



Article

# Proteomics Analysis of Early Developmental Stages of Zebrafish Embryos

Kathiresan Purushothaman <sup>1</sup>, Prem Prakash Das <sup>2</sup>, Christopher Presslauer <sup>1</sup>, Teck Kwang Lim <sup>2</sup>, Steinar D. Johansen <sup>1</sup>, Qingsong Lin <sup>2,\*</sup> and Igor Babiak <sup>1,\*</sup>

<sup>1</sup> Genomics Group, Faculty of Biosciences and Aquaculture, Nord University, 8049 Bodø, Norway; kathiresan.purushothaman@nord.no (K.P.); cpressla@hotmail.com (C.P.); steinar.d.johansen@nord.no (S.D.J.)

<sup>2</sup> Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore; dbspdd@nus.edu.sg (P.P.D.); dbslimtk@nus.edu.sg (T.K.L.)

\* Correspondence: dbslinqs@nus.edu.sg (Q.L.); igor.s.babiak@nord.no (I.B.)

Received: 6 November 2019; Accepted: 13 December 2019; Published: 17 December 2019



**Abstract:** Zebrafish is a well-recognized organism for investigating vertebrate development and human diseases. However, the data on zebrafish proteome are scarce, particularly during embryogenesis. This is mostly due to the overwhelming abundance of egg yolk proteins, which tend to mask the detectable presence of less abundant proteins. We developed an efficient procedure to reduce the amount of yolk in zebrafish early embryos to improve the Liquid chromatography–tandem mass spectrometry (LC–MS)-based shotgun proteomics analysis. We demonstrated that the de yolking procedure resulted in a greater number of proteins being identified. This protocol resulted in approximately 2-fold increase in the number of proteins identified in de yolked samples at cleavage stages, and the number of identified proteins increased greatly by 3–4 times compared to non-de yolked samples in both oblong and bud stages. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed a high number of functional proteins differentially accumulated in the de yolked versus non-de yolked samples. The most prominent enrichments after the de yolking procedure included processes, functions, and components related to cellular organization, cell cycle, control of replication and translation, and mitochondrial functions. This de yolking procedure improves both qualitative and quantitative proteome analyses and provides an innovative tool in molecular embryogenesis of polylecithal animals, such as fish, amphibians, reptiles, or birds.

**Keywords:** egg yolk; embryonic development; LC–MS/MS shotgun proteomics; proteome; zebrafish

## 1. Introduction

Zebrafish have become a prominent and broadly used model system to study developmental biology, neurogenetic disorders, genetics, toxicology, reproduction, pathology, and pharmacology [1–5]. The genome annotation is relatively well developed [6], and the embryonic transcriptome of zebrafish has been characterized in several studies [7–11]. However, knowledge about the comprehensive proteome dynamics during embryogenesis in zebrafish remains elusive.

Proteome in zebrafish is usually investigated in adult organs or tissues [12–15]. The overwhelming occurrence of vitellogenin yolk proteins is a limiting factor in a polylecithal embryo, such as in zebrafish, as it hinders global identification of less abundant proteins using mass spectrometry-based techniques [4,16]. Proteolytic peptides of yolk proteins can potentially subdue the ionization of the less abundant proteolytic peptides of non-yolk proteins [17,18]. Consequently, abundant yolk proteins can potentially interfere with the identification of cellular proteins, although the degree of such interference is unknown. To reduce the abundance of yolk proteins, de yolking protocols are employed; they have been used in a number of studies on zebrafish embryos and larvae from 3.3 h post-fertilization (hpf) to

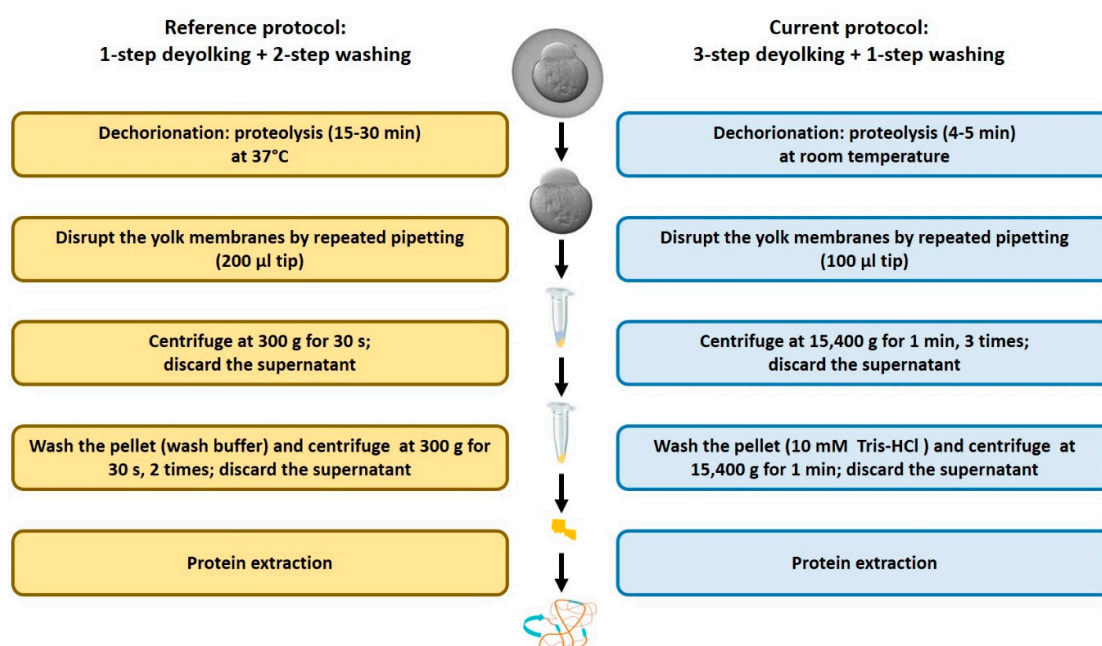
7 days post-fertilization (dpf) [16,19–21]. In most extensive studies to date, 5267 and 8363 proteins were identified in zebrafish de yolked embryos at 24 hpf [22,23].

So far, all the studies on zebrafish embryonic proteome were conducted on embryos being at a certain developmental advancement, and the information on the early stages, particularly before the maternal-to-zygotic transition (MZT), is missing. Pre-MZT stages of development are characterized by rapid, synchronous cell cycles (cleavages), and the development is driven by maternally-provided factors, including transcriptome and proteome [24]. Therefore, a knowledge of maternal proteome dynamics seems to be essential for understanding the regulation of early embryonic development in zebrafish. We improved the de yolking procedure, allowing the efficient capture and identification of proteins from the onset of development (1-cell stage). The protocol yielded 2 times more identified proteins compared to the non-de yolked counterparts in cleavage stages, and 3–4 times at oblong and bud stages. Also, the protocol caused minimal loss of proteins. Our improved protocol was effective for the subsequent systematic proteomics studies of zebrafish early embryonic development, and it is applicable to studies on other polylecithal animals.

