

Mid-Miocene formation of the Yangtze river-lake system revealed by ancestral spawning reconstruction of endemic cyprinids

Yeke Wang^{1,2, #}, Feng Chen^{1,2, #}, Jun He^{1, #}, Liang Chen^{1,3}, Ge Xue^{1,2}, Yan Zhao^{1,2}, Yanghui Peng^{1,2}, Jun Chen^{1, *} & Ping Xie^{1,4, *}

1. Donghu Experimental Station of Lake Ecosystems, State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, 430072, China
2. University of Chinese Academy of Sciences, Beijing, 100049, China
3. State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Faculty of Water Resources and Hydroelectric Engineering, Xi'an University of Technology, Xi'an, 710048, China
4. Institute of Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environment, Yunnan University, Kunming 650500, P. R. China

These authors contributed equally to this work.

* Authors for correspondence:

Ping Xie (Dr.), E-mail: xieping@ihb.ac.cn

Jun Chen (Dr.), E-mail: chenjun@ihb.ac.cn

Abstract

The Yangtze River is cross-linked with numerous lakes within its floodplain and is a worldwide biodiversity hotspot. There is no evidence indicating when this unique river-lake system developed. The endemic East Asian cyprinid clade has evolved diverse spawning adaptations to different flow conditions. Our ancestral egg-type reconstruction showed an ancestral state of adhesive eggs and later demersal eggs origination (both stream adaptations). Semi-buoyant eggs emerged ~18 Mya as a fast-flowing river adaptation, with increased hydration via three yolk protein degradation pathways, ion transport pathways and egg envelope permeability transition pores. Adhesive eggs evolved secondarily ~14 Mya with the egg envelope increasing to four layers and an adhesive layer, along with an increase in adhesiveness via microfilament/adhesive-related protein crosslinking and enhanced glycosaminoglycan biosynthesis, improving adherence to submerged lake plants, indicating that the cross-linked river-lake system formed in the mid-Miocene. This study provides a unique biological evidence for large-scale water system evolution.

Introduction

Cyprinidae (Teleostei: Cypriniformes), the largest fish family, comprises approximately 367 genera and 3006 species that are distributed widely in Eurasia, Africa, and North America¹. The East Asian endemic cyprinids (Teleostei: Ostariophysi: Cypriniformes) are a natural assemblage of freshwater fishes found in eastern Asia, especially in China. This group includes approximately 29 genera and 81 species, and the burst of its diversification may have been closely linked to the uplift of the Qinghai-Tibetan Plateau (QTP) and East Asian monsoon intensification². The QTP (mean elevation over 4,000 m a.s.l.) acts as a strong heat source in summer, generating upward airflow motion over its eastern flank, which combined with large amounts of moisture from the tropics, result in strong monsoons and a wet climate in East Asia^{3, 4}. Large amounts of water vapor from the Pacific and Indian Oceans are carried onto the continent in summer by the East Asian summer monsoon⁵; this phenomenon gave rise to a cross-linked river-lake system in East Asia and

led to the development of the Yangtze River in response to the surface uplift of southeastern Tibet⁶. However, it is still unclear when the cross-linked river-lake system of the Yangtze River developed, although the date at which the eastward flow of the Yangtze River began has been debated for centuries.

The development of the cross-linked river-lake system greatly increased the ecological and phenotypic diversity of the East Asian cyprinids within a short time, enabling these fish to exploit areal lakes and river drainages and leading to rapid diversification⁷. Some cyprinids, such as Hypophthalmichthyinae and Squaliobarbus Tribe, lay semi-buoyant eggs in the rivers during the rainy season, from which the hatched fish travel to lakes to mature, while Cultrinae and Xenocyprinae lay adhesive eggs that adhere mainly to aquatic plants in the lakes. The Opsariichthys Tribe produces demersal eggs in slow-flowing streams. The emergence of semi-buoyant eggs and adhesive eggs in East Asian endemic cyprinids might serve as an indicator of the formation of the cross-linked river-lake system under the strong monsoon in East Asia⁸. Few studies have investigated the structural and functional development of the egg envelope in East Asian endemic cyprinids via ancestral state reconstruction analysis^{9, 10}.

Semi-buoyant eggs are an ideal adaptation to the flowing waters of large rivers, such as the Yangtze River in East Asia. They hydrate after they are produced, absorbing water and expanding briefly to form a large perivitelline space. As they are slightly heavier than water, the semi-buoyant eggs must increase their buoyancy to become suspended in the river at a flow rate of ca. 0.63-1.83 m/s¹¹. Semi-buoyant eggs are beneficial for dispersal, the prevention of sinking under still water conditions, the completion of hatching and development into young fishes capable of swimming¹². Marine pelagic eggs take up large amounts of water during ovary maturation and float in seawater after ovulation, which is reported to be regulated by free amino acids (FAAs) and ions determining osmolality^{13, 14}. The hydration of semi-buoyant eggs in freshwater fish occurs in vitro in a hypotonic environment, and the mechanisms involved are still poorly understood. On the other hand, adhesive eggs show an increase in adhesiveness after fertilization to adhere to aquatic plants, gravel or other hard surfaces, which is an adaptation to the lake environment. The degree of egg adhesiveness and the structure of the egg envelope might reflect adaptations to variable environmental conditions¹⁵. Glycoproteins and mucopolysaccharides (glycosaminoglycan (GAG)) are reported to participate in the development of egg stickiness^{16, 17, 18}. In addition, the fish egg envelope (also referred to as the chorion or zona radiata) varies in thickness, structure and the number of layers between or within species¹⁹; the envelope consists of 2–4 ZP proteins and can protect the embryo from physical and environmental stressors^{20, 21}. However, the molecular mechanism underlying the adhesiveness of freshwater eggs remains unknown.

Knowledge about typical characteristics during evolution is conducive to understanding how organisms adapt to new environments and the origins of biodiversity^{22, 23}. Comparisons between individuals with different reproductive modes within the same species facilitate the identification of changes specifically related to the evolution of egg type²⁴ and can avoid interspecific interference in metabolomic or proteomic analyses²⁵. The topmouth culter (*Culter alburnus*) has two ecotypes, one (*C. alburnus*-A) lays adhesive eggs in lakes such as Lake Taihu, while the other (*C. alburnus*-B) that spawns semi-buoyant eggs in fast-flowing rivers such as the Yangtze River²⁶. Therefore, *C. alburnus* is an ideal model for investigating the adaptative evolution of egg types among East Asian endemic cyprinids.

