



OPEN

# Evolutionarily conserved ovarian fluid proteins are responsible for extending egg viability in salmonid fish

Aurélie Gueho<sup>1</sup>, Daniel Źarski<sup>2</sup>, H el ene Rime<sup>1</sup>, Blandine Gu evel<sup>3,4</sup>, Emmanuelle Com<sup>3,4</sup>, R egis Lavigne<sup>3,4</sup>, Thaovi Nguyen<sup>1</sup>, J er ome Montfort<sup>1</sup>, Charles Pineau<sup>3,4</sup> & Julien Bobe<sup>1</sup>✉

In contrast to most fishes, salmonids exhibit the unique ability to hold their eggs for several days after ovulation without significant loss of viability. During this period, eggs are held in the body cavity in a biological fluid, the coelomic fluid (CF) that is responsible for preserving egg viability. To identify CF proteins responsible for preserving egg viability, a proteomic comparison was performed using 3 salmonid species and 3 non-salmonid species to identify salmonid-specific highly abundant proteins. In parallel, rainbow trout CF fractions were purified and used in a biological test to estimate their egg viability preservation potential. The most biologically active CF fractions were then subjected to mass spectrometry analysis. We identified 50 proteins overabundant in salmonids and present in analytical fractions with high egg viability preservation potential. The identity of these proteins illuminates the biological processes participating in egg viability preservation. Among identified proteins of interest, the ovarian-specific expression and abundance in CF at ovulation of *N*-acetylneuraminic acid synthase a (Nansa) suggest a previously unsuspected role. We show that salmonid CF is a complex biological fluid containing a diversity of proteins related to immunity, calcium binding, lipid metabolism, proteolysis, extracellular matrix and sialic acid metabolic pathway that are collectively responsible for preserving egg viability.

**Keywords** Rainbow trout, Coelomic fluid, Gamete quality, Mass spectrometry, Fertilization

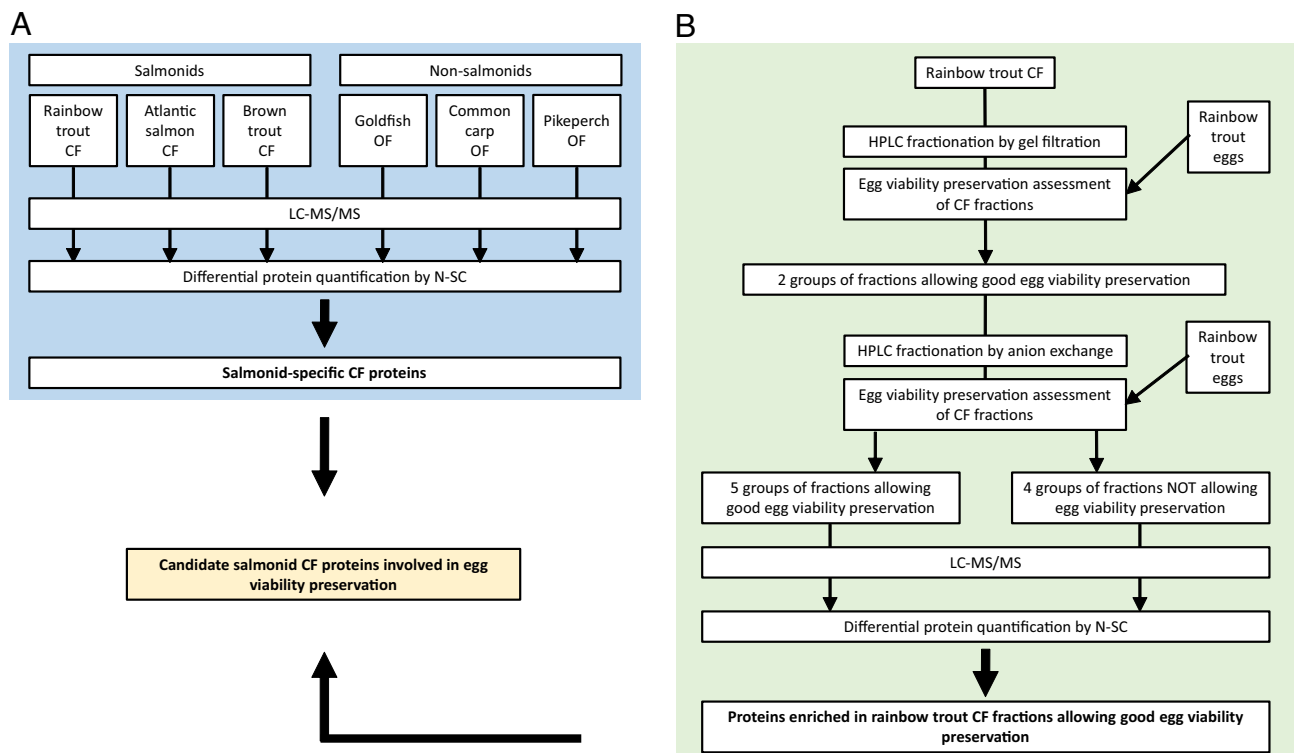
In most teleost fish species, egg viability and quality (i.e., the egg ability to be fertilized and subsequently develop into a normal individual) decreases rapidly after ovulation<sup>1,2</sup>. This is for instance the case in goldfish (*Carassius auratus*), in which a decrease in egg quality is detected as early as 3 h post-ovulation (hpo) at 23 °C, and a complete loss of viability is observed at 12 hpo<sup>3</sup>. In common carp (*Cyprinus carpio*), 50% of the eggs lose their viability within 6 h after ovulation, and the complete loss of viability is reached 12–14 hpo<sup>4</sup>. In sharp contrast with all of other teleost fish species, salmonids have the unique ability to hold their eggs for several days, or even weeks, following ovulation without any significant loss of their viability<sup>5–7</sup>. In contrast to most teleost fish, salmonid fish exhibit “open” ovaries (termed secondary gymnovarians) that are devoid of ovarian lumen. At ovulation, eggs are therefore released directly into the body cavity where they bath in a semi-viscous fluid. This biological fluid is called coelomic fluid (CF), or ovarian fluid (OF) in consistency with other teleost species that exhibit a closed ovary, and because some OF components are of ovarian origin<sup>8–10</sup>. Several studies have shown that salmonid egg viability and ability to be fertilized decreases rapidly after ovulation when eggs are held in vitro in artificial medium rather than in CF<sup>11</sup>. This indicates that CF components play a major role in maintaining egg viability. In addition to its unique ability to maintain egg viability both in vivo and in vitro, CF also exhibits several features associated with fertilization. In fresh water, eggs are water-activated within a minute, or even seconds. The micropyle (i.e., the canal through which the spermatozoa penetrates into the oocyte during fertilization) closes and the chorion hardens, making fertilization impossible<sup>12–14</sup>. As CF is expelled along with the eggs into the water during spawning, its presence around the eggs, even diluted, is believed to locally contribute to maintaining an optimal environment able to prevent or delay immediate egg activation to allow fertilization<sup>15</sup>.

<sup>1</sup>INRAE UR1037, Fish Physiology and Genomics, 35000 Rennes, France. <sup>2</sup>Department of Gamete and Embryo Biology, Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Tuwima Str. 10, 10-748 Olsztyn, Poland. <sup>3</sup>Inserm, EHESP, Irset, UMR\_S 1085, Univ Rennes, 35000 Rennes, France. <sup>4</sup>CNRS, Inserm, Biosit UAR 3480 US\_S 018, Protim Core Facility, Univ Rennes, 35000 Rennes, France. ✉email: Julien.bobe@inrae.fr

CF would also play a chemoattractant role during fertilization to guide the spermatozoa to the micropyle when male and female gametes are simultaneously released into the water<sup>16</sup>. This would be consistent with the role played by ovarian fluid during fertilization in teleost fishes<sup>17,18</sup>. In addition, salmonids coelomic fluid can also be used to maintain egg viability in non-salmonids species like goldfish<sup>19,20</sup> or zebrafish<sup>21</sup>.

The physicochemical properties of salmonid CF have been extensively studied. CF pH is around 8.5 and its osmolarity is about 290 mmol kg<sup>-1</sup>. The ionic composition of CF is similar to blood plasma. It contains sodium (Na<sup>+</sup>), potassium (K<sup>+</sup>), calcium (Ca<sup>2+</sup>), magnesium (Mg<sup>2+</sup>), and chloride (Cl<sup>-</sup>) with differences in the amount of K<sup>+</sup> as CF contains 3–4 times more K<sup>+</sup> than blood plasma. CF also contains glucose, fructose, proteins and free amino acids, cholesterol and phospholipids<sup>22–24</sup>. As indicated above, several lines of evidence strongly suggest that mimicking the ionic composition of the trout CF is not sufficient to fully preserve egg viability<sup>11,25</sup>. This also indicates that the unique biological features of salmonid CF are conferred by organic compounds and most probably by proteins. While the general physicochemical composition of salmonids CF is well known, our knowledge of CF proteome is, in contrast, extremely limited. CF is known to have different enzymatic activities like phosphatase, collagenase, gelatinase and lactate dehydrogenase activities<sup>22</sup>. Several CF proteases have been identified including a progastricsin<sup>10</sup> and a serine protease of the chymotrypsin family<sup>26</sup>. Some protease inhibitors, named Trout Ovulatory Proteins (TOP) were also identified in brook trout (*Salvelinus fontinalis*) and rainbow trout (*Oncorhynchus mykiss*) CF<sup>8,9</sup>. The TOP proteins have the ability to confer anti-bacterial activity to the CF. Using 2D proteomic analysis, potential markers of egg quality, such as vitellogenin fragments and lipoproteins that may originate from leaking eggs and be the consequence of oocyte postovulatory ageing in CF have been reported<sup>27</sup>. The presence of components from broken eggs was found to decrease the pH of trout CF<sup>28</sup> indicating that the biochemical composition of salmonids CF can indirectly reflect the quality of the eggs in the case of post-ovulatory ageing. More recently, shotgun proteomic analysis of rainbow trout CF<sup>29</sup> and of chinook salmon CF<sup>30</sup> led to the identification of 54 and 174 proteins, respectively. Despite this, our current knowledge of salmonid CF proteome is still partial and proteins responsible for its biological properties remain unknown.

The aim of this study was to comprehensively and functionally characterize proteins of rainbow trout CF and subsequently identify the proteins responsible for the extension of egg viability in salmonid CF. For this, the proteomic composition of different salmonid CF and non-salmonid OF was compared to identify proteins specifically present in salmonids CF (Fig. 1A). We reasoned that salmonid-specific proteins could be responsible, at least in part, for preserving egg viability, a feature found only in salmonids. In parallel, rainbow trout CF was fractionated using HPLC coupled to gel filtration or ion exchange columns in order to subsequently identify protein fractions with strong egg viability preservation potential (Fig. 1B). Proteins present in the fractions allowing the highest egg viability preservation were identified using mass spectrometry. Identifying the CF components involved in preserving egg viability would have important applied outputs with various biotechnological applications for aquaculture. It could also help understanding egg quality preservation mechanisms, and subsequently improving egg storage condition in salmonid and non-salmonid fish species.



**Figure 1.** Experimental workflow to identify salmonid-specific CF proteins (A) and rainbow trout CF proteins involved in egg viability preservation (B).