## 2. Results

### 2.1. Efficiency of the New Extraction Protocol

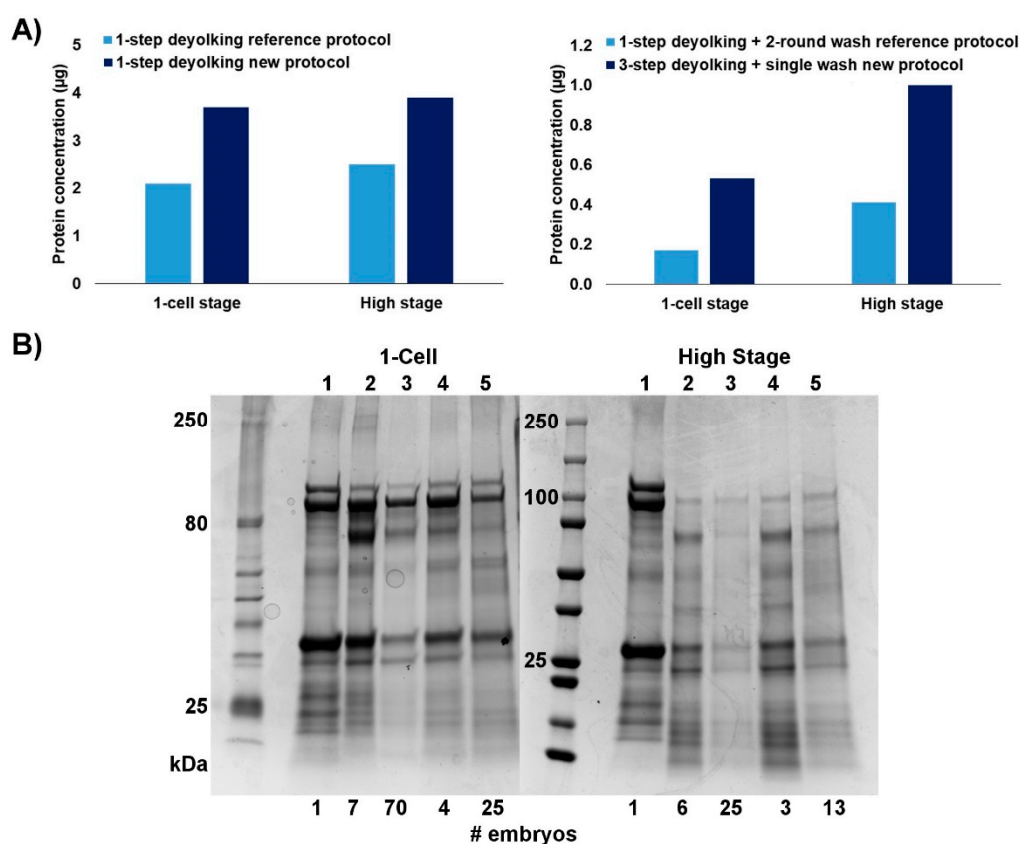
Application of the existing de yolking protocol [16] to zebrafish early embryos requires a considerable amount of embryos to be sampled, yet the representation of low-abundance proteins is reduced (unpublished observation). Therefore, we developed an improved protocol. The major differences are related to the timing and temperature of the de choriation step, separation of the protein pellet from a liquid fraction, and the wash step (Figure 1).



**Figure 1.** Schematic representation of major differences between the reference [16] and the current de yolking protocols. The detailed information is given in the text.

We compared our protocol to the protocols by Link et al. [16], which were based on 1-step de yolking with or without subsequent washing steps. For a fair comparison, we compared the reference 1-step de yolking procedure without washing [16] to our 1-step de yolking procedure without washing, and the reference 1-step de yolking plus double wash procedure [16] to our 3-step de yolking plus single wash procedure.

Both methods resulted in a reduction of yolk proteins, and the washing steps further depleted the protein content. Nevertheless, our new protocol yielded a larger number of unique proteins from a smaller number of embryos. We obtained approximately 1.7-fold increase in protein concentration per embryo sample when applying the 1-step deysolking process. When using our 3-step deysolking + single wash protocol, the protein yields per embryo sample were 3.1- and 2.5-fold higher (1-cell and high stage embryos, respectively) than those obtained with the reference protocols [16] with 1-step deysolking + double wash (Figure 2A). The effective number of 1-cell stage embryos needed to collect a workable amount of protein (30 µg) was approximately 2 or 3 times lower when using our 1-step deysolking or 3-step deysolking + single wash protocols, respectively; for the high stage embryos, this number of embryos was approximately 2 times lower than that of the respective reference protocols (Figure 2B). Also, compared to our new protocol, less amount of proteins was harvested with the reference protocol (Figure 2B).



**Figure 2.** Comparison of the efficiencies of deysolking protocols: reference [16] versus the new one. (A) Protein concentration obtained using protocols in 1-step versions (left chart) and in full versions (right chart). (B) SDS-PAGE of proteins extracted from zebrafish embryos at 1-cell stage (left panel) and high stage (right panel) using the reference protocols versus new protocols. Lane 1—non-deysolking embryo (control); Lane 2—1-step deysolking reference protocol [16]; Lane 3—1-step deysolking + double wash reference protocol [16]; Lane 4—1-step deysolking (new method); and Lane 5—3-step deysolking + single wash (new method). At the bottom line, number of embryos is given for each sample, from which the proteins were extracted.

## 2.2. Proteome in Deysolking Versus Non-Deysolking Samples

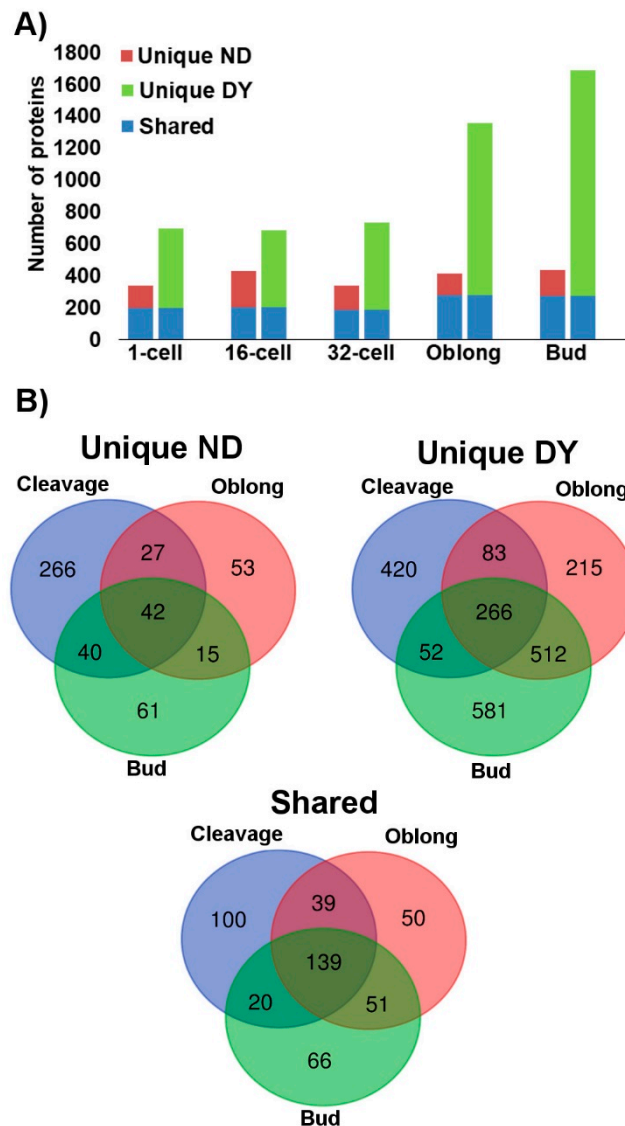
Generally, the amount of extracted total protein per embryo increased with the developmental advancement of the embryo, and the deysolking procedure greatly reduced the protein concentration. However, this reduction was decreasing from over 25-fold in cleavage stages (1-cell, 16-cell, and 32-cell stages) to approximately 15-fold in the bud stage (Table 1).