Here, based on ancestral state reconstruction, we attempt to reveal the Miocene history of egg-type evolution in East Asian endemic cyprinids, thus providing biological evidence indicating the timing of the formation of the cross-linked river-lake system in the middle and lower reaches of the Yangtze River. The molecular and physiological mechanisms underlying egg type evolution (from semi-buoyant eggs to adhesive eggs) were investigated by multiomics analyses and biochemical experiments in two ecotypes of *C. alburnus*.

Results

Ancestral state reconstruction for the egg types of East Asian endemic cyprinids

Maximum likelihood (ML) analyses incorporating mitochondrial genome data from 63 species (3 of which were outgroups) from GenBank resulted in highly resolved and well-supported phylogenetic trees of East Asian endemic group of cyprinids (Supplementary Figure S1 and Supplementary Data S1). The cladogram resulting from ML analyses was used for egg type state reconstruction among the East Asian endemic cyprinids. The reconstruction revealed that adhesive eggs were the ancestral state in East Asian cyprinids (ancestral adhesive egg, Node 1, Fig. 1a). Species from the Aphyocypris and Opsariichthys Tribes evolved to lay demersal eggs (Node 2, Fig. 1a). Simultaneously, the Hypophthalmichthyinae and Squaliobarbus Tribes evolved to spawn semi-buoyant eggs (Node 3, Fig. 1a). Moreover, species from Cultrinae and Xenocyprinae underwent further evolution to lay adhesive egg (Node 4, Fig. 1a). We also used the relaxed molecular clock method to estimate divergence times. The topology resulting from ML analysis was used for dating phylogenetic events due to relatively well-resolved phylogenetic relationships. Adhesive eggs (ancestral state) produced by the Metzia Tribe appeared approximately 28 Mya (with a 95% confidence interval of 23.7-33.7 Mya) (Node 1, Fig. 1a). Demersal eggs produced by the Opsariichthys Tribe and Aphyocypris Tribe appeared approximately 22 Mya (with a 95% confidence interval of 18.0-27.1 Mya) (Node 2, Fig. 1a). Semi-buoyant egg produced by Hypophthalmichthyinae and Squaliobarbus Tribe appeared approximately 18 Mya (with a 95% confidence interval of 15.6-21.3Mya) (Node 3, Fig. 1a). Adhesive eggs produced by Cultrinae and Xenocyprinae (resulting from secondary evolution) appeared approximately 14 Mya (with a 95% confidence interval of 11.1-17.4 Mya) (Node 4, Fig. 1a).

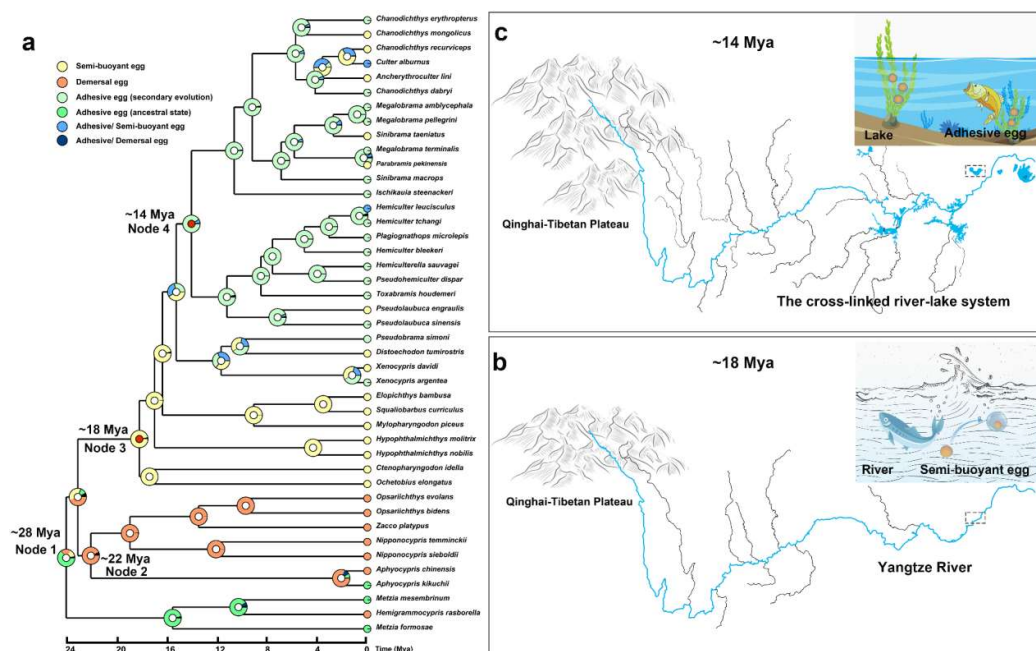


Fig. 1. (a) Chronogram of the East Asian endemic group of cyprinids, and reconstruction of ancestral character states for different egg types of the endemic East Asian cyprinid group. Divergence date estimates were based on the topology tree resulting from the maximum likelihood analyses. Pie charts represent the median posterior probability of each spawning type for the common ancestor. Yellow, semi-buoyant egg; orange, demersal egg; light green, adhesive egg (secondary evolution); dark green, adhesive egg (ancestral state); blue, adhesive/ semi-buoyant egg, dark blue, adhesive/ demersal egg. (b) Sketch map showing the initial formation of the drainage networks of the Yangtze River at approximately 18 Mya and the semi-buoyant egg, which absorbs a large amount of water and floats in the fast-flowing river. (c) Sketch map showing the formation of the cross-linked river-lake system in the middle and lower reaches of the Yangtze River at approximately 14 Mya and the adhesive egg, which adheres to aquatic plants.