## Results

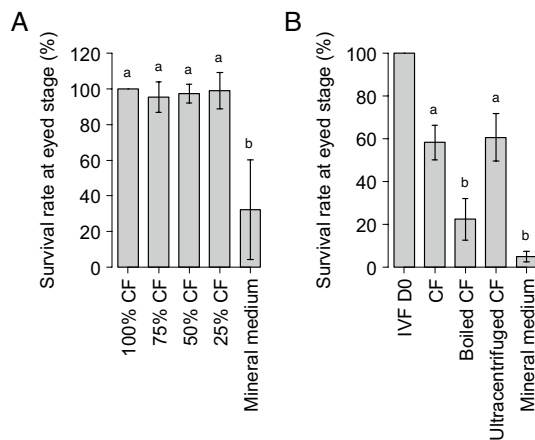
### Proteins are responsible for the biological activity of rainbow trout CF

Rainbow trout eggs were stored in different dilutions of rainbow trout CF (100%, 75%, 50%, and 25%) diluted in mineral medium (MM) mimicking the CF mineral composition<sup>31</sup> or in MM only. In vitro fertilization was carried after 3 days at 12 °C to assess the impact of storage on egg viability and ability to be fertilized. No significant differences were observed in the percentage of eggs reaching eyed-stage between eggs stored in the different CF dilutions and eggs stored in the non-diluted CF. The survival rate (i.e., embryos reaching eyed stage) was however dramatically reduced when eggs were stored in mineral medium alone (Fig. 2A). These observations indicate that organic (i.e., non-mineral) components of the CF are responsible for maintaining egg viability and ability to be fertilized for several days after ovulation. Our observations also suggest a strong biological activity of at least some of the non-mineral CF components as a 25% CF dilution was sufficient to maintain an egg viability similar to that observed using non-diluted CF (Fig. 2A).

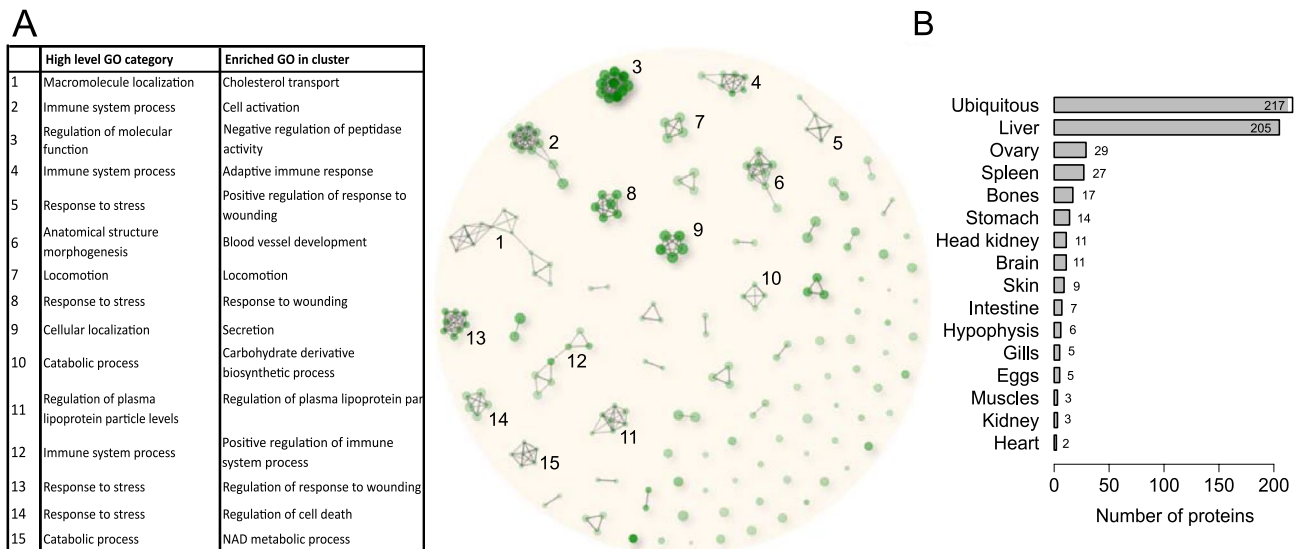
To identify the nature of the organic compound conferring its biological activity to the trout CF, eggs were stored for 3 days at 12 °C in CF (i.e., non-diluted), boiled CF, ultracentrifuged CF, or MM before fertilization. CF was boiled to denature the proteins while ultracentrifugation was used to remove the lipids in order to test the respective contribution of proteins and lipids. No significant difference was observed in the percentage of eggs reaching eyed stage between eggs stored in CF and eggs stored in ultracentrifuged CF. However, when CF proteins had been denatured before holding eggs, this percentage was significantly decreased (Fig. 2B). This suggests that proteins, but not lipids, are responsible for the biological activity associated with CF ability to maintain egg viability and ability to be fertilized.

### Proteomic characterization of rainbow trout CF

In order to understand how rainbow trout CF allows egg conservation, its proteomic composition was characterized by mass spectrometry analysis. Together, 573 proteins were identified (Table S1). A large majority of the proteins were newly identified in the trout CF compared to previous proteomic studies<sup>29,30</sup> allowing the characterization of a more complete trout CF proteome. Human orthologs could be found for 482 of those proteins, corresponding to 326 unique human genes. Those 326 unique genes were then used for biological processes GO-term annotation and the 500 most enriched GO-terms found were clustered into several processes (Fig. 3A, Table S2). Fifteen main clusters connecting at least 4 different GO-terms sharing 70% of their genes were identified. For each of them, the corresponding high-level GO categories as well as the most represented GO-term are indicated. Four of those clusters are related to “response to stress”, and more precisely to “response to wounding”. Three other clusters are related to “immune system processes”. Among the remaining ones are found biological processes involved in “blood vessel development”, “regulation of plasma lipoprotein particle levels”, “cholesterol transport”, “locomotion”, “secretion”, “negative regulation of peptidase activity” or more general catabolic processes like “carbohydrate derivative catabolic processes” and “NAD metabolic process”.



**Figure 2.** Effect of a 3-day in vitro storage of unfertilized rainbow trout eggs in different dilutions of CF on embryonic survival at eyed stage (A). Results were normalized to the percentage of embryos reaching eyed stage after incubation of unfertilized eggs in the non-diluted CF control condition (100% CF). Five independent experiments were performed using eggs collected from 5 different rainbow trout females. Percentages of embryos reaching eyed stage (mean  $\pm$  sd) superscripted with the same letter are not significantly different, Wilcoxon–Mann Whitney test ( $p$  value  $< 0.05$ ). Effect on embryonic survival of a 3-day in vitro storage of unfertilized rainbow trout eggs in coelomic fluid (CF), boiled coelomic fluid (Boiled CF), ultracentrifuged coelomic fluid (Ultracentrifuged CF) or trout mineral medium (MM) (B). Results were normalized to the percentage of eggs reaching eyed stage when the fertilization was performed immediately upon eggs collection (IVF D0). Percentages of embryos reaching eyed stage (mean  $\pm$  sem) superscripted with the same letter are not significantly different, Wilcoxon–Mann Whitney ( $p$  value  $< 0.05$ ). Seven independent experiments were performed using eggs collected from 7 different rainbow trout females.



**Figure 3.** Proteomic characterization of rainbow trout coelomic fluid. Hierarchical clustering of the 500 most enriched biological processes GO-terms (FDR < 0.05) in the rainbow trout CF proteome (A). Related GO-terms in a cluster share at least 70% of their proteins. Proteins of the main clusters (clusters with at least 4 GO-terms) were further subjected to GO enrichment analysis to identify high-level GO categories representing the different clusters, as well as the most represented biological processes in each cluster. Main tissues of expression of the rainbow trout CF proteins (B).

The amino acid sequences of all the proteins identified in the trout CF were blasted against the PhyloFish transcriptomic database<sup>32</sup>. For each protein, the main tissue of expression of the corresponding contig was manually checked (Fig. 3B, Table S3). Proteins were classified based on the tissue/organ exhibited predominant expression. Proteins exhibiting a ubiquitous expression were the most abundant. In addition, a high proportion of proteins exhibited a predominant expression in the liver. This observation is consistent with the previously reported presence of egg yolk proteins in rainbow trout fluid<sup>27</sup>. We also identified 29 proteins that appear to be predominantly expressed in the ovary. This is consistent with the ovarian origin of several abundant CF proteins previously described in salmonids, including progastresin<sup>10</sup> and trout ovulatory proteins<sup>8,9</sup>.

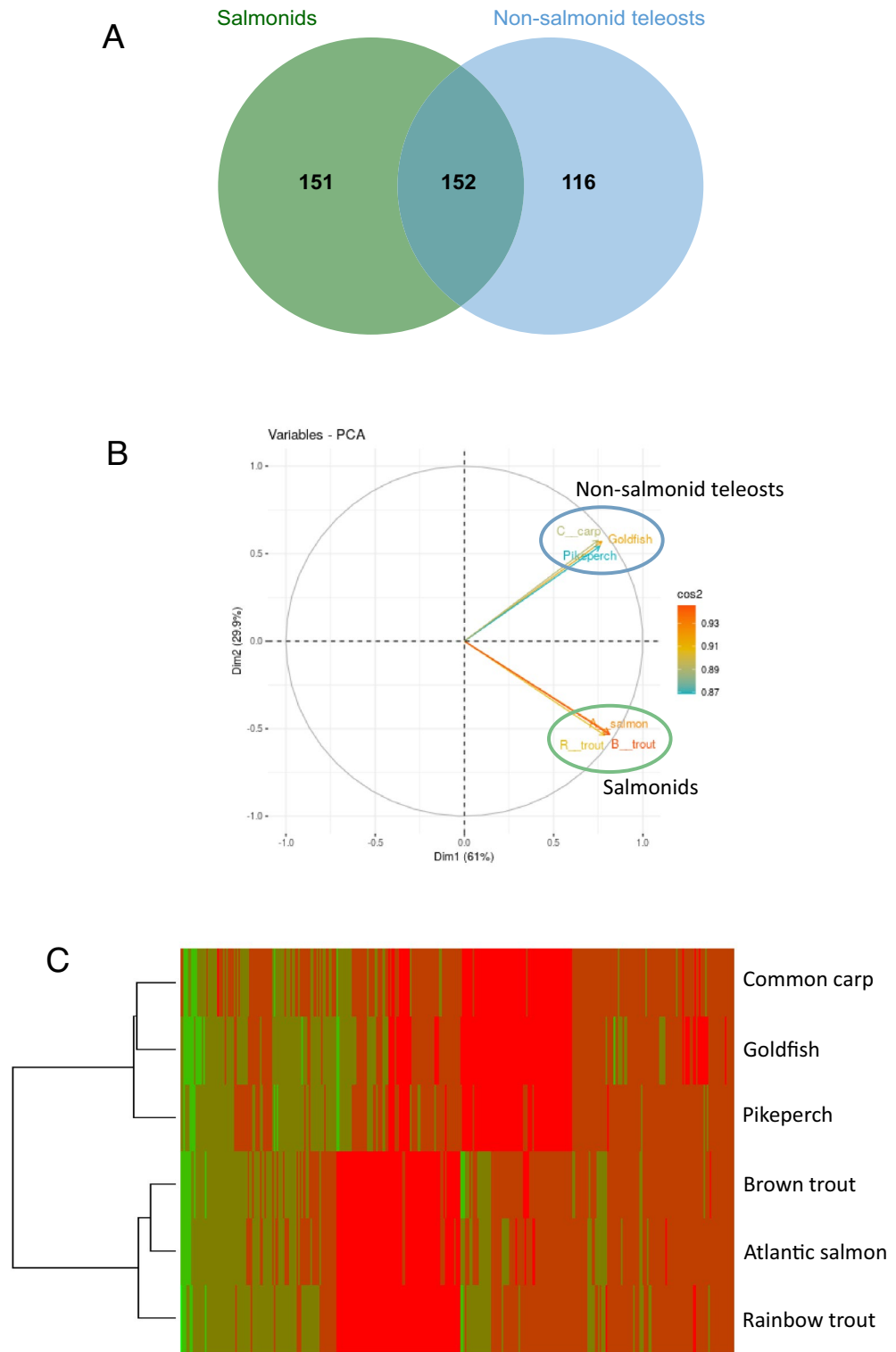
### Identification of salmonid specific CF proteins