**Table 1.** The amount of protein extracted from de-yolked versus non-de-yolked embryos.

Developmental Stage	Number of Embryos	Total Sample Volume ( $\mu$ L)	Amount of Extracted Protein ( $\mu$ g)		
			Total	Per $\mu$ L	Per Embryo
Non-de-yolked					
1-cell	28	119	395.08	3.32	14.11
16-cell	20	88	327.36	3.72	16.37
32-cell	40	170	697.00	4.10	17.42
Oblong	20	120	478.80	3.99	23.94
Bud	20	92	524.40	5.70	26.22
De-yolked					
1-cell	575	94	304.56	3.24	0.53
16-cell	300	58	191.41	3.30	0.63
32-cell	400	99	277.22	2.81	0.69
Oblong	225	42	246.54	5.87	1.09
Bud	250	70	413.70	5.91	1.65

Analysis of the digested protein samples using the one-dimensional (1D) mass spectrometry (MS)/MS shotgun proteomics approach (1D shotgun) consistently demonstrated that the de-yolking procedure resulted in a greater number of proteins being identified (Supplementary File 1). In the non-de-yolked samples, the total numbers of proteins identified throughout the developmental stages were relatively consistent, ranging from 338 to 434 proteins identified in the 1-cell and bud stages, respectively. By comparison, the numbers of proteins identified in de-yolked samples in all the developmental stages were considerably higher than in the non-de-yolked counterparts. In the cleavage stages, these differences were approximately 2-fold, and increased to over 3-fold in the later developmental stages, ranging from 696 to 1687 proteins identified in the 1-cell and bud stages, respectively (Figure 3A, Supplementary File 1). In contrast to the non-de-yolked samples, where there was no apparent correlation between the developmental progression and the total number of proteins identified, de-yolked samples resulted in a consistent number of proteins identified throughout cleavage stages (1-cell, 16-cell, and 32-cell), which considerably increased in the later developmental stages (Figure 3A). Most of the proteins identified in the non-de-yolked samples were also found in the de-yolked counterparts (Figure 3A, Supplementary File 1). The number of proteins unique to the non-de-yolked samples (that is, not found in the de-yolked counterparts) was relatively stable throughout the developmental stages. In contrast, most of the proteins identified in the de-yolked samples were unique, meaning that they were not found in the non-de-yolked counterparts, and the number of unique proteins apparently increased throughout the embryonic development from the cleavage stages to the bud stage (Figure 3A, Supplementary File 1).

When looking only to the proteins shared between the non-de-yolked and de-yolked samples, representation of vitellogenin in the de-yolked samples was substantially reduced (36–58 times, depending on developmental stage; Supplementary File 2). At the same time, the representation of non-vitellogenin proteins in the de-yolked samples was considerably elevated (2–6 times, depending on developmental stage; Supplementary File 3).



**Figure 3.** Numbers of proteins identified in samples from intact (non-deyolged, ND) versus deyolged (DY) zebrafish embryos. Unique proteins were found in either ND or DY embryos, whereas shared proteins were found in both ND and DY embryos. (A) Total number of unique and shared proteins in ND (left column) and DE embryos (right column) at 1-cell, 16-cell, 32-cell, oblong, and bud developmental stages. (B) Specificity and overlap of the identified proteins across the critical stages of early embryonic development: Cleavage stages (1-, 16-, and 32-cell stages combined), maternal–zygotic transition (oblong), and post-maternal–zygotic transition (bud).

Among the 504 proteins unique to non-deyolged samples, most of them were specific to the cleavage stages, and 42 proteins were found in all the developmental stages. By comparison, out of 2129 proteins unique to the deyolged samples, 420 proteins were found in the cleavage stages only, and 266 proteins were found commonly in all the deyolged samples. In contrast to the non-deyolged sample counterparts, a substantial proportion of unique proteins was found in either or both oblong and bud stages. In total, 465 proteins were present in both non-deyolged and deyolged samples across all the developmental stages (Figure 3B, Supplementary File 1).

### 2.3. Functional Annotations of the Proteome

In both non-deyolged and deyolged samples, the identified proteins were substantially involved in metabolic, ribosome, and biosynthesis of secondary metabolite and proteasome pathways,

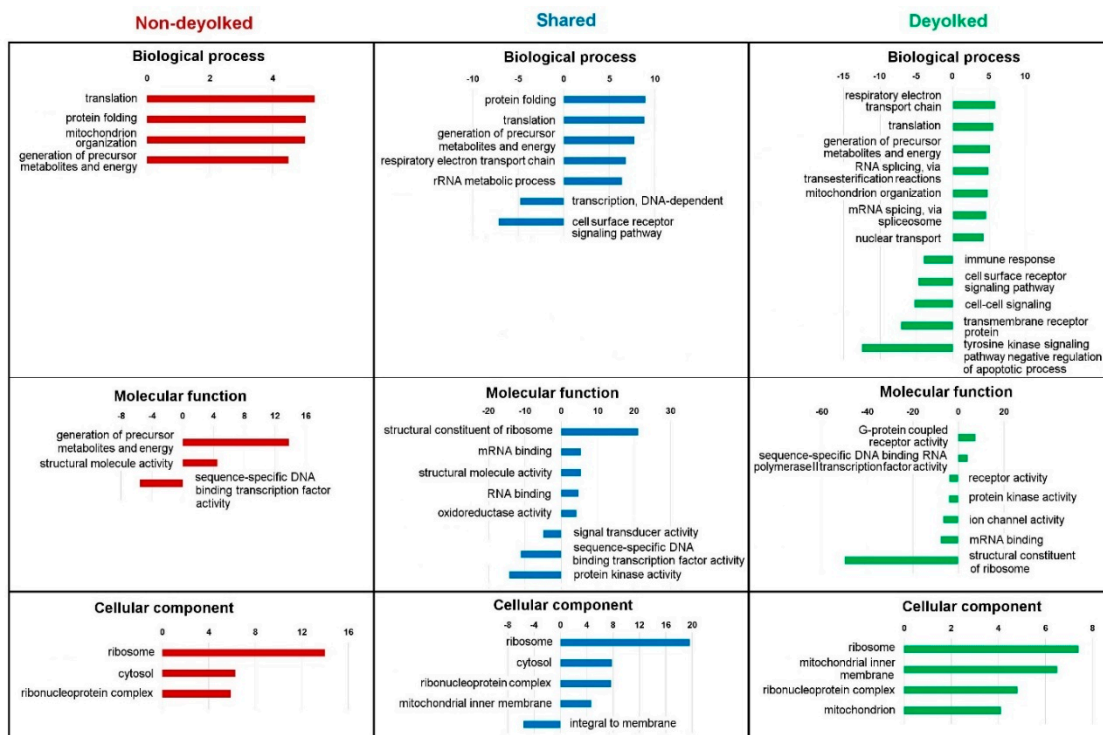
while enrichments specific to sampling protocol and/or developmental stage were found in certain pathways, such as in proteasome, RNA transport, or thermogenesis pathways (Table 2).