Biochemical adaptation for hydration in East Asian endemic cyprinids with different egg types

Ovulated mature eggs and fertilized eggs at 0, 0.5 and 1 h postfertilization were subjected to the measurement of perivitelline space volumes and water contents. These eggs included the semi-buoyant eggs of six representative species (*C. alburnus*-B, *Hypophthalmichthys molitrix*, *H. nobilis*, *Ctenopharyngodon idella*, *Mylopharyngodon piceus*, and *Squaliobarbus curriculus*), the demersal eggs of two representative species (*Zacco platypus* and *Opsariichthys evolans*) and the adhesive eggs (secondary evolution) of three representative species (*C. alburnus*-A, *Megalobrama amblycephala*, and *Chanodichthys dabryi*) of the endemic East Asian cyprinid group. The water contents of the semi-buoyant eggs were obviously higher than those of demersal and adhesive eggs by 0.5 h and 1 h after fertilization (Fig. 2a), and the perivitelline space volume (V_m/V_e) of semi-buoyant eggs was greater than those of demersal and adhesive eggs by 0.5 h and 1 h after fertilization (Supplementary Figure S2). The results of an osmotic gradient experiment performed on eggs at 1 h after fertilization showed that the membrane diameter changes (1 h-FE MD/0 h-FE MD) observed in semi-buoyant eggs with an increase in the osmotic concentration of the solution

were significantly smaller than those in demersal eggs and adhesive eggs, indicating that osmotic pressure played a major role in regulating the hydration of semi-buoyant eggs (Supplementary Figure S3). Moreover, the total free amino acid (T-FAA) and ion contents of the fish eggs were examined in this study. After fertilization, the T-FAA contents of semi-buoyant eggs obviously increased with time compared with those of demersal eggs and adhesive eggs (Fig. 2b). The total contents of Na^+ , K^+ , Ca^{2+} , and Mg^{2+} were also markedly increased in semi-buoyant and demersal eggs but reduced in adhesive eggs (Fig. 2c). The effect of the T-FAA content on osmolality was considerably lower than that of the ion content in freshwater semi-buoyant eggs (Supplementary Table S1).

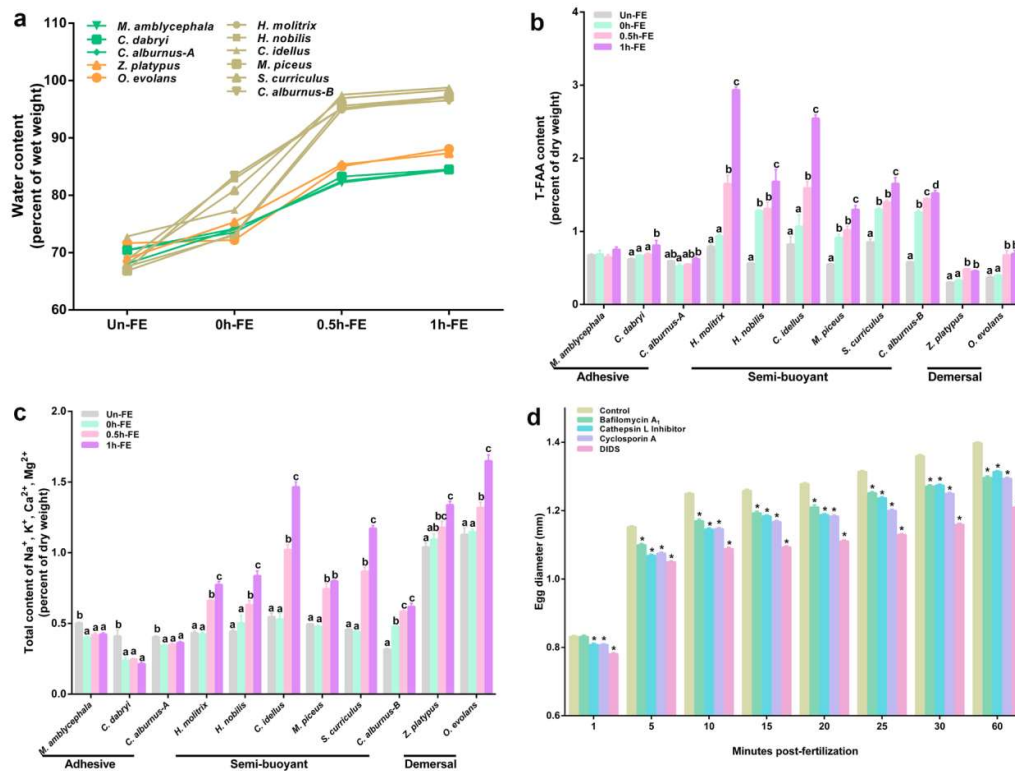


Fig. 2 The different egg types differ in their biochemical composition at the unfertilized (Un-FE), and 0, 0.5 and 1 h postfertilization (0 h-FE, 0.5 h-FE, and 1 h-FE). (a) The water contents of adhesive eggs (green lines), demersal eggs (orange line) and semi-buoyant eggs (yellow lines). (b) The T-FAA contents of three egg types, and (c) The total contents of Na^+ , K^+ , Ca^{2+} , Mg^{2+} of three egg types. Values are mean \pm SEM from three to five separate experiments. Different lower-case letters indicate significant differences between four times ($p < 0.05$, one-way analysis of variance). T-FAA, total free amino acid. (d) The membrane diameter of zebrafish at postfertilization of 1, 5, 10, 15, 20, 25, 30 and 60 minutes when exposed by 0.05% DMSO (control, banana yellow), 100 nM bafilomycin A1 (sea green), 1.5 μM cathepsin L inhibitor (cyan), 5 μM cyclosporine A (purple) 200 μM DIDS (pink). Values are the means \pm SEMs from three separate experiments (30-40 per group in each experiment). Asterisks indicate significant differences between the control and treatment groups ($p < 0.05$, Student's t-test). DMSO, dimethyl sulfoxide; DIDS, 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid disodium salt.

Convergent evolution of the morphological characteristics of the egg envelope

Ovulated mature eggs and fertilized eggs at 1 h postfertilization were subjected to microstructure analysis by transmission electron microscopy (TEM). These analyses included the semi-buoyant eggs of six representative species (*C. alburnus*-B, *H. molitrix*, *H. nobilis*, *C. idellus*, *M. piceus*, and *S. curriculus*), the demersal eggs of two representative species (*Z. platypus* and *O. evolans*) and the adhesive eggs (secondary evolution) of three representative species (*C. alburnus*-A, *M. amblycephala*, and *C. dabryi*) of the East Asian endemic group of cyprinids (Fig.3a and 3b) and the ovulated mature adhesive eggs (ancestral state) of a representative species (*Metzia formosae*) (Supplementary Figure S4). We observed that the egg envelope of semi-buoyant eggs and the ancestral adhesive eggs was divided into three layers (1, 2, and 3) (Fig. 3b and Supplementary Figure S4). The egg envelope of the demersal eggs was divided into two layers (1 and 3), while the egg envelope of the adhesive egg (secondary evolution) was further divided into four layers (1, 2, 3 and 4) (Fig. 3b). The TEM results indicated that the adhesive eggs (secondary evolution) possessed an adhesive layer (AL) at 1 h after fertilization, while the semi-buoyant eggs and demersal eggs did not develop an AL after fertilization (Fig. 3c). Our results showed that the egg envelope of the adhesive eggs (secondary evolution) possessed a new 4th layer and AL compared with the semi-buoyant egg envelope.