Extended egg viability following ovulation is a salmonid-specific trait. In order to identify proteins conferring this specific egg conservation property to the salmonids CF, a proteomic comparison of three different salmonid species CF (rainbow trout, *Oncorhynchus mykiss*; Atlantic salmon, *Salmo salar* and brown trout, *Salmo trutta*) to three non-salmonid teleost species OF (goldfish, *Carassius auratus*; common carp, *Cyprinus carpio*; and pikeperch, *Sander lucioperca*) was performed (Figs. 1A and 4). For each species, only proteins identified in all biological replicates, when available, were kept for further analysis to constitute the representative CF or OF proteome of the species. Together, 511 proteins were identified in Atlantic salmon CF, 388 in brown trout CF, 876 in goldfish OF, 423 in common carp OF and 392 in pikeperch OF (Table S1). To facilitate inter-species comparisons, all identified proteins were mapped to teleost common ancestors using the Genomicus<sup>33</sup> database in order to identify groups of orthologous proteins. A total of 368 unique common ancestors were found in rainbow trout CF, 311 in Atlantic salmon CF, 299 in brown trout CF, 527 in goldfish OF, 420 in common carp OF, and 295 in pikeperch OF (Table S1).

Ancestral proteome comparison was performed using common ancestors shared by different species. For comparison purposes, only common ancestors shared by at least two species of each group were kept for further analysis. This led to the identification of 303 and 268 common ancestors in the salmonid group and non-salmonid group, respectively. Each group shared roughly half of its identified common ancestors with the other group (Fig. 4A). The PCA of the quantified common ancestors confirmed the discrimination of the representative CF and OF proteomes of the different species into two distinct groups: salmonids and non-salmonid teleosts (Fig. 4B). This was further supported by the non-supervised hierarchical clustering of the quantified common ancestors (Fig. 4C) that clearly grouped salmonid and non-salmonid-species together. From the 419 common ancestors quantified, 78 were found significantly enriched in the salmonid group in comparison to the non-salmonid group (Fig. 4C, Table S4). In rainbow trout CF, 97 proteins corresponding to 66 common ancestors were found. Proteins corresponding to the 12 remaining enriched common ancestors were only found in the CF of the two other salmonid species (Table S4).

### Identification of rainbow trout candidate CF proteins involved in egg viability preservation

The proteomic comparison of salmonid CF to non-salmonid OF led to the identification of proteins specifically present in salmonids CF. This set of 78 common ancestor genes could be playing an important role in maintaining



**Figure 4.** Reference proteome comparison of salmonids CF and non-salmonid teleost OF. Venn diagram showing the common ancestor shared by salmonids, non-salmonids, or both groups. Only common ancestors identified in at least 2 of the 3 species of a group we used (A). Principal component analysis (PCA) of mean normalized spectral counts of the different species (B). Heat map of non-supervised hierarchical clustering of mean normalized spectral counts corresponding to the common ancestors of the different species. Red indicates high abundance while green indicates low abundance (C).

egg viability and ability to be fertilized after ovulation. In addition to this evolutionarily-supported analysis, we used a functional angle. We speculated that different protein fractions of the salmonid CF could play different role in preserving egg viability and ability to be fertilized (Fig. 4). In order to further identify proteins involved in maintaining egg viability, rainbow trout CF was analytically fractionated (Fig. 1B). Using a gel filtration column, CF proteins were separated in 30 different fractions according to their size. Then, the ability of the different fractions to preserve trout egg quality after 3 days of storage was evaluated. Storage in non-fractionated CF was used as a positive control.

Strong differences in embryonic survival rate could be observed between the fractions (Fig. 5A), confirming our working hypothesis that different CF fractions had different contribution to CF biological activity. Three main peaks of egg survival rate were observed in fractions F3, F12 and F28 with 42%, 61% and 61.5% of eggs reaching eyed-stage, respectively. These percentages were not significantly different from the one obtained with the non-fractionated CF. Proteins present in those fractions were subsequently analyzed by mass spectrometry. To tentatively identify candidate proteins involved in egg preservation, the two groups of fractions with the highest embryonic survival rates, F12 (group F11–F13) and F28 (group F27–F29), were selected for further analysis.

For these two fractions, a second fractionation was performed (Fig. 1B). An anion exchange column was used to separate the proteins according to their charge. The egg viability conservation property of the different fractions obtained was evaluated as described above. The results for the group F11–F13 fractionation are presented in Fig. 5B. Three main pics of embryonic survival rate were observed in the fractions F6, F12 and F17 with a high proportion (i.e., above 78%) of eggs reaching eyed-stage. The percentages of these 3 main peaks were not significantly different from the percentage obtained with the non-fractionated CF. The results for the group F27–F29 fractionation are presented in Fig. 5C. Observed survival rates were lower than in the two previous fractionation experiments. However, a pic of embryonic survival rate, reaching 62%, was observed in the fraction F6. This embryonic survival rate was not significantly different from the non-fractionated trout CF. Proteins present in the fractions allowing good embryonic survival rate were identified by MS: F6, F12 and F17 for the group F11–F13, and F6 for the group F27–F29. As a comparison, fractions with embryonic survival rate significantly lower than with non-fractionated CF were also analyzed by MS. Those fractions were selected preferentially “close” to the fractions allowing good egg conservation during the fractionation process. Fractions F11, F14 and F22 were selected for comparison with the F11–13 group while fraction F12 was selected for comparison with the F27–F29 group. Normalized spectral counts (N-SC) ratios of good versus bad egg conservation fractions were then calculated. A total number of 376 proteins were found specifically present or highly enriched (N-SC ratios > 2) in the fractions allowing good egg viability conservation (Table S5).

Interestingly, 50 of those proteins specifically present or enriched in the rainbow trout CF fractions allowing egg viability conservation were also found among the salmonid specific CF proteins (Table S5). They constitute a list of potential candidates conferring the exceptional egg conservation property to salmonid CF.

### Functional analysis of the *N*-acetylneuraminic acid synthase a (*Nansa*) protein

Among the potential candidates identified, some are known to play a role in the ovulation process. We decided to focus on a specific protein not previously reported in salmonids CF, and with no known function in the ovulation process: the sialic acid synthase protein (XP\_021429200.1). According to Ensembl, this protein is encoded by the *N*-acetylneuraminic acid synthase a gene (*nansa*) (ENSOMYG00000017265) that is orthologous to the human *NANS* (ENSG00000095380) gene. Sialic acid synthases are responsible for the formation of sugars on glycoproteins in eukaryotic cells and are involved in fertilization, embryogenesis, and differentiation<sup>34–36</sup>. The identified *Nansa* protein has a NeuB domain and a SAF NeuB-like domain. NeuB is the catalytic domain involved in *N*-acetylneuraminic acid (Neu5Ac) in procaryotes, whereas SAF NeuB-like domain would allow the binding to substrates. SAF NeuB-like domain is part of the superfamily domain SAF which also includes the domains similar to fish antifreeze proteins according to the NCBI conserved domain database<sup>37</sup>.

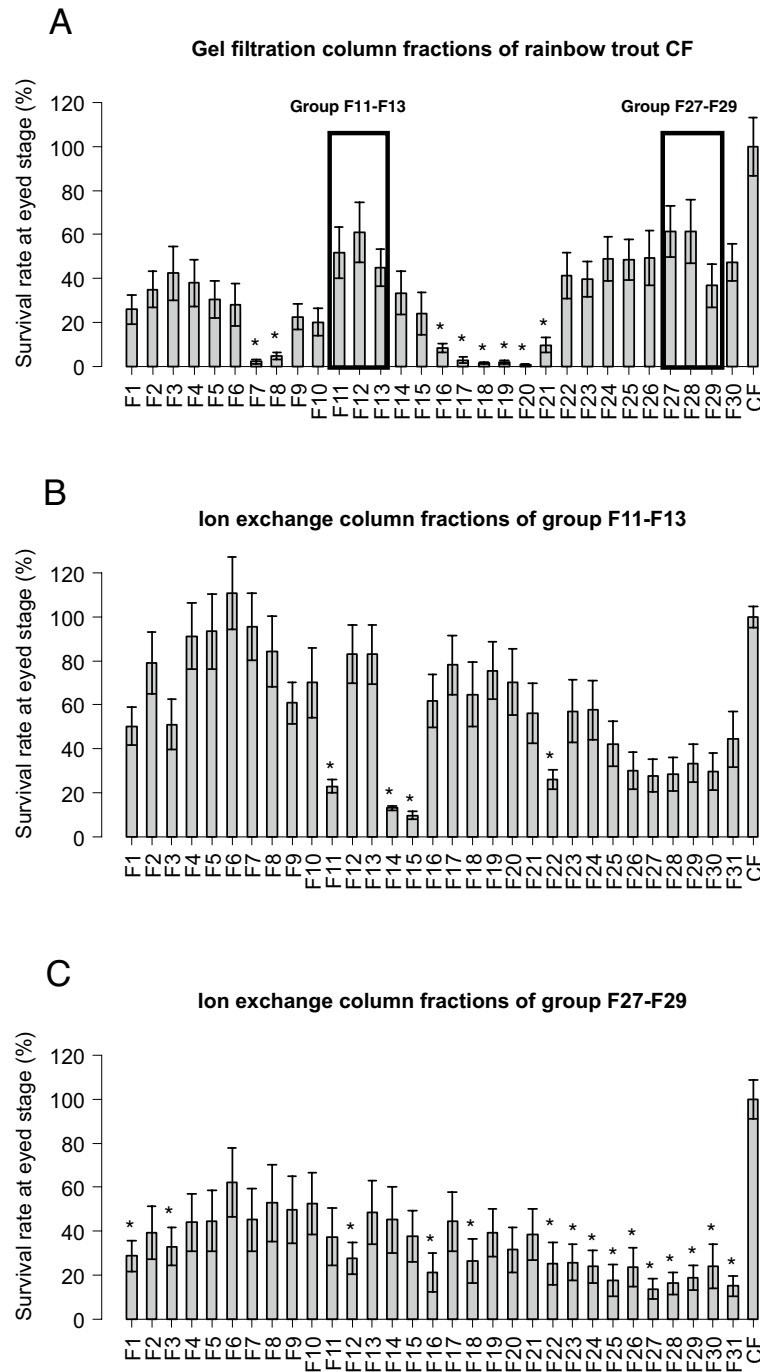
According to the Phylofish database<sup>32</sup>, the contig corresponding to *Nansa* is mainly expressed in the ovary (Table S3). Using quantitative RT-PCR we observed that *nansa* was strongly and predominantly expressed in the ovary, but also in the testis (Fig. 6A). In the ovary, *nansa* is progressively down-regulated (Fig. 6B) following ovulation. *Nansa* protein profile observed in ovary and CF are consistent with gene expression profiles. Our western blot analysis showed that *Nansa* is highly expressed in the ovary and the CF at ovulation. *Nansa* protein abundance then progressively decreased following ovulation (Fig. 6C,D, Fig. S1).