**Table 2.** Significant ( $p < 0.05$ ) pathways identified by Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of proteins from non-deyolked (ND) cleavage, oblong, and bud stage zebrafish embryos, and their deyolked (DY) counterparts. Numbers of proteins mapped to annotated pathways are given.

Pathway Name	ND-Cleavage Stage-Unique	ND- Oblong & Bud Stages Unique	ND-Common in All Stages	DY-Cleavage Stage Unique	DY-Oblong & Bud Stages Unique	DY-Common in All Stages	Shared Proteins
map01100 Metabolic pathways	22	23	6	66	62	87	69
map03010 Ribosome	27	11	2	21	19	11	52
map01110 Biosynthesis of secondary metabolites	12	11	3	21	25	27	32
map04714 Thermogenesis	2	6	1	28	6	37	20
map01200 Carbon metabolism	6	6	1	10	12	21	23
map04141 Protein processing in endoplasmic reticulum	4	0	0	5	14	11	20
map03050 Proteasome	14	4	1	1	15	0	14
map00010 Glycolysis / Gluconeogenesis	6	3	0	2	6	6	12
map00071 Fatty acid degradation	3	1	0	9	1	11	9
map01212 Fatty acid metabolism	8	3	0	7	0	10	8
map04530 Tight junction	5	5	2	6	13	2	7
map03013 RNA transport	7	0	0	3	22	1	6
map04110 Cell cycle	9	2	0	3	12	1	5
map04810 Regulation of actin cytoskeleton	5	1	1	6	12	1	4
map04144 Endocytosis	3	0	0	3	10	6	4
map00230 Purine metabolism	5	5	1	2	11	1	4
map03018 RNA degradation	3	0	0	0	8	1	4
map04210 Apoptosis	0	4	1	4	5	1	4
map00970 Aminoacyl-tRNA biosynthesis	2	1	1	8	12	1	3
map03030 DNA replication	4	3	0	0	14	0	3

Analysis of representation of the identified proteins annotated to functional Gene Ontology (GO) terms revealed multiple processes, functions, and components overrepresented and underrepresented in both non-deyolked and deyolked samples, with some of them specific to the developmental stage (Figure 4, Supplementary File 4).

To distinguish the effect of the extraction protocol (non-deyolked versus deyolked samples) from the biological features (natural representation of proteins at given developmental stage), we used functional annotations of proteins represented in both non-deyolked and deyolked samples from all the developmental stages as a filtering criterion. In this way, shared GO terms were established by: The same proteins identified in samples from both extraction methods (“Shared” dataset); different proteins in both datasets (“unique ND” and “unique DY” datasets) enriching the same terms; or partially the same and partially different proteins (“Shared” and “unique ND”, “Shared” and “unique DY”, and all the three datasets). Whereas, unique GO terms were established exclusively by proteins from either “unique ND” or “unique DY” datasets.



**Figure 4.** Significantly (false discovery rate (FDR) <0.05) enriched Gene Ontology (GO) terms from SLIM analysis for unique deyolked, unique non-deyolked, and shared proteins grouped by biological process, molecular function, and cellular component. Representation of GO terms containing a minimum 100 reference genes and a fold change  $\geq 4$  or  $\leq 4$  is given.

Clearly, the deyolking procedure yielded a considerable number of unique GO terms, which were not annotated with the proteins identified in the non-deyolked samples (Supplementary File 5). The most prominent, developmentally relevant examples included enrichment in: Cellular component organization, RNA splicing, DNA replication, intracellular transport, cell cycle, translational initiation, and mitochondrial organization, transport, and gene expression (biological process); ATP binding, GTP binding, NADH dehydrogenase activity, ribonucleoprotein complex binding, translation initiation factor activity, ribosome binding, and ligase activity (molecular function); chromosome, endoplasmic reticulum, Golgi-associated vesicle, polysome, spliceosomal complex, cytochrome, and mitochondrial ribosome, matrix, and respiratory chains I and II (cellular component). Similarly, underrepresentation in unique GO terms was developmentally relevant, and it included: Cell–cell signaling, chemical synaptic transmission, intracellular signal transduction, and immune response (biological process); DNA-binding transcription factor activity, transcription regulator activity, channel activity, G protein-coupled receptor activity, and kinase activity (molecular function); cell surface, extracellular region, and plasma membrane-bounded cell projection (cellular component). In contrast to the abundance of unique GO terms annotated with the “unique DY” dataset, there were very few unique GO terms associated with “unique ND” dataset, with the most notably enriched terms in molecular function: Carbohydrate binding and endopeptidase regulator activity (Supplementary Files 4 and 5).

A certain number of proteins was unique for a given developmental stage (that is, identified only in a single developmental stage), in both non-deyolked and deyolked samples. Interestingly, significantly enriched GO terms for these proteins were different for non-deyolked and deyolked samples, in all five developmental stages investigated (Supplementary File 6).

### 3. Discussion

The improved de yolking procedure resulted in a considerably high quantity of the extracted total protein (Figure 2A) We identified 2575 proteins in total. In the study by Link et al. [16], 57 proteins were found, but six of them were not identified, and two proteins had a duplicated ID. We manually retrieved these 50 IDs, and found that 47 (94%) proteins were present in our dataset. Two of the three proteins not found in our dataset were actually *Cyprinus carpio* and *Drosophila melanogaster* proteins, but their possible homologues in zebrafish were missing in our dataset as well. We used 3 times less embryos in our procedure (Figure 2B) than in the reference procedure [16]. Consequently, we were able to conduct the proteomics analysis of zygotic and cleavage stages of zebrafish for the first time. Most of the proteins identified in the cleavage stages were unique to these stages of development (786 out of 1494; Figure 2B). This indicates a massive dynamics of zebrafish developmental proteome. It needs to be noted that the protein sequence database, which we did not use for annotating MS data, does not include the sequences of micro-peptides. Therefore, we cannot determine whether the method is suitable for harvesting very small proteins and micro-peptides.