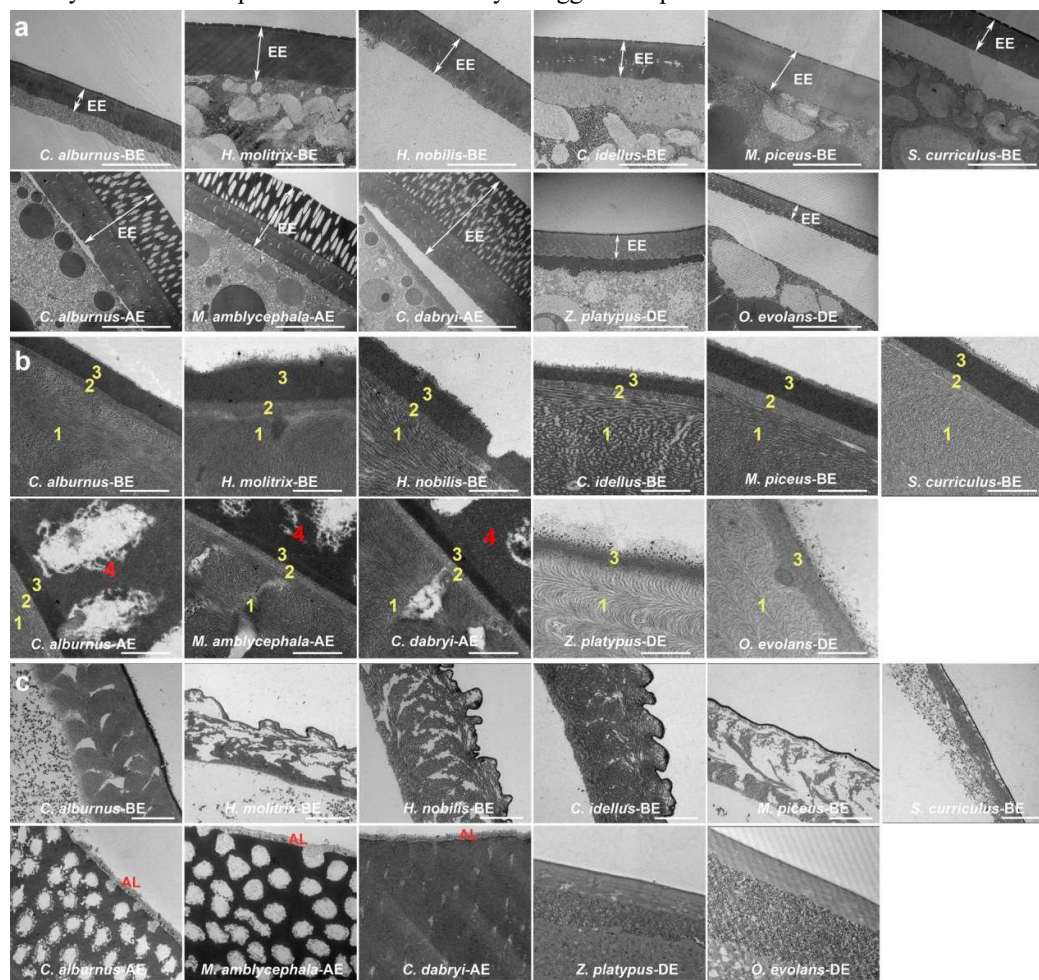


Fig. 3 The egg envelope ultrastructure of the semi-buoyant eggs of 6-representative species (*C. alburnus*-B, *H. molitrix*, *H. nobilis*, *C. idellus*, *M. piceus*, and *S. curriculus*), the demersal eggs of

2-representative species (*Z. platypus* and *O. evolans*) and the adhesive eggs (secondary evolution) of 3-representative species (*C. alburnus*-A, *M. amblycephala*, and *C. dabryi*) of the East Asian endemic group of cyprinids. (a) TEM images of the egg envelopes of unfertilized eggs (scale bar is 10 μ m). EE: egg envelope. (b) TEM images of the egg envelopes of unfertilized eggs (scale bar is 500 nm). The egg envelope of semi-buoyant eggs was divided into three layers (1, 2 and 3), the egg envelope of demersal eggs was divided into two layers (1 and 3), and the egg envelopes of adhesive egg was divided into four layers (1, 2, 3 and 4). (c) TEM images of the adhesive layer of eggs at 0.5 h postfertilization. (scale bar is 2 μ m). AL, adhesive layer; BE, semi-buoyant egg; AE, adhesive egg; DE, demersal egg.

Transcriptome sequence analysis, assembly and functional annotation

The Iso-Seq library was prepared according to the Isoform Sequencing protocol (Iso-Seq) using the Clontech SMARTer PCR cDNA Synthesis Kit and the BluePippin Size Selection System protocol, as described by Pacific Biosciences (PN 100-092-800-03). Pooled cDNA libraries derived from egg and ovarian tissues of *C. alburnus* were prepared and sequenced (Supplementary Table S2), and the total number of unigenes ranged from 123 to 17,562 bp, with an average length of 2,424 bp and an N50 of 2,981 bp. The length distribution of all transcripts is shown in Supplementary Figure S5. Annotation results of Venn diagram analysis and the statistical results of the GO and KEGG classification of all annotated unigenes are shown in Supplementary Figure S6.

Proteome analysis

Among the total identified egg proteins of *C. alburnus*, 7,017 proteins could be quantified through the use of tandem mass tag (TMT) labelling and HPLC fractionation followed by LC-MS/MS analysis. Among the quantified proteins, differentially expressed proteins were identified, which were involved in regulating the mechanism of hydration or adhesiveness (supplementary Data. S2). Egg envelope proteins of *C. alburnus* were identified, among which 1,714 proteins could be quantified through the use of label-free HPLC fractionation followed by LC-MS/MS analysis. Differentially expressed proteins were identified, which were involved in regulating the mechanism of adhesiveness (supplementary Table. S3).

