski, M.J. Luczynski, R. Kujawa, A. Mamcarz, E. Wyszomirska, T. Szabó, S. Ratajski, Artificial spawning of Eurasian perch, *Perca fluviatilis* L. using Ovopel, *Arch. Pol. Fish.* 9 (2001) 39–49.
- [47] M. Alix, D. Zarski, D. Chardard, P. Fontaine, B. Schaefer, D. Zarski, D. Chardard, P. Fontaine, B. Schaefer, Deformities in newly hatched embryos of Eurasian perch populations originating from two different rearing systems, *J. Zool.* 302 (2017) 126–137, <https://doi.org/10.1111/jzo.12447>.
- [48] D. Zarski, A. le Cam, J. Nynca, C. Klopp, S. Ciesielski, B. Sarosiek, J. Montfort, J. Król, P. Fontaine, A. Ciereszko, J. Bobe, Domestication modulates the expression of genes involved in neurogenesis in high-quality eggs of *Sander lucioperca*, *Mol. Reprod. Dev.* 87 (2020) 934–951, <https://doi.org/10.1002/mrd.23414>.
- [49] J. Nynca, D. Zarski, J. Bobe, A. Ciereszko, Domestication is associated with differential expression of pikeperch egg proteins involved in metabolism, immune response and protein folding, *Animal* 14 (2020) 2336–2350, <https://doi.org/10.1017/S1751731120001184>.
- [50] R. Edgar, M. Domrachev, A.E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, *Nucleic Acids Res.* 30 (2002) 207–210, <https://doi.org/10.1093/nar/30.1.207>.
- [51] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254, [https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3).
- [52] J. Cox, M.Y. Hein, C.A. Luber, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ, *Mol. Cell. Proteomics* 13 (2014) 2513–2526,



- <https://doi.org/10.1074/mcp.M113.031591>.
- [53] Y. Li, Y. Sun, A. Ni, L. Shi, P. Wang, A.M. Isa, P. Ge, L. Jiang, J. Fan, H. Ma, G. Yang, J. Chen, Seminal plasma proteome as an indicator of sperm dysfunction and low sperm motility in chickens, *Mol. Cell. Proteomics* 19 (2020) 1035–1046, <https://doi.org/10.1074/mcp.RA120.002017>.
- [54] J.A. Vizcaino, E.W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J. A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P.-A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R.J. Chalkley, H.-J. Kraus, J.P. Albar, S. Martinez-Bartolomé, R. Apweiler, G.S. Omenn, L. Martens, A.R. Jones, H. Hermjakob, ProteomeXchange provides globally coordinated proteomics data submission and dissemination, *Nat. Biotechnol.* 32 (2014) 223–226, <https://doi.org/10.1038/nbt.2839>.
- [55] S.X. Ge, D. Jung, R. Yao, ShinyGO: a graphical gene-set enrichment tool for animals and plants, *Bioinformatics*. 36 (2020) 2628–2629, <https://doi.org/10.1093/bioinformatics/btz931>.
- [56] P. Bardou, J. Mariette, F. Escudié, C. Djemiel, C. Klopp, Jvenn: an interactive Venn diagram viewer, *BMC Bioinformatics* 15 (2014) 293, <https://doi.org/10.1186/1471-2105-15-293>.
- [57] R.J. White, J.E. Collins, I.M. Sealy, N. Wali, C.M. Dooley, Z. Digby, D.L. Stemple, D. N. Murphy, K. Billis, T. Hourlier, A. Füllgrabe, M.P. Davis, A.J. Enright, E. M. Busch-Nentwich, A high-resolution mRNA expression time course of embryonic development in zebrafish, *ELife*. 6 (2017) 1–32, <https://doi.org/10.7554/eLife.30860>.
- [58] A. Untergasser, H. Nijveen, X. Rao, T. Bisseling, R. Geurts, J.A.M. Leunissen, Primer3Plus, an enhanced web interface to Primer3, *Nucleic Acids Res.* 35 (2007) W71–W74, <https://doi.org/10.1093/nar/gkm306>.