To evaluate the contribution of sialic acid-synthase protein to egg conservation, rainbow trout eggs were stored for 3 days at 12 °C in the dark, in MM complemented with two different concentrations of sialic acid-synthase or in a CF fraction not allowing a good egg conservation complemented with two different concentrations of sialic acid-synthase, before fertilization. As a control, eggs were also incubated in MM or in CF fraction complemented with BSA. Complementation with sialic acid-synthase did not improve embryonic survival rate (Fig. 6E).

## Discussion

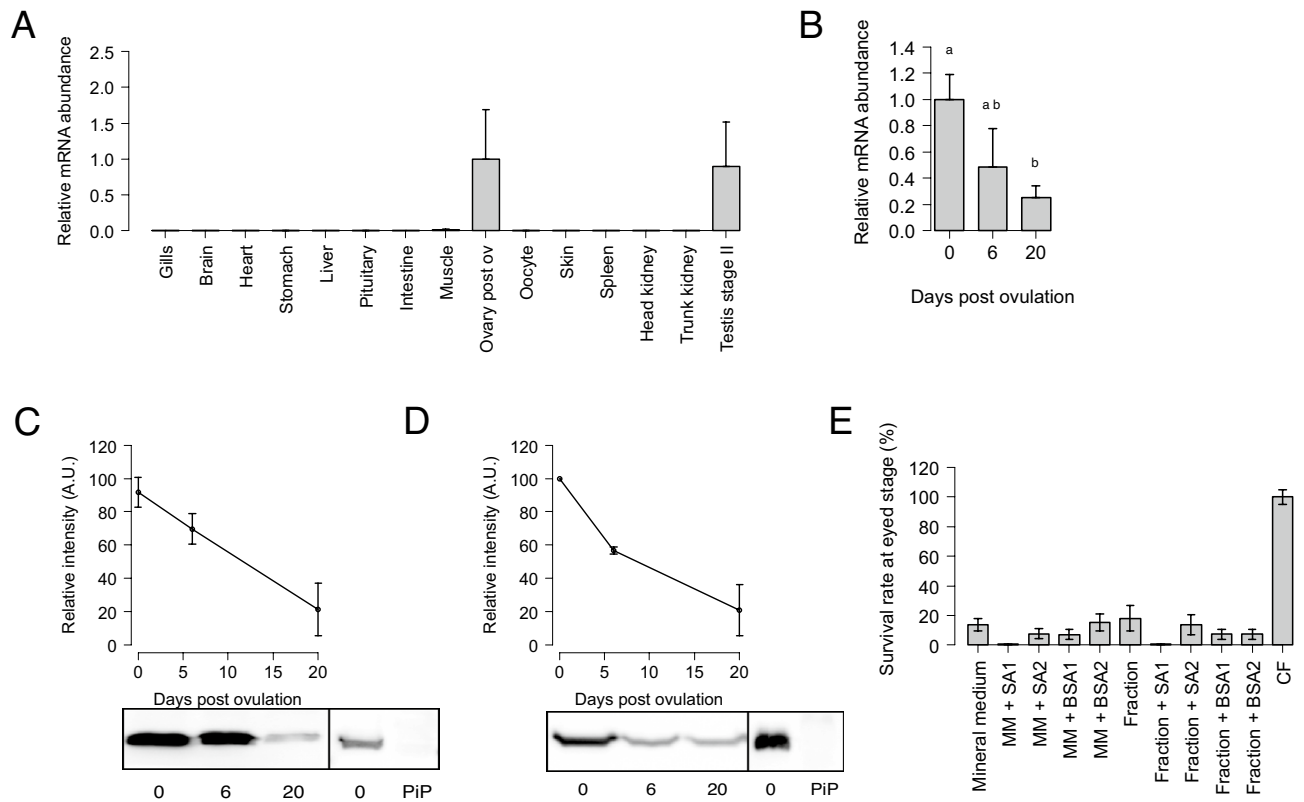
### Proteomic characterization of the rainbow trout CF

In our study, we demonstrate that using mineral medium mimicking the chemical composition of rainbow trout CF for egg storage is not sufficient to preserve egg viability and ability to be fertilized. We showed that organic compounds, and more specifically proteins, confer egg conservation properties to rainbow trout CF (Fig. 2). To better understand this biological property, the rainbow trout CF proteomic composition was analyzed. This led to the identification of 573 proteins (Table S1), which is consistent with data recently obtained for Chinook salmon CF proteome<sup>16</sup>. However, compared to the previously 54 proteins identified in the rainbow trout CF<sup>29</sup>, the present analysis provides a more comprehensive characterization of the rainbow trout CF proteome. Together,



**Figure 5.** Egg preservation potential of the rainbow trout coelomic fluid fractions. CF proteins were separated into 30 different fractions according to their size (A). Proteins present in fractions 11–13 (B) and fractions 27–29 (C) that exhibited the highest egg preservation potential were further separated according to their charge using an anion exchange column. The egg preservation potential of each fraction was estimated through the monitoring of embryonic survival at eyed stage after a 3-day *in vitro* storage in CF fraction. All bars represent the means  $\pm$  sem of 4 independent experiments conducted using the eggs of 4 different females. \*Significantly different from CF, *t*-test (*p* value < 0.05).

these results indicate that CF is a complex biological fluid. It should be noted that CF and OF used in this study could not be processed for proteomic analysis or HPLC fractionation immediately after sampling. For this reason, they had to be frozen and stored until further use. We cannot exclude that the freezing process had an impact on CF and OF composition or on their egg viability conservation property. However, to allow comparison of the results obtained with the different experiments performed, all CF and OF samples were processed similarly and frozen in similar conditions.



**Figure 6.** (A) Gene expression profile of *N-acetylneuraminic acid synthase a (nansa)* (means  $\pm$  sem) in different rainbow trout tissues. Samples were obtained from three different fish. The *nansa* mRNA abundance was determined by qPCR and normalized to the abundance of *18S*. Abundance was set to 1 for the maximal gene expression and data are expressed as a percentage of transcript abundance in the post-ovulatory (Ovary post ov). (B) Ovarian expression profile of *N-acetylneuraminic acid synthase a (nansa)* gene in the ovary of rainbow trout during the post-ovulatory period (means  $\pm$  sem). Ovaries were samples at ovulation, then 6 and 20 days after ovulation. The *nansa* mRNA abundance was determined by qPCR and normalized to the abundance of *18S*. Abundance is pressed as percentage of mRNA levels at ovulation. For each ovary stage, ovaries were sampled from 3 separate females. Different letters indicate significant differences (Welch t-test,  $p < 0.05$ ). (C) Representative cropped western blot analysis of *N-acetylneuraminic acid synthase a (Nansa)* expression in the rainbow trout ovary. Ovarian samples were analyzed at 0, 6, and 20 days after ovulation. For each ovary stage, ovaries were sampled from 3 separate females. Band intensity was normalized by setting the maximal signal obtained to 100%. Upper panel Graphs represent means  $\pm$  sd of 3 independent western blots performed on 3 biological replicates (Suppl Fig. S1A). Pre-immune serum (PiP) was used as a negative control (Lower panel). (D) Representative cropped western blot analysis of *N-acetylneuraminic acid synthase a (Nansa)* expression in the rainbow trout coelomic fluid. CF samples were analyzed at 0, 6, and 20 days after ovulation. CF series were collected from 3 separate females. Band intensity was normalized by setting the maximal signal obtained to 100% (Upper panel). Graphs represent means  $\pm$  sd of 3 independent western blots performed on 3 biological replicates (Suppl Fig. S1B). Pre-immune serum for antibody sialic acid synthase-like was represented (PiP) (Low panel). (E) Percentage of rainbow trout eggs reaching eyed stage after 3 days of in vitro storage in mineral medium (MM), mineral medium complemented with 2 different concentrations of *N-acetylneuraminic acid synthase a (Nansa)* (SA1, 4  $\mu$ g/mL; SA2 40  $\mu$ g/mL), a CF fraction exhibiting a low egg preservation potential (Fraction), the same fraction complemented with 2 different concentrations of *N-acetylneuraminic acid synthase a (Nansa)* (SA1, 4  $\mu$ g/mL; SA2 40  $\mu$ g/mL), and in CF (positive control). BSA was used as a negative control for mineral medium and fraction complementation at the same concentrations as sialic acid synthase-like (BSA1, 4  $\mu$ g/mL; BSA2 40  $\mu$ g/mL). All bars represent the means  $\pm$  sem of 3 independent experiments conducted on the eggs of 3 different females.

The gene ontology analysis of the rainbow trout CF proteome revealed fifteen clusters of biological processes that are highly consistent with previous proteomic analyses of rainbow trout and chinook salmon<sup>29,30</sup> (Fig. 3, Table S2). Four GO clusters were related to “response to stress” and more precisely to “response to wounding”. This group of proteins are involved in injury response, which is consistent with the response to the rupture of follicular somatic layers at ovulation. Three other clusters were related to “immune system processes”. In this group are notably found proteins of the Serpin family, Fibrinogen proteins, proteins of the complement, and of the von Willebrand factor. This group of proteins would be directly involved in egg conservation and more precisely in the egg protection against infections. Indeed, it was previously demonstrated that the trout CF has anti-bacterial properties<sup>9</sup>. However, complement proteins, which are present in this group, are not only restricted to immunity.



They are also frequently associated to ovarian function. Complement component 3 (C3) has been identified as a potential biomarker for in vitro fertilization (IVF) outcome<sup>38</sup> while iC3b, its product of cleavage, is involved in oocyte maturation in pigs<sup>39</sup>. Several complement proteins have been proposed to be oocyte maturation factors in human<sup>40</sup>. All the above-listed proteins are acute phase proteins (APP) that are expressed in response to injuries or infections. In mammals, the ovulation process shares similarities with the inflammation process, and numerous APP are commonly involved in both processes<sup>41,42</sup>. Similarly to mammals, it has been suggested that ovulation in fish would also be an inflammation-like process<sup>43,44</sup>. Together, these observations are consistent with the identification of APPs in the two main represented biological processes in the rainbow trout CF: “response to wounding” and “immune system processes”.