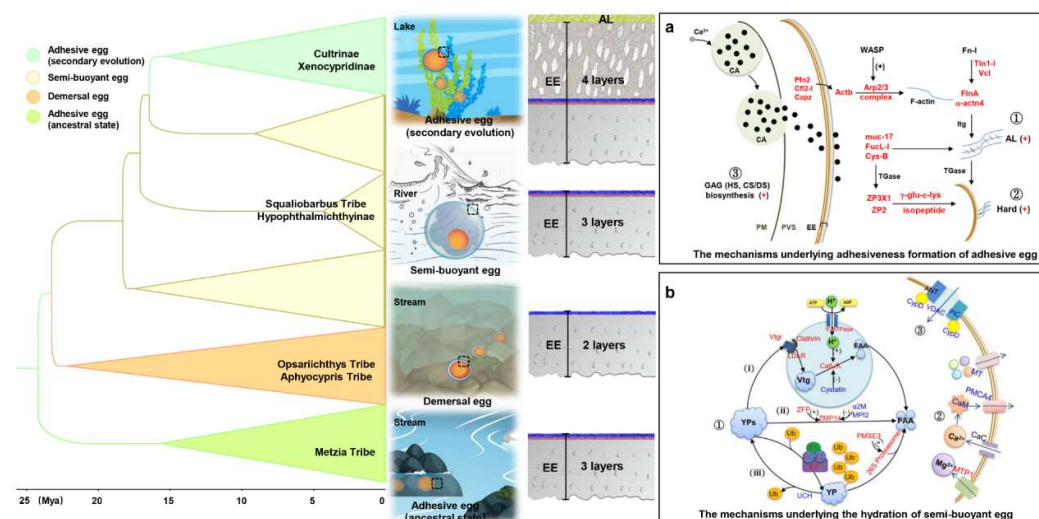


Fig. 4 A schematic summary of the evolutionary history of different egg types among endemic East Asian cyprinid clades and the corresponding adaptive mechanisms. (a) The mechanisms

underlying the hydration of semi-buoyant eggs. ① Three pathways of yolk protein degradation. (i) Vtg degradation pathway; (ii) zinc metalloproteinase pathway; (iii) ubiquitin-proteasome pathway. ② The pathways of Ca^{2+} and Mg^{2+} active transport. ③ The molecular structure of the egg envelope permeability transition pore. (b) The mechanisms underlying adhesiveness formation of adhesive eggs. ① The crosslinks of microfilament-associated proteins and adhesive-related proteins. ② The hardening of the egg envelope. Proteins in red show significant upregulation in adhesive eggs. Details are shown in Supplementary Data 2 and Table S3. ③ The biosynthesis of glycosaminoglycan (GAG) in the ovary, including the biosynthesis of immediate precursors of GAG determined by proteomics analysis (details are shown in Supplementary Table S4 and Figure S11) and the synthesis and modification of heparan sulphate (HS) and chondroitin/dermatan sulphates (CS/DS) by transcriptomics analysis (details are shown in Supplementary Table S5 and Figure S12). EE, egg envelope; PVS, perivitelline space; Yp, yolk protein; FAA, free amino acids; Vtg, vitellogenin; Ub, ubiquitin; CaC, Ca^{2+} channel; CA, cortical alveoli; PM, plasma membrane; AL, adhesive layer.

Discussion

The results of our phylogenetic reconstruction showed that the burst of diversification in the endemic East Asian cyprinids occurred from the early to middle Miocene (Fig. 1a). The endemic East Asian cyprinid clade evolved with diverse spawning strategies (adhesive, demersal and semi-buoyant eggs) as adaptations to different flow conditions in the Yangtze river-lake system. In the present study, the reconstruction of the ancestral state (egg types) revealed that adhesive eggs were the ancestral state, which was consistent with a previous study¹⁰. The Metzia Tribe lays adhesive eggs (ancestral state) adhering to stones, and the Aphyocypris and Opsariichthys Tribes produce demersal eggs falling into stone cracks; both of these strategies are adaptations to the environment of slow-flowing streams. These Tribes are mainly distributed in southern East Asia², where the landscape consists of mountain ranges with many streams. Southern East Asia has been subject to the Asian monsoon since the Eocene, and the landscape in this area formed when the climate changed to a subtropical humid climate²⁷. Coincidentally, our results showed that the ancestor of the endemic East Asian cyprinid clade probably laid adhesive eggs that originated evolutionarily and diversified approximately 28 Mya under stream conditions.

Subsequently, the semi-buoyant eggs spawned by Hypophthalmichthyinae and the Squaliobarbus Tribe emerged approximately 18 Mya (Node 3, Fig. 1a), a time period coincident with the major phase of the birth of the Yangtze River, estimated in the pre-Miocene⁶, and the time of East Asian summer monsoon intensification, close to the start of the Miocene²⁸. As semi-buoyant eggs are adapted to a river flow of ca. 0.63-1.83 m/s¹¹, the length of the river needs to meet the requirements for semi-buoyant eggs to complete hatching, to develop into young fishes capable of swimming, and to avoid entering the ocean. Therefore, the appearance of semi-buoyant eggs could be regarded as an indicator marking the formation of a long river, suggesting an early Miocene birth time of the Yangtze River (Fig. 1b). To adapt to fast-flowing rivers, semi-buoyant eggs undergo hydration after they are produced. Decreased protein contents (Supplementary Figure S7) and increased T-FAA levels (Fig. 2b) result from yolk protein degradation to regulate the process of hydration in semi-buoyant eggs, which is similar to what is observed in marine pelagic eggs²⁹. Based on proteomic analysis of *C. alburnus*, three pathways of yolk protein degradation were found to regulate the mechanisms of hydration in semi-buoyant

eggs (Fig. 4b): (i) the “Vtg degradation pathway”, (ii) the “zinc metalloproteinase pathway” and (iii) the “ubiquitin-proteasome pathway”. As vitellogenin is the main component of yolk protein, the “Vtg degradation pathway” is a major mechanism of yolk protein degradation in freshwater semi-buoyant eggs, which are highly similar to marine pelagic eggs¹³. In a functional verification test, the membrane diameter of zebrafish eggs was shown to be significantly reduced by bafilomycin A1 and cathepsin L inhibitor treatment relative to those of the control (Fig. 2d), suggesting that V-ATPase and cathepsin L could affect egg hydration. However, the zinc metalloproteinase pathway and the ubiquitin-proteasome pathway were initially found to be involved in yolk protein degradation in freshwater semi-buoyant eggs (Fig. 4b), which is not reported in marine pelagic eggs.