KEGG analysis showed that ribosome, biosynthesis of secondary metabolites, carbon metabolism, and proteasome pathways were detectable in all the samples (Table 2). Also, a number of GO terms were detected in both de yolked and non-de yolked datasets (Supplementary File 5). Nevertheless, we observed a substantial increase in the number of identified unique proteins in the de yolked samples as compared to the non-de yolked counterparts. Consequently, they enriched a number of developmentally relevant GO terms, such as the cell cycle, mitochondrial organization, and functions or translation initiation, which were not enriched in the non-de yolked samples (Supplementary File 5). These functional terms are essential for the proper growth and development of the early stage of embryos [25–28] Knowledge of developmentally relevant proteome will aid understanding the regulation of early embryonic development. The underrepresented GO terms in the de yolked samples were mainly related to cellular signaling, transcription, G protein-coupled receptor activity, and cell surface (Supplementary File 5). These terms were not found underrepresented in the non-de yolked samples. In contrast to the significant GO terms found uniquely in the de yolked samples, there were very few unique GO terms associated with “unique ND” dataset (Supplementary File 5); this indicates that the presence of many embryonic proteins is masked due to the high abundance of yolk.

The functional annotation of cleavage stage proteome is concordant with the canonical knowledge of the catabolism, cell cycle, subcellular organization, and the transcriptional quiescence of pre-MZT embryos [24,29]. Moreover, our data suggest active translation-related processes in the very early embryos. Since zygotic transcripts are not produced yet [8], maternally-provided mRNAs [30] were used to produce the translational machinery and perform the translation. Quantitative proteome analysis throughout the development would be needed to determine the extent of this process, though.

Although the dechoriation/de yolking procedure generally resulted in a substantial increase in the number of identified proteins, it also resulted in a loss of certain proteins as compared to the non-de yolked counterparts (Figure 3A), similarly to a study on 5 dpf zebrafish larvae [21]. Most of the previous proteomic studies did not address the problem of protein depletion due to the de yolking process, and they only used de yolked embryos for the analyses [16,21,22]. In the present study, approximately 30% of the proteins at cleavage stages and 12% at oblong and bud stages were not identified after the de yolking (Figure 3, Supplementary File 1). The GO analysis revealed that these lost proteins are involved in a number of biological processes (translation, protein folding, and mitochondrial organization), molecular functions (generation of precursor metabolites and energy), and cellular component (ribosome and cytosol; Supplementary File 4). Moreover, developmental stage-unique proteins enrich GO terms different for non-de yolked and de yolked samples, in all investigated developmental stages (Supplementary File 6). Altogether, our results suggest that de yolked and non-de yolked samples should be analyzed in parallel to extract a reliable information on the proteome in embryonic development.



## 4. Materials and Methods

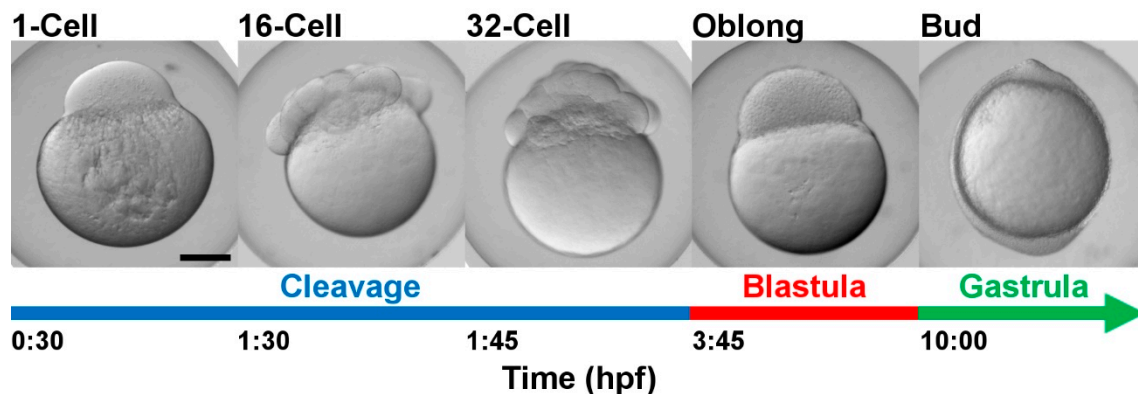
### 4.1. Fish

The samples were collected at the zebrafish facility of the Nord University, Bodo, Norway. The experimental process and husbandry were performed in agreement with the Norwegian Regulation on Animal Experimentation (The Norwegian Animal Protection Act, No. 73 of 20 December 1974). This was certified by the National Animal Research Authority, Norway, General License for Fish Maintenance and Breeding no. 17.

The maintenance of zebrafish was done using an Aquatic Habitats recirculating system (Pentair, Apopka, FL, USA) and following established protocols [31]. The fish were fed newly hatched *Artemia* sp. nauplii (Pentair) and SDS zebrafish-specific diet (Special Diet Services, Essex, UK) according to the manufacturers' instruction. The zebrafish used in the experiment were from the AB line.

### 4.2. Sample Collection

Embryos originated from natural spawning and were collected at five developmental stages (Figure 5). Embryo development was monitored and staged according to Kimmel et al. [32]. For each developmental stage, embryo batches were divided into two variants: Non-deyolged and deyolged. The non-deyolged (intact) embryos were promptly snap-frozen in liquid nitrogen and subsequently stored at  $-80^{\circ}\text{C}$ . The deyolged embryo variants went through the process of dechoriation (removal of chorion) and deyolking. Additionally, the 1-cell (0.5 hpf) and high-stage (3.3 hpf) embryos were collected to compare our deyolking protocol with that by [16].



**Figure 5.** Developmental stages of zebrafish embryos sampled in the present study. hpf = hours post-fertilization at  $28.5^{\circ}\text{C}$ .

### 4.3. Dechoriation and Deyolking

Embryos were placed in a Petri dish in phosphate-buffered saline (PBS) supplemented with  $1.0\text{ mg/mL}$  Pronase (Sigma Aldrich, St. Louis, MO, USA) [31]. The enzymatic digestion of chorion was performed for 5 min at  $37^{\circ}\text{C}$  with gentle shaking. Embryos were washed a minimum of 5 times with PBS or until all visible chorion fragments were removed.

The dechorionated embryos were processed using our modified protocol with 3-step deyolking and a single wash. The embryos were transferred to  $1.5\text{ mL}$  Eppendorf tubes containing  $1.0\text{ mL}$  of deyolking buffer ( $55\text{ mM NaCl}$ ,  $3.6\text{ mM KCl}$ , and  $1.25\text{ mM NaHCO}_3$ ) and were mechanically disrupted by pipetting repeatedly through a  $100\text{ }\mu\text{L}$  tip. The content was gently mixed by inverting the tube several times before centrifugation at  $13,000\text{ RPM}$  for 1 min at  $4^{\circ}\text{C}$ . The supernatant containing the yolk was discarded, and the pellet was re-suspended with the deyolking buffer, vortexed, and centrifuged as above. The procedure was repeated two times. After this, the pellet was re-suspended with  $10\text{ mM Tris-HCl}$  ( $\text{pH } 7.4$ ), vortexed, and centrifuged as above. The supernatant was discarded and

the pellet (deyolked embryos) was snap-frozen in liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$ . Additionally, for comparison of our protocol with that of [16], the dechorionated embryos at 1-cell and high stage were subjected to two types of deyolking protocols reported by Link et al. [16]: (1) 1-step deyolking, and (2) 1-step deyolking with two additional wash steps.