Two GO clusters were related to lipids binding and transport: “cholesterol transport” and “regulation of plasma lipoprotein particles level”. Lipids play important dual functions during oogenesis. Cholesterol is an essential substrate for normal steroidogenesis and consequently for normal follicle development and ovulation. Furthermore, lipids are sequestered from the VLDL plasma and incorporated in the growing oocyte in the form of lipid granule where they constitute an energy reserve for the future developing embryo. As in previous trout and salmon CF proteomic studies, the most abundant proteins that we identified were vitellogenins (Table S1). Those glycolipophosphoproteins are the major proteins of egg yolk. After incorporation in the growing oocyte during vitellogenesis, they are further processed and also constitute a metabolic energy reserve for the future developing embryo.

Among the remaining GO clusters are notably found “blood vessel development” and “regulation of peptidase activity” processes. Blood vessel development is important for the local delivery of proteins involved in the ovulation process or necessary for oocyte growth. In the GO “regulation of peptidase activity” are found proteases and their inhibitors involved in the extra cellular matrix (ECM) remodeling and degradation like Mmp9 and Timp2. This process is necessary to the follicular rupture. In this GO term are also found proteases involved in the coagulation process, in response to follicular rupture.

### Identification of candidate CF proteins involved in egg conservation

More than a comprehensive proteomic characterization of the rainbow trout CF, the main goal of this study was to identify proteins conferring its exceptional egg preservation ability to salmonid CF. We first used an evolutionary approach to identify proteins specifically enriched in salmonid CF. To our knowledge, this is the first study comparing the proteomic composition of CF/OF in different fish species. The main challenge was to find a way to perform inter species comparisons. To address this issue, all identified proteins were mapped to teleost common ancestors using the Genomicus database that provides synteny-supported orthology assignments. Then, species were separated in 2 groups: salmonid species and non-salmonid species. This comparison led to the identification of 78 unique teleost common ancestors significantly enriched in the CF of salmonid species (Fig. 4, Table S4) in comparison to non-salmonid species. We then used a functional approach to identify CF proteins specifically involved in egg conservation using a biological test. This analysis was conducted in rainbow trout. For this approach, rainbow trout CF was fractionated to allow an independent evaluation of the biological activity of the different fraction. This led to the identification of 376 rainbow trout CF proteins specifically enriched in fractions allowing good egg conservation (Fig. 5, Table S5).

After compiling the results from evolutionary and functional analyses, 50 rainbow trout CF proteins were commonly identified by both approaches (Table S5). We speculated that these proteins are major players of egg viability preservation. Based on their functions, they can be separated in different categories that will be discussed below.

#### *Immunity*

This category includes complement system components C1q-like (C1q1), complement factor D (Cfd), complement factor H (Cfh), and precerebellin-like protein precursor (Cbln1). Cbln1 has previously been associated with acute phase response in rainbow trout<sup>45</sup>. Together with previous proteomic analyses<sup>16,29</sup> our data show that proteins involved in immunity are abundant in the CF of salmonid fish. In addition, we provide functional evidence that these players would play important roles in preserving eggs viability when eggs are held in CF.

#### *Calcium-dependent and calcium-binding proteins*

Our results also suggest that several calcium-dependent and calcium-binding CF proteins play important roles in preserving egg viability, including 45 kDa calcium-binding protein, calcium/calmodulin-dependent protein kinase type II subunit delta (Camk2d), calumenin-A (Calu), ependymin<sup>46</sup>, lithostathine-1-like<sup>47</sup>, and B-cadherin. Calcium (Ca<sup>2+</sup>) is an important parameter modulating sperm velocity. Indeed, there is a negative correlation between duration of sperm motility and OF Ca<sup>2+</sup> concentration<sup>23</sup>. Chemical composition, including Ca<sup>2+</sup> concentration, of several salmonids CF has been extensively studied<sup>22,23</sup>. In fact, a high proportion of calcium would not be free but bound to proteins<sup>23</sup>. The calcium-binding protein that we identified in our study might be involved in the regulation of free Ca<sup>2+</sup> concentration in salmonid CF. High levels of calcium binding proteins in CF would presumably help extending sperm motility and increase fertilization success. Consistent with this hypothesis is the abundance of several calcium-binding proteins in Chinook salmon CF<sup>16</sup>.

#### *Sialic acid related proteins*

Several sialic acid related proteins were identified in our analysis, including sialic acid synthase, alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1-like. Sialic acids are 9 carbons sugar structures present as terminal residues of glycoproteins. Specific trout egg glycoproteins are progressively sialylated during oogenesis<sup>48</sup>. Several sialyltransferases are involved in this process, notably alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase

1. It was hypothesized that sialic acid structures could have several functions in egg protection including protection from maternal complement attack<sup>36</sup>. Interestingly, complement factor H, that we also identified in our screen, might also be involved in egg protection from maternal complement. Indeed, factor H down-regulates the activation of the alternative complement pathway. By binding sialic acids expressed at cell surface, it allows self-recognition and prevent cells to be attacked by the complement system<sup>49</sup>. Sialic acid might also be involved in calcium binding. Ependymin, one of our candidates, is a Ca binding protein that is a sialylated. It seems that sialic acid residues are responsible for this Ca binding property<sup>46</sup>.

#### *Lipid transport and catabolism*

Our screen also led to the identification of several proteins involved in lipid transport and catabolism/hydrolysis including apolipoprotein B-100, beta-2-glycoprotein 1, carboxylesterase 3, phospholipase B1, and prosaposin. As previously discussed, lipids play important function during the oogenesis. Apo B-100 is involved in lipid transport, and notably cholesterol, the substrate for steroid hormones synthesis by the ovarian follicle<sup>44</sup>. It may also be involved in the storage of neutral lipid in the form of lipid globules in the oocyte cytoplasm. This step would also require the action of lipid hydrolases<sup>50,51</sup>. The identification of proteins of this category in fish ovarian fluid is therefore not surprising. However, it is not clear why they are specifically enriched in Salmonids CF and in CF fractions allowing a better egg conservation. Their role in egg protection is currently unknown and would require further investigations.

#### *Extracellular matrix (ECM) proteins and ECM degrading enzymes*

Many ECM proteins and ECM degrading enzyme were also identified as possible major players responsible for salmonid CF ability to preserve egg viability including cathepsin K, matrix metalloproteinase-9 (Mmp9), olfactomedin-4, versican b, thrombospondin-1, cytoskeleton proteins (tubulin alpha chain, scinderin like b, plastin-2) and cell adhesion proteins (integrin beta-1, B-cadherin). These results are consistent with the previous report of protease activity and notably collagenase activity in salmonid CF<sup>22</sup>. Mmp9 is a gelatinase expressed in the teleost ovary<sup>52</sup> and involved in the ovulation process<sup>53</sup>. CtsK is a collagenase involved in bone resorption<sup>54</sup>. CtsK has been detected in the mouse ovary but its involvement in ECM degradation in the ovary has not been demonstrated<sup>55</sup>. At the time of ovulation, proteases play a role in the lysis of the follicular epithelium to allow the release of mature oocytes into the coelomic cavity. Proteolytic enzymes and components of the extracellular matrix might also be released in the coelomic fluid during this process. Whereas their abundance in salmonids CF can be explained, their role in the egg conservation process remains unclear. However, some of those proteins are associated to oocyte quality or fertilization process in vertebrates. In humans, MMP-9 levels in follicular fluid are positively correlated with oocyte quality and fertilization rate<sup>56</sup>. Furthermore, spermatozoa capacitation appears to require interaction with ECM proteins<sup>57</sup>. In summary, while their precise contribution is unclear, it cannot be excluded that some of the identified ECM candidate proteins might contribute to preserving unfertilized egg viability.

### **Preservation of egg viability and quality is a multi-protein process**

Our study indicates that the CF ability to preserve egg viability during the post-ovulatory period is associated with the presence of multiple abundant proteins. Several studies have previously led to the identification of abundant CF proteins, including CF-specific proteins<sup>10,58,59</sup>. In the present study we have specifically investigated sialic acid synthase (Nansa) due to its presence in a CF fraction exhibiting high biological activity, because several sialic acid related proteins have been identified as being abundant only in salmonid species. Yet, supplementation of mineral medium with Nansa at two different concentrations did not improve egg viability and quality. While this could be due to methodological limitations, including concentration and biological activity, we can speculate that Nansa, alone, is not sufficient to preserve egg viability and quality. Furthermore, Nansa might not act directly on the eggs. Instead, it might be involved in the sialylation of other CF proteins that contribute more directly to egg viability preservation. For example, we identified ependymin among our candidates. As mentioned above, ependymin is a calcium binding protein and its binding property is dependent of its sialylation state. Together with the high number and functional diversity of proteins found in biologically active fractions and in the 3 salmonid species investigated these observations suggest that salmonid CF is a complex biological fluid. We propose that a significant number of proteins present in CF are collectively responsible for providing its unique ability to preserve egg viability and quality. Among these key players are proteins involved in immunity, calcium binding, lipid metabolism, proteolysis, extracellular matrix and sialic acid metabolic pathway.

## **Methods**

### **Coelomic fluid, egg and sperm collection**

Two-year old rainbow trout (*Oncorhynchus mykiss*) were obtained from the INRAE PEIMA experimental fish farm (Sizun, France) approximately 1 month before spawning. Fish were held during spawning season in a recirculated water system at 12 °C under a natural photoperiod (ISC LPGP, Rennes, France) until ovulation. Twice a week, fish were anaesthetized using tricaine methanesulfonate (MS222) at a concentration of 80 mg/L and checked for ovulation using gentle abdominal pressure. At the time of detected ovulation, eggs and coelomic fluid were collected by hand stripping (i.e., applying pressure onto the abdomen). Eggs were separated from CF with using a 0.5 mm mesh screen. A turbidity test was performed as previously described<sup>20</sup> for each collected CF to avoid any bias due to damaged or leaking eggs and only non-precipitating CF were kept. CF sample were then centrifuged at 3000g for 15 min at 4 °C to remove red blood cells and debris (pure CF). The collected supernatants were stored at -20 °C until use. When mentioned, the supernatants were either boiled at 95 °C for 10 min to denature proteins (boiled CF) or ultracentrifuged 1h, at 100,000g, at 4 °C to float up lipids (CF ultracentrifuged). CF depleted in lipids was collected from the bottom of the tube with the help of a needle