Moreover, the major contents of Ca^{2+} and Mg^{2+} ions were prominently increased in semi-buoyant eggs after fertilization relative to those in adhesive and demersal eggs (Supplementary Figure S8), which indicated that semi-buoyant eggs took up more Ca^{2+} and Mg^{2+} to increase their hydration, while the cationic osmolality of pelagic eggs was determined by K^+ and Na^+ in the seawater^{30, 31}. Our proteomic results indicated that the active transport pathways of Ca^{2+} and Mg^{2+} were enhanced in semi-buoyant eggs compared with those in adhesive eggs (Fig. 4b and Supplementary Data S2). This difference is likely related to cations, as more Ca^{2+} and Mg^{2+} are present in freshwater, and more Na^+ and K^+ are found in the ovarian fluid of marine fish³². Additionally, the permeability transition pore is permeable to solutes smaller than 1.5 kDa, such as ions, and is composed of cyclophilin D (CypD), voltage-dependent anion-selective channel protein (VDAC), adenine nucleotide transporter (ANT) and phosphate carrier protein (PiC) in the vitelline envelope^{33, 34}. The protein contents of CypD, VDAC, ANT and PiC were significantly lower in semi-buoyant eggs than in adhesive eggs (Fig. 4b and Supplementary Data S2). In the functional verification test, the membrane diameter of zebrafish was significantly reduced by cyclosporine A and DIDS relative to that in controls (Fig. 2d), which revealed that CypD and VDAC could regulate egg hydration. Therefore, our results indicated that the vitelline envelope permeability transition pore not only affected ion transport but also regulated water permeability in semi-buoyant eggs. In contrast, the water permeability of marine pelagic eggs was dependent on aquaporins³⁵, revealing a different adaptation in response to different environmental pressures between marine pelagic and freshwater semi-buoyant eggs. In the hyperosmotic sea environment, oocytes take up water during oocyte maturation in the ovary, while in the freshwater hypotonic environment, the osmolality of semi-buoyant eggs is higher than that of freshwater eggs. Therefore, semi-buoyant eggs increase osmotic pressure by taking up water quickly by accumulating ions and degrading yolk proteins to achieve suspension in the river and avoid floating into the ocean or other adverse conditions.

Finally, the adhesive eggs produced by Cultrinae and Xenocyprinae evolved secondarily approximately 14 Mya (Node 4, Fig. 1a), a time period coincident with the progressive intensification of the East Asian monsoon following the mid-Miocene Climatic Optimum³⁶. The Yangtze River system developed in response to the strengthening of the East Asian summer monsoon²⁸, which might have promoted the formation of the cross-linked river-lake system. Therefore, both the burst of lacustrine species diversification and the secondary evolution of adhesive eggs suggest a mid-Miocene timing of the formation of the cross-linked river-lake system in the middle and lower reaches of the Yangtze River (Fig. 1c). Adhesive eggs (secondary evolution) exhibited better adherence on submerged plants in the lake, with the egg envelope

increasing to four layers and with an AL (Fig. 3). Our TEM results showed that the egg envelope of adhesive eggs (secondary evolution) possessed a 4th layer, in contrast to the semi-buoyant egg envelope (Fig. 3), which was regulated by ZP3X1. The results of our proteomics analysis showed that the levels of ZP2 and ZP3X1 were significantly upregulated in adhesive egg envelopes (Supplementary Table S3) and that ZP2 was located in two types of egg envelopes (the 2nd/3rd layer). Additionally, ZP3X1 was located in the egg envelope (the 4th layer) of adhesive eggs according to our immunofluorescence localization analysis (Supplementary Figure S9). ZP2 could be crosslinked with ZP3 by the γ -glutamyl- ϵ -lysine isopeptide catalysed by transglutaminase (TGase), which is closely related to fish egg envelope hardening^{37, 38}. Our proteomics results showed that the amount of TGase was obviously increased in adhesive eggs (Supplementary Data 2), which suggested that the adhesive egg envelope was likely to change from fragile to tough after fertilization, thereby resisting stressors present in lake environments.

Based on our proteomic results, multiple microfilament-associated proteins were confirmed in *C. alburnus* eggs, including actin, Arp 2, WASP proteins, filamin-A, α -actinin, vinculin, profilin, cofilin and capping protein (Supplementary Data 2 and Table S3). The Arp 2/3 complex, activated by WASP proteins, mediated actin polymerization and induced the initiation of new filaments^{39, 40}. Filamin-A and α -actinin stabilize the entire network by cross-linking actin filaments³⁹. Vinculin and talin, which bind to fibronectins, form a network with actin filaments and may be beneficial to fish egg morphology⁴¹, while profilin, cofilin and capping protein could restrict polymerization to new filaments close to the plasma membrane^{39, 40, 42}. Fibronectin, collagen, actin and myosin are also substrates of TGase⁴³⁻⁴⁵. Moreover, adhesion-related proteins were detected in the *C. alburnus* egg envelope; these proteins included mucin, fucoselectin, and cystatin-B (Supplementary Table S3), which might present a close relationship with egg adhesion⁴⁶⁻⁴⁸. Further recruitment of cystatins via the regulation of surface TGase is essential for cell adhesion^{49, 50}. Therefore, the crosslinking catalysed by TGase among microfilament-associated proteins, adhesive-related proteins and egg envelope proteins (ZPs) might increase egg envelope hardness and adhesiveness in adhesive eggs (Fig. 4a). Furthermore, in the present study, CA and the outer 4th layer of fertilized adhesive eggs showed positive reactions in Alcian blue-periodic acid Schiff (AB-PAS) histochemical assays, in contrast to the semi-buoyant eggs of *C. alburnus* (Supplementary Figure S10). Our proteomics results indicated that the relative protein expression levels of multiple enzymes (Gfpt1, Gpi1, Nagk, Udgh, and Galt) were significantly upregulated in the *C. alburnus*-A ovary (Supplementary Table S4); these enzymes catalysed the early steps in GAG biosynthesis (Supplementary Figure S11), and their presence suggested that adequate precursor pools were present during oogenesis for the large increase in GAG biosynthesis during the oogenesis of *C. alburnus*-A. Our transcriptomics analyses showed that the mRNA expression enzymes involved in GAG (heparan sulphate (HS), chondroitin sulphate (CS) /dermatin sulphate (DS)) biosynthesis and modification was detectable in the *C. alburnus* ovary, which suggested that the biosynthesis and modification of GAG (HS and CS/DS) might be enhanced in *C. alburnus*-A (Supplementary Figure S12 and Table S5). Glucosamine is one of the units of the repeating disaccharides of HS, which is the preferred substrate for cell adhesion⁵¹, and our monosaccharide analysis showed that the contents of glucosamine in the egg envelope of *C. alburnus*-A were significantly higher than those in *C. alburnus*-B after fertilization (Supplementary Figure S13). Therefore, the enhancement of GAG biosynthesis in *C. alburnus*-A might provide a material foundation for the development of adhesiveness, and GAG might interact with the proteins of the egg envelope or adhesive layer,