#### 4.4. Protein Extraction

Both intact (non-deyolked) and deyolked embryo samples were lysed by adding 100  $\mu\text{L}$  of sodium dodecyl sulphate (SDS) lysis buffer (1% SDS; Sigma-Aldrich, St. Louis, MO, USA), 0.5 M triethylammonium bicarbonate buffer pH 8.5 (TEAB; Sigma Aldrich), and 1  $\times$  Protease Inhibitor cocktail (Thermo Scientific, Rockford, IL, USA). The tubes were vortexed and incubated at  $90\text{ }^{\circ}\text{C}$  for 30 min, then cooled on crushed ice for 5 min. The lysed material was centrifuged at 13,000 RPM for 20 min at  $4\text{ }^{\circ}\text{C}$ . The supernatant, containing the proteins, was collected and transferred to a new Eppendorf tube. The total protein concentration was quantified using a Qubit<sup>®</sup> 3.0 Fluorometer (Invitrogen, Eugene, OR, USA) and the Qubit<sup>™</sup> Protein Assay Kit (Invitrogen) according to the manufacturer's instructions. After the quantification, the samples were freeze-dried (VirTis BenchTop<sup>™</sup> K, Warminster, USA) at  $-80\text{ }^{\circ}\text{C}$  for 18 h before being shipped to the Department of Biological Sciences, National University of Singapore for proteomics analysis.

#### 4.5. Polyacrylamide Gel Electrophoresis

One-dimensional gel electrophoresis was performed to check the efficiency of deyolking protocol, as well as to compare the efficiency of our protocol with the previous ones. Approximately equal concentrations of proteins from each sample were supplemented with 2 $\times$  SDS loading dye. The samples were denatured by incubation at  $95\text{ }^{\circ}\text{C}$  for 10 min and then the proteins were separated by SDS gel electrophoresis (4%–20% Mini-PROTEAN<sup>®</sup> TGX<sup>™</sup> Precast Protein Gels, Bio-Rad, Hercules, California, USA) in SDS running buffer for 1 h. Afterwards, the gel was washed with deionized water for 10 min. The gel was stained with Coomassie Brilliant Blue R-250, Bio-Rad) for 20 min, and de-stained with de-staining solution (40% methanol + 10% acetic acid) overnight at room temperature.

#### 4.6. Tube-Gel Digestion and Sample Clean up

For each sample, 30  $\mu\text{g}$  of proteins were used for downstream proteomics analyses. The samples were polymerized in a 10% polyacrylamide gel containing 4% SDS and subsequently fixed with a fixing reagent (50% methanol, 12% acetic acid) for 30 min at room temperature. The gel was cut into small pieces (1  $\text{mm}^3$ ) before being washed three times with 50 mM TEAB/50 % acetonitrile (*v/v*) and dehydrated with 100% acetonitrile. Next, samples were reduced using 5 mM Tris(2-carboxyethyl) phosphine (TCEP) at  $57\text{ }^{\circ}\text{C}$  for 60 min, followed by alkylation with 10 mM methyl methanethiosulfonate (MMTS) for 60 min at room temperature with occasional vortexing. The gel pieces were washed in 500  $\mu\text{L}$  of 50 mM TEAB, dehydrated in 500  $\mu\text{L}$  acetonitrile, and re-hydrated with 500  $\mu\text{L}$  of 50 mM TEAB. The final dehydration step was performed using 100  $\mu\text{L}$  acetonitrile. Trypsinization (1.5  $\mu\text{g}$  trypsin) was performed at  $37\text{ }^{\circ}\text{C}$  for 16 h. The digested peptides were centrifuged at  $6000\times g$  for 10 min to collect the supernatant and stored at  $-20\text{ }^{\circ}\text{C}$  (protocol modified from [17]). The samples were lyophilized and 30  $\mu\text{L}$  of the dissolution buffer (0.5 M TEAB, pH 8.5) was added to each sample.

#### 4.7. 1D LC-MS/MS Analysis

The separation of peptides was performed with an Eksigent nanoLC Ultra and ChiPLC-nanoflex (Eksigent, Dublin, CA, USA) in Trap-Elute configuration. The samples were desalted with a Sep-Pak tC 18  $\mu\text{L}$  Elution Plate (Waters, Milford, MA, USA), and reconstituted using 20  $\mu\text{L}$  of 2% acetonitrile and 0.05% formic acid. Five microliters ( $\mu\text{L}$ ) of each sample was loaded on a 200  $\mu\text{m} \times 0.5\text{ mm}$  trap column and eluted on a 75  $\mu\text{m} \times 15\text{ cm}$  analytical column (ChromXP C18-CL, 3  $\mu\text{m}$ ). A gradient formed by mobile phase A (2% acetonitrile, 0.1% formic acid) and mobile phase B (98% acetonitrile, 0.1% formic acid) was used to separate the sample content at a 0.3  $\mu\text{L}/\text{min}$  flow rate. The following gradient elution

was used for peptide separation: 0–5% of mobile phase B in 1 min, 5–12% of mobile phase B in 15 min, 12–30% of mobile phase B in 104 min, 30–90% of mobile phase B in 2 min, 90–90% in 7 min, 90–5% in 3 min and held at 5% of mobile phase B for 13 min (protocol modified from [33]).

#### 4.8. Protein Identification and Quantification

Peptide identification was carried out with the ProteinPilot 5.0 software Revision 4769 (AB SCIEX) using the Paragon database search algorithm (5.0.0.0.4767) and the integrated false discovery rate (FDR) analysis function. The data were searched against protein sequence databases downloaded from UniProt on May 2018 (total 119,356 entries). The MS/MS spectra obtained were searched using the following user-defined search parameters: Sample Type: Identification; Cysteine Alkylation: MMTS; Digestion: Trypsin; Instrument: TripleTOF5600; Special Factors: None; Species: *None*; ID Focus: Biological Modification; Database for 2018\_May\_uniprot-zebrafish.fasta; Search Effort: Thorough; FDR Analysis: Yes. The MS/MS spectra were searched against a decoy database to estimate FDR for peptide identification. The decoy database consisted of reversed protein sequences from the UniProt zebrafish database. FDR analysis was performed on the dataset and peptides identified with a confidence interval  $\geq 95\%$  were taken into account.

#### 4.9. KEGG and Gene Ontology (GO) Functional Pathways Analysis

To analyse functional pathways associated with protein identified from de-yolked and non-de-yolked samples, KEGG analysis was performed. The FASTA files were submitted to online server “KAAS - KEGG Automatic Annotation Server” (<https://www.genome.jp/kegg/kaas/>) in order to get KEGG Orthology (KO) assignments [34]. To map KEGG pathways, the obtained KO numbers were submitted to KEGG mapper web server ([http://www.genome.jp/kegg/tool/map\\_pathway2.html](http://www.genome.jp/kegg/tool/map_pathway2.html)) [35].

GO annotation results and pathway of differentially expressed proteins in pairwise comparisons were obtained using Panther (Panther14.0, 2018\_04) [36]. The web conversion tool (<https://biobdnet-abcc.ncifcrf.gov>) was used to convert unmapped UniProt Accession IDs to ZFIN ID. The web tool Biomart was used to convert unmapped ZFIN IDs to Gene stable ID and to manually identify the unmapped IDs by gene names [37]. UniProt was used to identify protein IDs discontinued (deleted) in the 2018\_11 release [38].