and stored at  $-20\text{ }^{\circ}\text{C}$  until use. Before storage at  $-20\text{ }^{\circ}\text{C}$ , protein concentration of the CF was determined using the Coomassie (Bradford) Protein assay Kit (Thermo Fisher Scientific, Rockford, IL, USA) and bovine serum albumin (BSA) as a calibration standard. To collect rainbow trout sperm, the genital region was wiped and sperm was obtained by manual stripping. A pool of sperm was made from 3 males and stored on ice until fertilization. For proteomic analysis, coelomic fluid (CF) samples were obtained from different Atlantic salmon (*Salmo salar*) ( $n=2$ ) and brown trout (*Salmo trutta*) ( $n=3$ ) individual females. For rainbow trout, CF samples were obtained as described above and whole proteome analysis was performed using a pool of CF samples originating from three different females. Goldfish (*Carassius auratus*) were obtained from the INRAE U3E facility in Rennes, France. Two-year-old fish raised in outdoor tanks were transferred into  $0.3\text{ m}^3$  tanks and reared several weeks in recycled water at  $14\text{ }^{\circ}\text{C}$  under a 14 h light and 10 h dark photoperiod (ISC-LPGP, Rennes, France). Fish were fed with carp pellets at a ratio of 1% of body weight. Ovarian fluid samples were obtained as previously described<sup>20</sup>. OF samples were collected from 3 different females and pooled before proteomic analysis. Common carp (*Cyprinus carpio*) OF samples were provided by the University of South Bohemia in České Budějovice, Faculty of Fisheries and Protection of Waters, České Budějovice, Czechia. Female fish were stocked in indoor tanks with controlled water and air supply, and temperature of  $18\text{--}22\text{ }^{\circ}\text{C}$ . OF samples were collected as previously described. For the proteomic analysis, OF samples originating from 3 different females were used<sup>60</sup>. Pikeperch (*Sander lucioperca*) OF samples obtained from domesticated pikeperch broodstock reproduced following standardized controlled reproduction protocol as previously described<sup>61</sup>. Fish were held at the Asialor fish farm where they were subjected to a specific annual photo-thermal program enabling proper course of gonadal development. Just prior to spawning ovulation was stimulated using salmon gonadolibertine analog ( $50\text{ }\mu\text{g}/\text{kg}$  of body weight; Bachem, Switzerland). When ovulation was observed, eggs were hand-stripped into a dry container and OF was aspirated with a pipet into cryotubes which were immediately after snap frozen in liquid nitrogen. OF samples were then stored at  $-80\text{ }^{\circ}\text{C}$ . For the current experiment OF was collected only from egg portions characterized by highest quality (fertilization rate above 70%). For the proteomic analysis, OF samples originating from 3 different females were used. CF and OF samples collected from those different species were processed as described above for rainbow trout CF.

### Tissue collection

For tissue collection, female rainbow trout were euthanized using a lethal dose of tricaine methanesulfonate (MS222) at a concentration of  $200\text{ mg}/\text{L}$ . Ovaries were dissected and ovarian samples were either frozen in liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$  until RNA extraction or homogenized in RIPA buffer (NaCl  $150\text{ mM}$ , deoxycholic acid  $0.5\%$ , NP40  $1\%$ , SDS  $0.1\%$ , Tris  $50\text{ mM}$ , pH 8) using a Precellys Evolution Homogenizer (Ozyme, Bertin Technologies) to extract proteins. After centrifugation, protein concentration in the supernatant was determined using the Coomassie (Bradford) Protein assay Kit (Thermo Fisher Scientific, Rockford, IL, USA) and bovine serum albumin (BSA) as a calibration standard. Supernatants were then stored at  $-20\text{ }^{\circ}\text{C}$  until use.

### Biological test to assess egg viability preservation potential

In vitro storage of unfertilized eggs was performed, under semi-sterile conditions in 6-well polystyrene culture plates (Falcon). Before storage, eggs were washed with filter-sterilized ( $0.22\text{ }\mu\text{m}$ ) trout mineral medium (MM)<sup>31</sup> ( $124.1\text{ mM NaCl}$ ,  $5.1\text{ mM KCl}$ ,  $1\text{ mM MgSO}_4\cdot 7\text{H}_2\text{O}$ ,  $1.6\text{ mM CaCl}_2\cdot 2\text{H}_2\text{O}$ ,  $5.6\text{ mM Glucose}$ ,  $26.5\text{ mM Hepes}$ , pH 8) to remove residual CF. A volume of  $3\text{ mL}$  of medium (CF, boiled CF, ultracentrifuged CF, diluted CF, CF fractions, or MM) was used in each well and eggs were added at a ratio of 35 eggs per well. Technical duplicates ( $2\times 35$  eggs) were used for each condition. Plates were then stored 3 days in the dark, in a  $12\text{ }^{\circ}\text{C}$  incubator, under gentle and constant agitation ( $50\text{ rpm}$ ) until fertilization. For CF fractions complemented with BSA or sialic acid synthase (Nansa), two concentrations were tested: 4 or  $40\text{ }\mu\text{g}/\text{mL}$ . The sialic acid synthase protein (NCBI accession # XP\_021429200) was commercially produced using the insect expression vector pFastBac1 (GenScript U488UFB180).

### Fertilization and developmental success monitoring

Fertilization was performed as previously described with minor modification<sup>7</sup>. After 3 days of in vitro storage, technical duplicates of 35 rainbow trout eggs were removed from their storage medium (CF, MM, CF fractions) using a fine-mesh screen and pooled into plastic tubes. Sperm was added ( $5\text{ }\mu\text{L}$ ) along with Actifish solution (IMV, Laigle, France). A pool of sperm originating from three different males was used. Tubes were gently swirled. After 5 min, Actifish solution was discarded, eggs were transferred into plastic incubators and held in the dark at  $10\text{ }^{\circ}\text{C}$  in a recirculated water system during early development. The number of eggs was counted. The number of embryos reaching eyed stage was counted 4 weeks after fertilization (dpf).

### RNA extraction and reverse transcription

RNA extraction was performed as previously described with minor modifications<sup>62</sup>. Briefly, tissues were homogenized in TRI Reagent (TR118, Euromedex) at a ratio of  $100\text{ mg}$  per  $\text{mL}$  of TRI Reagent. Total RNA was then extracted using the TRI Reagent procedure and resuspended in water. RNA concentration and purity was evaluated using a NanoDrop spectrophotometer. Reverse transcription was performed from  $1.6\text{ }\mu\text{g}$  of total RNA using Maxima First strand kit (ThermoScientific) according to manufacturer's instructions and using the following steps: 10 min at  $25\text{ }^{\circ}\text{C}$ , 30 min at  $60\text{ }^{\circ}\text{C}$ , and 5 min at  $85\text{ }^{\circ}\text{C}$ . Control reactions were run without reverse transcriptase and used as negative real-time PCR controls.

### Real-time PCR

RT products and negative control reaction products were diluted to 1/25 (1/2000 in the case of 18S) and 4  $\mu$ L were used in each PCR reaction. For each sample, triplicates were analyzed. Real-time PCR was performed using a real-time PCR kit provided with SYBR green fluorophore (Powerup SYBER Green Master Mix, Applied Biosystems) with 600 mM of each primer. For sialic acid synthase, the following primers were used ACCTGG CCCATCACTTTACG/ACATCACCCAGTGCCATGTT. Relative abundance of target cDNA in the sample was evaluated from a serially diluted cDNA pool (standard curve) using the StepOne + cycler software (Applied Biosystems). We observed a very stable expression of 18S with a coefficient of variation of 2.7% in the different rainbow trout tissues and 7.6% in the trout ovary at 0, 6, and 20 days post ovulation. Furthermore, we did not observe any significant difference between groups. Real-time PCR data were all normalized to 18S RNA abundance in the samples.

### SDS-PAGE and western blot

SDS-PAGE and western blot were performed as previously described<sup>62</sup> with minor changes. After resuspension in 2 $\times$  Laemmli buffer (0.125 M Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 10%  $\beta$ -mercaptoethanol)<sup>63</sup> at a final concentration of 1  $\mu$ g/ $\mu$ L, samples were reduced for 5 min at 95 °C. For each sample, 20  $\mu$ L were loaded per well on a 10% SDS-PAGE. A custom-made antibody directed against peptide CKVGEPRGVSPEDMG of the sialic acid synthase (NCBI accession # XP\_021429200) was purchased (GenScript U881EER070) and used at a dilution of 1/2000. Anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (Jackson ImmunoResearch, West Grove, USA) was used as secondary antibody at 1/5000 dilution. Detection was performed with Uptima Uptilight chemiluminescent reagent kit (Uptima-Interchim) using a Fusion FX7 imager (Vilbert Lourmat) with the Fusion software (v 15.11). Band intensities were measured using the ImageJ software. For each temporal profile, intensities were normalized by setting the maximal signal obtained to 100%.

Sample preparation of rainbow trout coelomic fluid before nanoLC-MS/MS analysis.

#### *HPLC fractionation of the rainbow trout coelomic fluid*

Coelomic fluid (CF) samples were desalted and concentrated by diafiltration (10 kDa Vivaspine, Sartorius) according to manufacturer instructions and then fractionated by UHPLC using sequential chromatography (gel filtration and anion exchange). First, samples (60 mg of proteins) were resuspended in 2 mL of 10 mM sodium phosphate/0.15 mM sodium chloride, pH 6.8 and loaded onto a gel filtration column (TSK gel G3000SW 21.5 mm i.d. $\times$ 30 cm, 13  $\mu$ m; Tosoh Bioscience, Tokyo, Japan) at a flow rate of 4 mL/min. Bound proteins were eluted isocratically using the phosphate buffer during 40 min. We monitored the UV absorbance of the eluent at 210 and 280 nm. Thirty fractions were collected and systematically assessed for egg quality preservation using the *in vitro* egg storage test followed by fertilization and developmental success monitoring.

Fractions of interest were further selected for a second chromatography fractionation. They were desalted and resuspended in 250  $\mu$ L of 10 mM Tris pH 7.6. The samples were then loaded onto an anion exchange column (ProSwift SAX-1S 4.6  $\times$  50 mm PK/SS; Thermo Fisher Scientific, Courtaboeuf, France) at a flow rate of 1 mL/min. The mobile phases used were (A) 10 mM Tris, pH 7.6 and (B) 1 M NaCl in 10 mM Tris, pH 7.6. After sample loading, the gradient applied was (min:%B); 0:0, 5:2, 30:50, 40:100, 50:100, 60:0, 120:0. Fractions were collected every 60 s from 0 to 46 min of the gradient, resulting in 31 fractions. The detection of protein was carried out at 210 nm and 280 nm. These fractions were also tested on ovulated eggs and proteins in fractions of interest were further analyzed by mass spectrometry after liquid trypsin digestion for protein identification.

The fractions of interest were desalted by diafiltration (10 kDa Vivaspine), and resuspended in 20  $\mu$ L of 50 mM ammonium bicarbonate/0.01% ProteaseMax (Promega, Charbonnières-les-Bains, France) then incubated in 2.5  $\mu$ L of 65 mM DTT at 37 °C for 15 min followed by incubation at room temperature in the dark for 15 min after addition of 2.5  $\mu$ L of 135 mM iodoacetamide. Finally, 2  $\mu$ L of sequencing grade modified trypsin (Promega) at a concentration of 0.1  $\mu$ g/ $\mu$ L and 23  $\mu$ L of 50 mM ammonium bicarbonate/0.01% ProteaseMax were added. The trypsin digestion was performed overnight at 37 °C and resulted peptides were analyzed by nanoLC-MS/MS using a short nanoLC gradient.

#### *Proteomic comparison of salmonid CF to non-salmonid OF*

Samples of salmonid CF and non-salmonid OF (2  $\mu$ g of proteins for each sample) were subjected to a short migration on a NuPAGE system (Invitrogen, 3 min at 200 V–400 mA–23 W) according to the manufacturer's instructions, in order to stack the proteins of each sample in one band. After coloration using the EZBlue Gel Staining Reagent (Sigma-Aldrich, Saint-Quentin Fallavier, France) bands were excised from the gel and processed for tryptic digestion and peptide extraction as previously described<sup>64</sup>. Resulting peptides were analyzed by nanoLC-MS/MS using a short nanoLC gradient (see below).