which could contribute to egg envelope hardening (Fig. 4a).

In conclusion, this study provides an integrative view of the adaptive evolution of the egg types of the endemic East Asian cyprinids, from phenotypic (including egg biochemical and morphological alterations) and molecular mechanisms to convergent evolution, by using a combination of bioinformatics and multiomics analyses. The endemic East Asian cyprinid clade has evolved diverse spawning strategies (adhesive, demersal and semi-buoyant eggs) for adaptation to different hydrological and ecological conditions in the East Asia cross-linked river-lake system. The adhesive eggs were indicated to represent the ancestral state, followed by the development of demersal eggs, both representing likely adaptations to static or slow-flowing waters. Semi-buoyant eggs began to emerge ~18 Mya as an adaptation to fast-flowing conditions, thus indicating an early Miocene birth of the Yangtze River. The buoyancy of the eggs was enhanced by increasing hydration through 3 yolk protein degradation pathways, ion transport pathways and egg envelope permeability transition pores. The adhesive eggs evolved again from semi-buoyant eggs ~14 Mya, with the egg envelope increasing to 4 layers and an adhesive layer. The adhesiveness of the adhesive eggs was significantly increased by the crosslinking of microfilament/adhesive-related proteins and enhanced glycosaminoglycan biosynthesis increasing their adherence to submerged plants in the lake; these findings in turn indicate the cross-linked river-lake system was most likely formed in the middle Miocene with the intensification of the East Asian monsoon. The endemic clade of East Asian Cyprinidae displays diverse ecological and phenotypic traits, enabling us to study the large-scale evolution of the river-lake system, which is usually difficult to understand on the basis of traditional geological research.

Methods

Phylogenetic analysis and evolution of egg types. The ingroup samples included in the phylogenetic analysis in this study included 60 species from the East Asian endemic groups of Cyprinidae, and 3 Cyprinidae species not belonging to the East Asian endemic groups of Cyprinidae (*Danio rerio*, *Barilius bendelisis*, and *Rasbora lateristriata*) were selected as outgroups. The mitochondrial genome data of 63 species were downloaded from GenBank (Supplementary Data S1). Twelve extracted protein-coding genes (PGCs), in addition to ND6⁵², were aligned in batches with MAFFT in the codon-alignment mode of PhyloSuite⁵³. The GTR + I + G4 model of evolution was chosen according to the Bayesian information criterion (BIC) implemented in ModelFinder⁵⁴. The ML analyses were performed using the program IQ-TREE⁵⁵ under the best-fit model with 200,000 ultrafast bootstraps as well as the Shimodaira–Hasegawa-like approximate likelihood-ratio test⁵⁶. Bayesian inference (BI) analyses were performed using MrBayes 3.2.6⁵⁷ with default settings and 4×10^6 MCMC generations (effective sample size (ESS) for each parameter >200). The initial 25% of sampled data were discarded as burn-in.

To trace the evolution of egg types within the East Asian endemic groups of Cyprinidae, we mapped the states of fish egg types onto the cladogram from our best ML tree topology. The ancestral states for egg types were reconstructed via BI implemented in RASP v.4.0⁵⁸ with Bayesian binary MCMC (BBM). Ten MCMC chains were run simultaneously for 5×10^6 generations with the fixed Jukes-Cantor (JC) model, and the maximum number of areas was set to 1.

Divergence time estimation. Based on the harmonic mean (HM) of the likelihood values of the

MCMC samples⁵⁹ and a more accurate stepping-stone (SS) sampling method⁶⁰ in MrBayes 3.2.6, a relaxed molecular clock method was found to be more suitable for estimating divergence time. Divergence times were estimated by using the ML topology in BEAST ver 1.8.1 under an uncorrelated lognormal relaxed molecular clock model, with the Yule speciation process as the tree prior. Based on the best-fit model of GTR + I + G4, we concurrently ran four independent MCMC tree searches for 1×10^9 generations and sampled every 100,000 generations. Then, four tree-subsets were combined after discarding the initial 10% as burn-in using LogCombiner v1.8.0. We calculated the maximum clade credibility tree using TreeAnnotator v1.8.0. The convergence and effective sample size (ESS) were measured by Tracer ver. 1.5. All ESS values were greater than 200. The following constraints were used for time calibration: (i) since the fossil of *Ecocarpia ningmingensis* is the earliest and most reliable fossil available for the East Asian endemic groups of Cyprinidae⁶¹, the age constraint of 23.03-33.9 Mya was assigned to the root of this group; (ii) the age constraint of 16-20.4 Mya was set for *Ctenopharyngodon*, because the available fossil of *Ctenopharyngodon* sp. is the oldest fossil of this lineage to our knowledge⁶²; (iii) the earliest reported fossil of *Hypophthalmichthys molitrix* was dated to 3.4-5.2 Mya, which was assigned as the age constraint⁶³.