## 5. Conclusions

We established an effective de-yolking procedure for the proteome analysis of the early stages of zebrafish embryos. Elimination of most of the yolk from early stages of embryos significantly enhanced the identification of cellular proteins with LC–MS-based shotgun proteomics analysis. The improved protocol is applicable to low-input material, enabling investigation of the earliest stages of development. Also, we demonstrated that the de-yolking procedure results in the depletion of certain parts of the proteome that can be important in embryonic development. Thus, we suggest that both de-yolked and non-de-yolked samples should be processed in parallel to ensure a reliable coverage of the proteome during the embryogenesis. Our de-yolking procedure will improve both qualitative and quantitative proteome analyses throughout embryonic development of polylecithal animals, such as fish, amphibians, reptiles, and birds.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/24/6359/s1>. Supplementary File S1. Complete list of proteins identified in the study. DY, de-yolked samples; ND, non-de-yolked samples; 1-cell, 16-cell, 32-cell, oblong, and bud stages of development were sampled; Supplementary File S2. Relative quantification of vitellogenin detected in non-de-yolked and de-yolked samples; Supplementary File S3. Relative quantification of non-vitellogenin proteins shared between non-de-yolked and de-yolked samples; Supplementary File S4. Gene ontology analyses of DY- de-yolked samples; ND- non-de-yolked; SH- shared samples; 1-cell, 16-cell, 32-cell, oblong and bud stages of embryos; Supplementary File S5. Gene ontology terms significantly overrepresented and underrepresented: Unique for the protein extraction method (non-de-yolked or de-yolked), or common for the two methods; Supplementary File S6. List and relative quantification of proteins unique for each developmental stage, and Gene Ontology analysis of terms overrepresented and underrepresented.

**Author Contributions:** Conceptualization, I.B., Q.L., and S.D.J.; methodology, K.P., C.P., and P.P.D.; validation, K.P., P.P.D., T.K.L., Q.L., and I.B.; formal analysis, K.P., T.K.L., P.P.D., Q.L., and I.B.; investigation, K.P., T.K.L., and P.P.D.; resources, I.B. and Q.L.; writing—original draft preparation, K.P.; writing—review and editing, all; visualization, K.P., C.P., and I.B.; supervision, I.B., S.D.J., and Q.L.; project administration, I.B. and Q.L.; funding acquisition, I.B. and Q.L.

**Funding:** This research was funded by the Research Council of Norway, *InnControl* project (grant #275786, granted to I.B.), Nord University (scholarship for K.P.), and Protein and Proteomics Centre, Department of Biological Sciences, National University of Singapore (granted to Q.L.).

**Acknowledgments:** The authors would like to thank Felix Müller, Research Fellow at Nord University, who helped in embryo collections. K.P. wishes to thank Nord University for funding the PhD scholarship and travel grant.

**Conflicts of Interest:** The authors declare no competing interests. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

ND	Non-deyolked
DY	Deyolked
SH	Shared
iTRAQ	Isobaric tag for relative and absolute quantitation
LC–MS/MS	Liquid chromatography–tandem mass spectrometry
1D LC–MS/MS	One-dimension liquid chromatography–tandem mass spectrometry
MZT	Maternal-to-zygotic transition
PBS	Phosphate-buffered saline
SDS	Sodium dodecyl sulfate
TCEP	Tris-(2-carboxyethyl) phosphine
MMTS	Methyl methane-thiosulfonate
TEAB	Triethylammonium bicarbonate
ACN	Acetonitrile
FDR	False discovery rate
COG	Clusters of orthologous groups
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes

## References

1. Amsterdam, A.; Hopkins, N. Mutagenesis strategies in zebrafish for identifying genes involved in development and disease. *Trends Genet.* **2006**, *22*, 473–478. [[CrossRef](#)] [[PubMed](#)]
2. Brownlie, A.; Donovan, A.; Pratt, S.J.; Paw, B.H.; Oates, A.C.; Brugnara, C.; Witkowska, H.E.; Sassa, S.; Zon, L.I. Positional cloning of the zebrafish sauternes gene: A model for congenital sideroblastic anaemia. *Nat. Genet.* **1998**, *20*, 244. [[CrossRef](#)] [[PubMed](#)]
3. Grosser, T.; Yusuff, S.; Cheskis, E.; Pack, M.A.; FitzGerald, G.A. Developmental expression of functional cyclooxygenases in zebrafish. *Proc. Natl. Acad. Sci. USA.* **2002**, *99*, 8418–8423. [[CrossRef](#)] [[PubMed](#)]
4. Gündel, U.; Benndorf, D.; von Bergen, M.; Altenburger, R.; Küster, E. Vitellogenin cleavage products as indicators for toxic stress in zebra fish embryos: A proteomic approach. *Proteom.* **2007**, *7*, 4541–4554. [[CrossRef](#)] [[PubMed](#)]
5. Hanisch, K.; Küster, E.; Altenburger, R.; Gündel, U. Proteomic signatures of the zebrafish (*Danio rerio*) embryo: Sensitivity and specificity in toxicity assessment of chemicals. *Int. J. Proteom.* **2010**. [[CrossRef](#)] [[PubMed](#)]
6. Howe, K.; Clark, M.D.; Torroja, C.F.; Tarrance, J.; Berthelot, C.; Muffato, M.; Collins, J.E.; Humphray, S.; McLaren, K.; Matthews, L. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **2013**, *496*, 498–503. [[CrossRef](#)]
7. Aanes, H.; Winata, C.L.; Lin, C.H.; Chen, J.P.; Srinivasan, K.G.; Lee, S.G.; Lim, A.Y.; Hajan, H.S.; Collas, P.; Bourque, G. Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition. *Genome. Res.* **2011**, *21*, 1328–1338. [[CrossRef](#)]