#### *NanoLC-MS/MS analysis*

For each sample, 300 ng of tryptic peptide mixtures were separated onto a 75  $\mu$ m  $\times$  250 mm IonOpticks Aurora 2 C18 column (Ion Opticks Pty Ltd., Bundoora, Australia). A gradient of basic reversed-phase buffers (Buffer A: 0.1% formic acid, 98% H<sub>2</sub>O MilliQ, 2% acetonitrile; Buffer B: 0.1% formic acid, 100% acetonitrile) was run on a NanoElute HPLC System (Bruker Daltonik GmbH, Bremen, Germany) at a flow rate of 300 nL/min at 50 °C for the HPLC fractions and 400 nL/min at 50 °C for non-fractionated samples. For the HPLC fractions, the liquid chromatography (LC) run lasted for 40 min (2–11% of buffer B during 19 min; up to 16% at 26 min; up to 25% at 30 min; up to 85% at 33 min and finally 85% for 7 min to wash the column). For the non-fractionated samples, the liquid chromatography (LC) run lasted for 120 min (2–15% of buffer B during 60 min; up to 25% at 90 min; up to 37% at 100 min; up to 95% at 110 min and finally 95% for 10 min to wash the column).

The column was coupled online to a tims TOF Pro (Bruker Daltonik GmbH, Bremen, Germany) with a Captive Spray ion source (Bruker Daltonik). The temperature of the ion transfer capillary was set at 180 °C. Ions were accumulated for 114 ms, and mobility separation was achieved by ramping the entrance potential from –160 to –20 V within 114 ms. The acquisition of the MS and MS/MS mass spectra was done with average resolutions of 60,000 and 50,000 full width at half maximum (mass range 100–1700 m/z), respectively. To enable the PASEF method<sup>65</sup>, precursor m/z and mobility information was first derived from full scan TIMS-MS experiments (with a mass range of m/z 100–1700). The quadrupole isolation width was set to 2 and 3 Th and, for fragmentation, the collision energies varied between 31 and 52 eV depending on the precursor mass and charge. Tims, MS operation and PASEF were controlled and synchronized using the control instrument software OtofControl 5.1 (Bruker Daltonik). LC-MS/MS data were acquired using the PASEF method with a total cycle time of 1.31 s, including 1 tims MS scan and 10 PASEF MS/MA scans. The 10 PASEF scans (100 ms each) containing, on average, 12 MS/MS scans per PASEF scan. Ion mobility-resolved mass spectra, nested ion mobility versus m/z distributions, as well as summed fragment ion intensities were extracted from the raw data file with DataAnalysis 5.1 (Bruker Daltonik).

Generated spectra were analyzed with the Mascot database search engine (v2.6.2; <http://www.matrixscience.com>) for peptide and protein identification, using its automatic decoy database search to estimate a false discovery rate (FDR) and calculate the threshold at which the e-values of the identified peptides were valid. According to the samples, the spectra were queried simultaneously in their species corresponding proteome database (NCBI *Oncorhynchus mykiss*—20200923—97,743 sequences; NCBI *Carassius auratus*—20180809—96,703 sequences; NCBI *Salmo salar*—20210421—97,555 sequences; NCBI *Cyprinus carpio*—20210512—63,928 sequences; NCBI *Salmo trutta*—20210424—87,841 sequences; NCBI *Sander lucioperca*—20200814—56,557 sequences) and in a database containing known contaminants such as human keratins or porcine trypsin (247 sequences). Mass tolerance for MS and MS/MS was set at 15 ppm and 0.05 Da, respectively. The enzyme selectivity was set to full trypsin with one mis-cleavage allowed. Protein modifications were fixed carbamidomethylation of cysteines, variable oxidation of methionine. Identified proteins are validated with an FDR < 1% at PSM level and an e-value < 0.01, using Proline Studio v2.1.2 software<sup>66</sup>. Proteins identified with the same set of peptides were automatically grouped together.

#### Protein quantitation and annotation

Two groups were compared, salmonids (rainbow trout, Atlantic salmon, brown trout) and non-salmonid teleosts (goldfish, common carp, pikeperch). Only proteins identified in all the replicates (when available) of a specific species with a minimum of 2 peptides were kept for further analysis. When biological replicates of the same species were available, the mean value was used. Proteins of the different species sharing a common ancestor (post teleost-specific whole genome duplication, TGD) were identified using Genomicus (GenomicusFish v04.02—Gene Search)<sup>33</sup>. Common ancestors were then used to perform protein repertoire comparison between salmonids and non-salmonids. For each protein, the mean number of spectral counts was calculated and normalized to the length of the protein (N-SC). In several cases, proteins had the same common ancestor, the sum of their N-SCs was therefore used for the quantification. The proteomic quantitative and statistical analysis was performed using ProStaR v1.28.0<sup>67</sup>. Only proteins that were identified in at least two species of one of the two groups were kept for further analysis. Then, data were normalized by centering N-SC distributions on the median. Missing values were inferred as follows: partially observed values (POV) were inferred using SLSA algorithm and values missing on entire condition (MEC) were inferred using DetQuantile method. Between-groups comparisons were performed with a limma test and Benjamini–Hochberg *p* value correction. A  $-\log_{10}(p \text{ value})$  cut-off of 2.35 ( $p \text{ value} = 0.00447$ ) was calculated for a FDR < 1%.

Gene ontology analysis was performed using ShinyGO<sup>68</sup>. First, human orthologs of the proteins identified in the rainbow trout coelomic fluid were obtained using genomicus. (GenomicusFish v04.02—Gene Search). The corresponding human Ensembl gene numbers were then used for gene ontology annotation. The 500 most significantly enriched biological processes GO-terms (FDR < 0.05) were retrieved and hierarchically clustered. GO-terms sharing 70% of their proteins were clustered together. Proteins of the main clusters, composed of at least 4 GO-terms, were further subjected to GO annotation to identify the high-level GO categories representing the different clusters, and the most represented GO-term in each cluster.

#### Statistics

Statistical analyzes were performed using R software. Normal distribution of the variables was evaluated using Shapiro–Wilk test. When the distribution was normal, statistical difference of the means was evaluated using a t-test. Otherwise, a Wilcoxon–Mann Whitney test was performed. The proteomic quantitative and statistical analysis was performed as described above, using the ProStaR v1.28.0 software<sup>67</sup>, dedicated to quantitative differential analysis of proteomics data.

#### Ethical statement

Experiments and procedures were fully compliant with French and European animal welfare policies and followed guidelines of the INRAE LPGP Institutional Animal Care and Use Ethical Committee, which specifically approved this study. The study was carried out in compliance with the ARRIVE guidelines.

#### Data availability

All mass spectrometry proteomics data have been deposited on the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository<sup>69</sup> under the dataset identifier PXD034989 and <https://doi.org/10.6019/PXD034989>.