- [59] Z. Dwight, R. Palais, C.T. Wittwer, uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application, *Bioinformatics*. 27 (2011) 1019–1020, <https://doi.org/10.1093/bioinformatics/btr065>.
- [60] M. Mommens, J.M. Fernandes, K. Tollefsen, I.A. Johnston, I. Babiak, Profiling of the embryonic Atlantic halibut (*Hippoglossus hippoglossus* L.) transcriptome reveals maternal transcripts as potential markers of embryo quality, *BMC Genomics* 15 (2014) 829, <https://doi.org/10.1186/1471-2164-15-829>.
- [61] V. Colson, M. Cousture, D. Damasceno, C. Valotaire, T. Nguyen, A. le Cam, J. Bobe, Maternal temperature exposure impairs emotional and cognitive responses and triggers dysregulation of neurodevelopment genes in fish, *PeerJ*. 2019 (2019), <https://doi.org/10.7717/peerj.6338>.
- [62] J. Cosson, Fish sperm physiology: structure, factors regulating motility, and motility evaluation, in: *Biological Research in Aquatic Science*, IntechOpen, 2019, <https://doi.org/10.5772/intechopen.85139>.
- [63] E.S. Litscher, P.M. Wassarman, *The Fish Egg's Zona Pellucida*, 1st ed., Elsevier Inc., 2018 <https://doi.org/10.1016/bs.ctdb.2018.01.002>.
- [64] T. Rocha de Almeida, M. Alix, A. le Cam, C. Klopp, J. Montfort, L. Toomey, Y. Ledoré, J. Bobe, D. Chardard, B. Schaerlinger, P. Fontaine, Domestication may affect the maternal mRNA profile in unfertilized eggs, potentially impacting the embryonic development of Eurasian perch (*Perca fluviatilis*), *PLoS One* 14 (2019), e0226878, <https://doi.org/10.1371/journal.pone.0226878>.
- [65] M. Lopes-Ferreira, G.S. Magalhães, J.H. Fernandez, I.D.L.M. Junqueira-De-Azevedo, P. le Ho, C. Lima, R.H. Valente, A.M. Moura-Da-Silva, Structural and biological characterization of Nattectin, a new C-type lectin from the venomous fish *Thalassophryne nattereri*, *Biochimie*. 93 (2011) 971–980, <https://doi.org/10.1016/j.biochi.2011.03.001>.
- [66] A. Wargelius, T. Furmanek, J. Montfort, A. le Cam, L. Kleppe, A. Juanchich, R. B. Edvardsen, G.L. Taranger, J. Bobe, A comparison between egg transcriptomes of cod and salmon reveals species-specific traits in eggs for each species, *Mol. Reprod. Dev.* 82 (2015) 397–404, <https://doi.org/10.1002/mrd.22487>.
- [68] C. Sullivan, R.W. Chapman, B.J. Reading, P.E. Anderson, Transcriptomics of mRNA and egg quality in farmed fish: some recent developments and future directions, *Gen. Comp. Endocrinol.* 221 (2015) 23–30, <https://doi.org/10.1016/j.ygcen.2015.02.012>.
- M. Mommens, J.M. Fernandes, T.T. Bizuayehu, S.L. Bolla, I. a Johnston, I. Babiak, Maternal gene expression in Atlantic halibut (*Hippoglossus hippoglossus* L.) and its relation to egg quality, *BMC Research Notes* 3 (2010) 138, <https://doi.org/10.1186/1756-0500-3-138>.
- [69] H. Izumi, K. Gen, P.M. Lokman, S. Hagihara, M. Horiuchi, T. Tanaka, S. Ijiri, S. Adachi, Maternal transcripts in good and poor quality eggs from Japanese eel, *Anguilla japonica*—their identification by large-scale quantitative analysis, *Mol. Reprod. Dev.* 86 (2019) 1846–1864, <https://doi.org/10.1002/mrd.23273>.
- [70] J.N. Myers, N. Chatakondi, R.A. Dunham, I.A.E. Butts, Genetic architecture of early life history traits for channel catfish, *Ictalurus punctatus* ♀ × blue catfish, *I. furcatus* ♂ hybrid production, *Aquaculture* (2019) 734436, <https://doi.org/10.1016/j.aquaculture.2019.734436>.