8. Heyn, P.; Kircher, M.; Dahl, A.; Kelso, J.; Tomancak, P.; Kalinka, A.T.; Neugebauer, K.M. The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Rep.* **2014**, *6*, 285–292. [[CrossRef](#)]
9. Mehjabin, R.; Xiong, L.; Huang, R.; Yang, C.; Chen, G.; He, L.; Liao, L.; Zhu, Z.; Wang, Y. Full-Length Transcriptome Sequencing and the Discovery of New Transcripts in the Unfertilized Eggs of Zebrafish (*Danio rerio*). *G3* **2019**. [[CrossRef](#)]
10. Nudelman, G.; Frasca, A.; Kent, B.; Sadler, K.C.; Sealfon, S.C.; Walsh, M.J.; Zaslavsky, E. High resolution annotation of zebrafish transcriptome using long-read sequencing. *Genome. Res.* **2018**, *28*, 1415–1425. [[CrossRef](#)]
11. White, R.J.; Collins, J.E.; Sealy, I.M.; Wali, N.; Dooley, C.M.; Digby, Z.; Stemple, D.L.; Murphy, D.N.; Billis, K.; Hourlier, T. A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* **2017**. [[CrossRef](#)] [[PubMed](#)]
12. De Souza, A.G.; Mac Cormack, T.J.; Wang, N.; Li, L.; Goss, G.G. Large-scale proteome profile of the zebrafish (*Danio rerio*) gill for physiological and biomarker discovery studies. *Zebrafish* **2009**, *6*, 229–238. [[CrossRef](#)] [[PubMed](#)]
13. Groh, K.J.; Nesatyy, V.J.; Segner, H.; Eggen, R.I.; Suter, M.J.-F. Global proteomics analysis of testis and ovary in adult zebrafish (*Danio rerio*). *Fish Physiol. Biochem.* **2011**, *37*, 619–647. [[CrossRef](#)]
14. Singh, S.K.; Sundaram, C.S.; Shanbhag, S.; Idris, M.M. Proteomic profile of zebrafish brain based on two-dimensional gel electrophoresis matrix-assisted laser desorption/ionization MS/MS analysis. *Zebrafish* **2010**, *7*, 169–177. [[CrossRef](#)] [[PubMed](#)]
15. Zhang, J.; Lanham, K.A.; Peterson, R.E.; Heideman, W.; Li, L. Characterization of the adult zebrafish cardiac proteome using online pH gradient strong cation exchange-RP 2D LC coupled with ESI MS/MS. *J. Sep. Sci.* **2010**, *33*, 1462–1471. [[CrossRef](#)] [[PubMed](#)]
16. Link, V.; Shevchenko, A.; Heisenberg, C.-P. Proteomics of early zebrafish embryos. *Bmc. Dev. Biol.* **2006**. [[CrossRef](#)] [[PubMed](#)]
17. Lu, W.; Yin, X.; Liu, X.; Yan, G.; Yang, P. Response of peptide intensity to concentration in ESI-MS-based proteome. *Sci. China Chem.* **2014**, *57*, 686–694. [[CrossRef](#)]
18. Vaidyanathan, S.; Kell, D.B.; Goodacre, R. Selective detection of proteins in mixtures using electrospray ionization mass spectrometry: Influence of instrumental settings and implications for proteomics. *Anal. Chem.* **2004**, *76*, 5024–5032. [[CrossRef](#)]
19. Lucitt, M.B.; Price, T.S.; Pizarro, A.; Wu, W.; Yocum, A.K.; Seiler, C.; Pack, M.A.; Blair, I.A.; FitzGerald, G.A.; Grosser, T. Analysis of the zebrafish proteome during embryonic development. *Mol. Cell Proteom.* **2008**, *7*, 981–994. [[CrossRef](#)]
20. Rahlouni, F.; Szarka, S.; Shulaev, V.; Prokai, L. A Survey of the Impact of Deyolking on Biological Processes Covered by Shotgun Proteomic Analyses of Zebrafish Embryos. *Zebrafish* **2015**, *12*, 398–407. [[CrossRef](#)]
21. Tay, T.L.; Lin, Q.; Seow, T.K.; Tan, K.H.; Hew, C.L.; Gong, Z. Proteomic analysis of protein profiles during early development of the zebrafish, *Danio rerio*. *Proteomics* **2006**, *6*, 3176–3188. [[CrossRef](#)] [[PubMed](#)]
22. Shaik, A.A.; Wee, S.; Li, R.H.X.; Li, Z.; Carney, T.J.; Mathavan, S.; Gunaratne, J. Functional mapping of the zebrafish early embryo proteome and transcriptome. *J. Proteome Res.* **2014**, *13*, 5536–5550. [[CrossRef](#)] [[PubMed](#)]
23. Lößner, C.; Wee, S.; Ler, S.G.; Li, R.H.; Carney, T.; Blackstock, W.; Gunaratne, J. Expanding the zebrafish embryo proteome using multiple fractionation approaches and tandem mass spectrometry. *Proteomics* **2012**, *12*, 1879–1882. [[CrossRef](#)] [[PubMed](#)]
24. Tadros, W.; Lipshitz, H.D. The maternal-to-zygotic transition: A play in two acts. *Development* **2009**, *136*, 3033–3042. [[CrossRef](#)]
25. Gong, Y.; Mo, C.; Fraser, S.E. Planar cell polarity signalling controls cell division orientation during zebrafish gastrulation. *Nature* **2004**, *430*, 689. [[CrossRef](#)]
26. Sugiyama, M.; Sakaue-Sawano, A.; Iimura, T.; Fukami, K.; Kitaguchi, T.; Kawakami, K.; Okamoto, H.; Higashijima, S.-i.; Miyawaki, A. Illuminating cell-cycle progression in the developing zebrafish embryo. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 20812–20817. [[CrossRef](#)]
27. Artuso, L.; Romano, A.; Verri, T.; Domenichini, A.; Argenton, F.; Santorelli, F.M.; Petruzzella, V. Mitochondrial DNA metabolism in early development of zebrafish (*Danio rerio*). *Biochim. Biophys. Acta.* **2012**, *1817*, 1002–1011. [[CrossRef](#)]

28. Sun, J.; Yan, L.; Shen, W.; Meng, A. Maternal Ybx1 safeguards zebrafish oocyte maturation and maternal-to-zygotic transition by repressing global translation. *Development* **2018**. [[CrossRef](#)]
29. Lee, M.T.; Bonneau, A.R.; Giraldez, A.J. Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 581–613. [[CrossRef](#)]
30. Lubzens, E.; Bobe, J.; Young, G.; Sullivan, C.V. Maternal investment in fish oocytes and eggs: The molecular cargo and its contributions to fertility and early development. *Aquaculture* **2017**, *472*, 107–143. [[CrossRef](#)]
31. Westerfield, M. A guide for the laboratory use of zebrafish (*Danio rerio*). *The Zebrafish Book* **2000**, *4*.
32. Kimmel, C.B.; Ballard, W.W.; Kimmel, S.R.; Ullmann, B.; Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **1995**, *203*, 253–310. [[CrossRef](#)] [[PubMed](#)]
33. Suriyanarayanan, T.; Qingsong, L.; Kwang, L.T.; Mun, L.Y.; Seneviratne, C.J. Quantitative proteomics of strong and weak biofilm formers of *Enterococcus faecalis* reveals novel regulators of biofilm formation. *Mol. Cell Proteom.* **2018**, *17*, 643–654. [[CrossRef](#)] [[PubMed](#)]
34. Moriya, Y.; Itoh, M.; Okuda, S.; Yoshizawa, A.C.; Kanehisa, M. KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic. Acids Res.* **2007**, *35*, W182–W185. [[CrossRef](#)] [[PubMed](#)]
35. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic. Acids Res.* **2015**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]
36. Mi, H.; Huang, X.; Muruganujan, A.; Tang, H.; Mills, C.; Kang, D.; Thomas, P.D. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic. Acids Res.* **2016**, *45*, D183–D189. [[CrossRef](#)]
37. Zerbino, D.R.; Achuthan, P.; Akanni, W.; Amode, M.R.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C.G.; et al. Ensembl 2018. *Nucleic. Acids Res.* **2017**, *46*, D754–D761. [[CrossRef](#)]
38. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M. UniProt: The universal protein knowledgebase. *Nucleic. Acids Res.* **2004**, *32*, D115–D119. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).