Received: 5 September 2023; Accepted: 18 April 2024

Published online: 26 April 2024

## References

1. Bobe, J. & Labbé, C. Egg and sperm quality in fish. *Gen. Comp. Endocrinol.* **165**, 535–548 (2010).
2. Migaud, H. *et al.* Gamete quality and broodstock management in temperate fish. *Rev. Aquac.* **5**, S194–S223 (2013).
3. Formacion, M. J., Hori, R. & Lam, T. J. Overripening of ovulated eggs in goldfish. *Aquaculture* **114**, 155–168 (1993).
4. Samarin, A. M., Gela, D., Bytuytsky, D. & Policar, T. Determination of the best post-ovulatory stripping time for the common carp (*Cyprinus carpio* Linnaeus, 1758). *J. Appl. Ichthyol.* **31**, 51–55 (2015).
5. Springate, J. R. C., Bromage, N. R., Elliott, J. A. K. & Hudson, D. L. The timing of ovulation and stripping and their effects on the rates of fertilization and survival to eying, hatch and swim-up in the rainbow trout (*Salmo gairdneri* R.). *Aquaculture* **43**, 313–322 (1984).
6. Sakai, K., Nomura, M., Takashima, F. & Oto, H. The over-ripening phenomenon of rainbow trout-II, changes in the percentage of eyed eggs, hatching rate and incidence of abnormal alevins during the process of over-ripening. *Nippon Suisan Gakkaishi* **41**, 855–860 (1975).
7. Aegerter, S., Jalabert, B. & Bobe, J. Large scale real-time PCR analysis of mRNA abundance in rainbow trout eggs in relationship with egg quality and post-ovulatory ageing. *Mol. Reprod. Dev.* **72**, 377–385 (2005).
8. Coffman, M. A. & Goetz, F. W. Trout ovulatory proteins are partially responsible for the anti-proteolytic activity found in trout coelomic fluid. *Biol. Reprod.* **59**, 497–502 (1998).
9. Coffman, M. A., Pinter, J. H. & Goetz, F. W. Trout ovulatory proteins: Site of synthesis, regulation, and possible biological function. *Biol. Reprod.* **62**, 928–938 (2000).
10. Bobe, J. & Goetz, F. W. An ovarian progesterone is present in the trout coelomic fluid after ovulation. *Biol. Reprod.* **64**, 1048–1055 (2001).
11. Bonnet, E., Jalabert, B. & Bobe, J. A 3-day in vitro storage of rainbow trout (*Oncorhynchus mykiss*) unfertilised eggs in coelomic fluid at 12°C does not affect developmental success. *Cybio* **27**, 47–51 (2003).
12. Billard, R. *et al.* Biology of the gametes of some teleost species. *Fish Physiol. Biochem.* **2**, 115–120 (1986).
13. Luchi, I., Masuda, K. & Yamagami, K. Change in component proteins of the egg envelope (chorion) of rainbow trout during hardening: (Fish egg envelope/chorion/chorion proteins/chorion hardening/egg activation). *Dev. Growth Differ.* **33**, 85–92 (1991).
14. Masuda, K., Luchi, I. & Yamagami, K. Analysis of hardening of the egg envelope (chorion) of the fish, *Oryzias latipes*: (Egg envelope (chorion)/egg activation/chorion hardening/fish egg/chorion proteins). *Dev. Growth Differ.* **33**, 75–83 (1991).
15. Lahnsteiner, F. The influence of ovarian fluid on the gamete physiology in the Salmonidae. *Fish Physiol. Biochem.* **27**, 49–59 (2002).
16. Johnson, S. L. *et al.* Ovarian fluid proteome variation associates with sperm swimming speed in an externally fertilizing fish. *J. Evol. Biol.* **33**, 1783–1794 (2020).
17. Myers, J. N., Senior, A., Zadmajid, V., Sørensen, S. R. & Butts, I. A. E. Associations between ovarian fluid and sperm swimming trajectories in marine and freshwater teleosts: A meta-analysis. *Rev. Fish. Sci. Aquac.* **28**, 322–339 (2020).
18. Zadmajid, V., Myers, J. N., Sørensen, S. R. & Ernest Butts, I. A. Ovarian fluid and its impacts on spermatozoa performance in fish: A review. *Theriogenology* **132**, 144–152 (2019).
19. Bail, P.-Y. *et al.* Optimization of somatic cell injection in the perspective of nuclear transfer in goldfish. *BMC Dev. Biol.* **10**, 64 (2010).
20. Depince, A., Marandel, L., Goardon, L., Le Bail, P.-Y. & Labbe, C. Trout coelomic fluid suitability as Goldfish oocyte extender can be determined by a simple turbidity test. *Theriogenology* **75**, 1755–1761 (2011).
21. Corley-Smith, G. E., Lim, C. J. & Brandhorst, B. P. Production of androgenetic zebrafish (*Danio rerio*). *Genetics* **142**, 1265–1276 (1996).
22. Lahnsteiner, F., Weismann, T. & Patzner, R. Composition of the ovarian fluid in 4 salmonid species: *Oncorhynchus mykiss*, *Salmo trutta flaccustris*, *Salvelinus alpinus* and *Hucho hucho*. *Reprod. Nutr. Dev.* **35**, 465–474 (1995).
23. Rosengrave, P. *et al.* Chemical composition of seminal and ovarian fluids of chinook salmon (*Oncorhynchus tshawytscha*) and their effects on sperm motility traits. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **152**, 123–129 (2009).
24. Hatef, A., Niksirat, H. & Alavi, S. M. H. Composition of ovarian fluid in endangered Caspian brown trout, *Salmo trutta caspius*, and its effects on spermatozoa motility and fertilizing ability compared to freshwater and a saline medium. *Fish Physiol. Biochem.* **35**, 695–700 (2009).
25. Goetz, F. W. & Coffman, M. A. Storage of unfertilized eggs of rainbow trout (*Oncorhynchus mykiss*) in artificial media. *Aquaculture* **184**, 267–276 (2000).
26. Hajnik, C. A., Goetz, F. W., Hsu, S.-Y. & Sokal, N. Characterization of a ribonucleic acid transcript from the brook trout (*Salvelinus fontinalis*) ovary with structural similarities to mammalian adipisin/complement factor D and tissue kallikrein, and the effects of kallikrein-like serine proteases on follicle contraction. *Biol. Reprod.* **58**, 887–897 (1998).
27. Rime, H. *et al.* Post-ovulatory ageing and egg quality: A proteomic analysis of rainbow trout coelomic fluid. *Reprod. Biol. Endocrinol.* **2**, 26 (2004).
28. Dietrich, G. J. *et al.* Broken eggs decrease pH of rainbow trout (*Oncorhynchus mykiss*) ovarian fluid. *Aquaculture* **273**, 748–751 (2007).
29. Nynca, J., Arnold, G. J., Fröhlich, T. & Ciereszko, A. Shotgun proteomics of rainbow trout ovarian fluid. *Reprod. Fertil. Dev.* **27**, 504–512 (2015).
30. Johnson, S. L. *et al.* Proteomic analysis of chinook salmon (*Oncorhynchus tshawytscha*) ovarian fluid. *PLoS ONE* **9**, e104155 (2014).
31. Wolf, K. & Quimby, M. C. 5 Fish cell and tissue culture. In *Fish Physiology*, vol. 3, 253–305 (Elsevier, 1969).
32. Pasquier, J. *et al.* Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genom.* **17**, 368 (2016).
33. Parey, E. *et al.* Genome structures resolve the early diversification of teleost fishes. *Science* **379**, 572–575 (2023).
34. Schwarzkopf, M. *et al.* Sialylation is essential for early development in mice. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5267–5270 (2002).
35. Asahina, S., Sato, C. & Kitajima, K. Developmental expression of a sialyltransferase responsible for sialylation of cortical alveolus glycoprotein during oogenesis in rainbow trout (*Oncorhynchus mykiss*). *J. Biochem. (Tokyo)* **136**, 189–198 (2004).
36. Abeln, M. *et al.* Sialic acid is a critical fetal defense against maternal complement attack. *J. Clin. Investig.* **129**, 422–436 (2019).
37. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–226 (2015).
38. Chen, F., Spiessens, C., D'Hooghe, T., Peeraer, K. & Carpentier, S. Follicular fluid biomarkers for human in vitro fertilization outcome: Proof of principle. *Proteome Sci.* **14**, 17 (2016).
39. Georgiou, A. S. *et al.* Effects of complement component 3 derivatives on pig oocyte maturation, fertilization and early embryo development in vitro. *Reprod. Domest. Anim. Zuchtig.* **46**, 1017–1021 (2011).
40. Yoo, S. W. *et al.* Complement factors are secreted in human follicular fluid by granulosa cells and are possible oocyte maturation factors. *J. Obstet. Gynaecol. Res.* **39**, 522–527 (2013).
41. Espey, L. L. Ovulation as an inflammatory reaction—A hypothesis. *Biol. Reprod.* **22**, 73–106 (1980).
42. Duffy, D. M., Ko, C., Jo, M., Brannstrom, M. & Curry, T. E. Ovulation: Parallels with inflammatory processes. *Endocr. Rev.* **40**, 369–416 (2019).

43. Bobe, J., Montfort, J., Nguyen, T. & Fostier, A. Identification of new participants in the rainbow trout (*Oncorhynchus mykiss*) oocyte maturation and ovulation processes using cDNA microarrays. *Reprod. Biol. Endocrinol. RBE* **4**, 39 (2006).
44. Lubzens, E., Young, G., Bobe, J. & Cerdà, J. Oogenesis in teleosts: How eggs are formed. *Gen. Comp. Endocrinol.* **165**, 367–389 (2010).
45. Gerwick, L., Reynolds, W. S. & Bayne, C. J. A precerebellin-like protein is part of the acute phase response in rainbow trout. *Oncorhynchus mykiss. Dev. Comp. Immunol.* **24**, 597–607 (2000).
46. Ganss, B. & Hoffmann, W. Calcium binding to sialic acids and its effect on the conformation of ependymins. *Eur. J. Biochem.* **217**, 275–280 (1993).
47. Lee, B.-I., Mustafi, D., Cho, W. & Nakagawa, Y. Characterization of calcium binding properties of lithostathine. *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.* **8**, 341–347 (2003).
48. Kitazume, S., Kitajima, K., Inoue, S., Inoue, Y. & Troy, F. A. Developmental expression of trout egg polysialoglycoproteins and the prerequisite alpha 2,6-, and alpha 2,8-sialyl and alpha 2,8-polysialyltransferase activities required for their synthesis during oogenesis. *J. Biol. Chem.* **269**, 10330–10340 (1994).
49. Blaum, B. S. *et al.* Structural basis for sialic acid-mediated self-recognition by complement factor H. *Nat. Chem. Biol.* **11**, 77–82 (2015).
50. Babin, P. J., Carnevali, O., Lubzens, E. & Schneider, W. J. Molecular aspects of oocyte vitellogenesis in fish. In *The Fish Oocyte: From Basic Studies to Biotechnological Applications* (eds Babin, P. J. *et al.*) 39–76 (Springer, 2007). [https://doi.org/10.1007/978-1-4020-6235-3\\_2](https://doi.org/10.1007/978-1-4020-6235-3_2).
51. Menn, F. L., Cerdà, J. & Babin, P. J. Ultrastructural aspects of the ontogeny and differentiation of ray-finned fish ovarian follicles. In *The Fish Oocyte: From Basic Studies to Biotechnological Applications* (eds Babin, P. J. *et al.*) 1–37 (Springer, 2007). [https://doi.org/10.1007/978-1-4020-6235-3\\_1](https://doi.org/10.1007/978-1-4020-6235-3_1).
52. Thomé, R., dos Santos, H. B., Sato, Y., Rizzo, E. & Bazzoli, N. Distribution of laminin  $\beta$ 2, collagen type IV, fibronectin and MMP-9 in ovaries of the teleost fish. *J. Mol. Histol.* **41**, 215–224 (2010).
53. Curry, T. E. & Osteen, K. G. The matrix metalloproteinase system: Changes, regulation, and impact throughout the ovarian and uterine reproductive cycle. *Endocr. Rev.* **24**, 428–465 (2003).
54. Saftig, P. *et al.* Functions of cathepsin K in bone resorption. Lessons from cathepsin K deficient mice. *Adv. Exp. Med. Biol.* **477**, 293–303 (2000).
55. Oksjoki, S., Söderström, M., Vuorio, E. & Anttila, L. Differential expression patterns of cathepsins B, H, K, L and S in the mouse ovary. *Mol. Hum. Reprod.* **7**, 27–34 (2001).
56. Bilen, E. *et al.* Do follicular fluid gelatinase levels affect fertilization rates and oocyte quality?. *Arch. Gynecol. Obstet.* **290**, 1265–1271 (2014).
57. Roa-Espitia, A. L. *et al.* Focal adhesion kinase is required for actin polymerization and remodeling of the cytoskeleton during sperm capacitation. *Biol. Open* **5**, 1189–1199 (2016).
58. Coffman, M. A. & Goetz, F. W. Trout ovulatory proteins are partially responsible for the anti-proteolytic activity found in trout coelomic fluid. *Biol. Reprod.* **59**, 497–502 (1998).
59. Matsubara, T., Hara, A. & Takano, K. Immunochemical identification and purification of coelomic fluid-specific protein in chum salmon (*Oncorhynchus keta*). *Comp. Biochem. Physiol. Part B Comp. Biochem.* **81**, 309–314 (1985).
60. Kholodnyy, V. *et al.* Common carp spermatozoa performance is significantly affected by ovarian fluid. *Aquaculture* **554**, 738148 (2022).
61. Żarski, D. *et al.* 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**, 527–536 (2019).
62. Rime, H., Nguyen, T., Ombredane, K., Fostier, A. & Bobe, J. Effects of the anti-androgen cyproterone acetate (CPA) on oocyte meiotic maturation in rainbow trout (*Oncorhynchus mykiss*). *Aquat. Toxicol.* **164**, 34–42 (2015).
63. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
64. Yilmaz, O. *et al.* Scrambled eggs: Proteomic portraits and novel biomarkers of egg quality in zebrafish (*Danio rerio*). *PLoS ONE* **12**, e0188084 (2017).
65. Meier, F. *et al.* Online parallel accumulation-serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer. *Mol. Cell. Proteom.* **17**, 2534–2545 (2018).
66. Bouyssie, D. *et al.* Proline: An efficient and user-friendly software suite for large-scale proteomics. *Bioinformatics* **36**, 3148–3155 (2020).
67. Wiczorek, S. *et al.* DAPAR & ProStaR: Software to perform statistical analyses in quantitative discovery proteomics. *Bioinformatics* **33**, 135–136 (2017).
68. Ge, S. X., Jung, D. & Yao, R. ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
69. Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

## Acknowledgements

The authors thank the INRAE ISC-LPGP for fish breeding. This study was funded by Agence Nationale de Recherche (ANR) Grant EggPreserve # ANR-16-CE20-0001) to JB and CP.

## Author contributions

A.G., D.Z. and J.B. wrote the main manuscript text. A.G. and D.Z. prepared the figures. C.P. and J.B. conceived the study. J.B. coordinated the study. A.G., H.R., E.C., B.G., R.L. and T.N. performed the experiments. A.G., E.C., D.Z., and J.M. analyzed the data. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-60118-2>.

**Correspondence** and requests for materials should be addressed to J.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024