- [71] J. Duan, L. Zhu, H. Dong, X. Zheng, Z. Jiang, J. Chen, X.C. Tian, Analysis of mRNA abundance for histone variants, histone- and DNA-modifiers in bovine in vivo and in vitro oocytes and embryos, *Sci. Rep.* 9 (2019) 1217, <https://doi.org/10.1038/s41598-018-38083-4>.
- [72] H.-Y. Wang, L. Zhou, J.-F. Gui, Identification of a putative oocyte-specific small nuclear ribonucleoprotein polypeptide C in gibel carp, *Comp. Biochem. Physiol. B: Genomics* 113 (2021) 3811–3826
- Biochem. Mol. Biol.* 146 (2007) 47–52, <https://doi.org/10.1016/j.cbpb.2006.09.002>.
- [73] D. Gao, T. Xu, X. Qi, W. Ning, S. Ren, Z. Ru, K. Ji, Y. Ma, T. Yu, Y. Li, Z. Cao, Y. Zhang, CLAUDIN7 modulates trophectoderm barrier function to maintain blastocyst development in pigs, *Theriogenology*. 158 (2020) 346–357, <https://doi.org/10.1016/j.theriogenology.2020.09.038>.
- [74] A. Vite, J. Li, G.L. Radice, New functions for alpha-catenins in health and disease: from cancer to heart regeneration, *Cell Tissue Res.* 360 (2015) 773–783, <https://doi.org/10.1007/s00441-015-2123-x>.
- [75] A.P.A. van Montfoort, J.P.M. Geraedts, J.C.M. Dumoulin, A.P.M. Stassen, J.L. H. Evers, T.A.Y. Ayoubi, Differential gene expression in cumulus cells as a prognostic indicator of embryo viability: a microarray analysis, *Mol. Hum. Reprod.* 14 (2008) 157–168, <https://doi.org/10.1093/molehr/gam088>.
- [76] S.J. Bultman, T.C. Gebuhr, H. Pan, P. Svoboda, R.M. Schultz, T. Magnuson, Maternal BRG1 regulates zygotic genome activation in the mouse, *Genes Dev.* 20 (2006) 1744–1754, <https://doi.org/10.1101/gad.1435106>.
- [77] R.M. Schultz, J.J. Eppig, Challenges to making an egg, *Nat. Cell Biol.* 23 (2021) 9–10, <https://doi.org/10.1038/s41556-020-00622-5>.
- [78] A. Rodriguez, S.M. Briley, B.K. Patton, S.K. Tripurani, K. Rajapaksh, C. Coarfa, A. Rajkovic, A. Andrieux, A. Dejean, S.A. Pangas, Loss of the E2 SUMO-conjugating enzyme Ube2i in oocytes during ovarian folliculogenesis causes infertility in mice, *Development* 146 (2019), <https://doi.org/10.1242/dev.176701> dev176701.
- [79] J.-L. Zhu, S.-L. Lin, M. Li, Y.-C. Ouyang, Y. Hou, H. Schatten, Q.-Y. Sun, Septin2 is modified by SUMOylation and required for chromosome congression in mouse oocytes, *Cell Cycle* 9 (2010) 1607–1616, <https://doi.org/10.4161/cc.9.8.11463>.
- [80] I. Adrian-Kalchauer, S.E. Sultan, L.N.S. Shama, H. Spence-Jones, S. Tiso, C. I. Keller Valsecchi, F.J. Weissing, Understanding “non-genetic” inheritance: insights from molecular-evolutionary crosstalk, *Trends Ecol. Evol.* 35 (2020) 1078–1089, <https://doi.org/10.1016/j.tree.2020.08.011>.

## 7. Acknowledgements

The authors are sincerely grateful to ASIALOR fish farm for their support and access to their facilities as well as Jean Baptiste Mullilo, Miroslav Blecha and Maud Alix for their assistance during the reproduction of pikeperch.