Sample collection. Eggs of representative East Asian endemic groups of Cyprinidae, including *H. molitrix*, *H. nobilis*, *C. idellus*, *M. piceus*, *S. curriculus*, *C. alburnus*-B, *M. amblycephala*, *C. dabryi*, and *C. alburnus*-A, were collected by artificial propagation from fisheries in Jinzhou and Ezhou, Hubei Province (in the Yangtze River Basin), from April to July. Mature females and males were stimulated to spawn by using the maturation-inducing steroid (MIS) 17 α , 20 β -dihydroxy-4-pregnen-3-one. Eggs of *Z. platypus* were obtained by natural propagation were collected from Jiyuan City, Henan Province and eggs of *O. evolans* produced by natural propagation were collected from Xingtai City, Hebei Province. Ovulated eggs were divided into four groups: unfertilized eggs and eggs at 0, 0.5, and 1 h postfertilization, and then maintained in aquaculture water (osmolality was 9.3 ± 1.2 mosmol kg⁻¹, and Mg²⁺, K⁺, Ca²⁺ and Na⁺ concentrations were 23.10 ± 1.73 , 9.53 ± 1.52 , 44.72 ± 8.41 , 20.27 ± 2.12 mg L⁻¹, respectively). Unfertilized *M. formosae* eggs were collected from Nanning, Guangxi Province. One portion of the samples was placed in Bouin's solution or 2.5% glutaraldehyde, and the remaining eggs were immediately frozen in liquid nitrogen and stored at -80 °C until analysis. All procedures were performed with the approval of the Animal Care and Use Committee of the Institute of Hydrobiology, Chinese Academy of Sciences (Approval ID: IHB 20140724).

Water content and osmolality. The weighed eggs (wet mass, W_m , g, n=3 to 5) were lyophilized in a vacuum freeze drier (ALPHA 1-4, Christ, Germany). After 48 h, their dry mass (W_d , g) was determined with a milli-balance (BSA124S, Sartorius, Germany), and they were then stored in a desiccator. The egg water content was calculated using the following equation: water content (%) = $100 (W_m - W_d) W_m^{-1}$. According to the “egg squash” technique⁶⁴, fresh eggs (n=3 to 6) were added to the barrel of a 1 or 2 ml syringe with a fine needle. Next, the squash liquid was centrifuged at 10,000 g for 1 min, and 50 μ l of the supernatant was collected and used for measuring osmolality on an OSMOMAT 010 instrument (Gonotec GmbH, Germany).

Free amino acid and inorganic ion content. Egg samples of 0.5-1 g per replicate (n= 3 or 4) were homogenized in 5 ml of 15 mM HCl for 3-5 min by using a standard micro-homogenizing package. Samples were centrifuged at 12,000 g for 10 min at 4 °C, and the supernatant was collected. One millilitre of the supernatant was collected and mixed thoroughly with 1 ml of

sulfosalicylic acid (Sinopharm Chemical Reagent Co., Ltd, China) for 1 h at 4 °C. After centrifugation at 12,000 g for 15 min at 4 °C, the supernatant was collected and filtered through a 0.22 µm sieve (JINTENG, China). One portion of the supernatant was then subjected to free amino acid measurement with an amino acid analyser (A300, Membrapure, Germany), and the remaining portion of the supernatant used for inorganic ion content measurement was diluted tenfold with deionized H₂O, followed by measurement on an optical emission spectrometer (Optima 8000, PerkinElmer, USA).

Protein content. Egg samples of 0.2-0.5 g per replicate (n= 3 or 4) were homogenized in radioimmunoprecipitation assay buffer (RIPA, Beyotime Biotechnology, China) with 1 mM phenylmethanesulfonyl fluoride (PMSF, Beyotime Biotechnology, China) for 2 min by using a standard microhomogenizing package (PRO-PK-02200S, PRO Scientific, USA). The homogenate was centrifuged at 12,000 g for 10 min at 4 °C, and the supernatant was collected. The supernatant was diluted tenfold using RIPA buffer, and the protein content was determined with an Enhanced BCA Protein Assay Kit (Beyotime Biotechnology, China).

Electron microscopy. For TEM, fish eggs were prefixed in 2.5% glutaraldehyde solution and then fixed in 1% aqueous osmium tetroxide. The specimens were embedded in epon-araldite after dehydration. Ultrathin sections were cut with diamond knives on an EM UC7 ultramicrotome (Leica, Germany), treated with the contrast agents uranyl acetate and lead citrate and examined on a Tecnai G² 20 TWIN microscope (FEI, USA).

Histological analysis. For light microscopy, fish eggs were fixed in Bouin's solution and embedded in paraffin, and blocks were sectioned on a microtome. The 5 µm-thick sections with the largest cut surface were stained with haematoxylin and eosin (H&E). Alcian blue (pH 2.5)-periodic acid Schiff (AB-PAS) staining was used to determine the histochemical contents of the egg structures. Histological observations were performed using light microscopy (Nikon Eclipse E100, Japan).

RNA extraction and transcriptome sequencing, assembly and annotation. The total RNA (n=3) of eggs from two individuals was isolated using TRIzol (Invitrogen, USA) according to the manufacturer's instructions. The quantity and quality of the total RNA were verified on an Agilent 2100 Bioanalyzer and by gel electrophoresis, respectively. Approximately 10 µg of each RNA sample was pooled for transcriptome library preparation. The Iso-Seq library was prepared according to the Iso-Seq protocol using the Clontech SMARTer PCR cDNA Synthesis Kit and the BluePippin Size Selection System protocol, as described by Pacific Biosciences (PN 100-092-800-03). Gene functions were annotated based on the following databases: Nr (NCBI nonredundant protein sequences); Nt (NCBI nonredundant nucleotide sequences); Pfam (Protein family); KOG/COG (Clusters of Orthologous Groups of proteins); Swiss-Prot (a manually annotated and reviewed protein sequence database); KO (KEGG Ortholog); GO (Gene Ontology). For CDS prediction, the ANGEL pipeline, a long-read implementation of ANGEL, was used to determine protein coding sequences from cDNAs. We used confident protein sequences from the studied species or closely related species for ANGEL training and then ran the ANGEL prediction for given sequences. For transcription factor (TF) analysis, animal TFs were analysed using the animal TFDB 2.0 database. For coding potential analysis, Coding-Non-Coding-Index (CNCI) profiles adjoining nucleotide triplets were used to effectively distinguish protein-coding and noncoding sequences independent of known annotations. We use CNCI with default parameters. Coding Potential Calculator (CPC) mainly through assesses the extent and quality of ORFs in a

transcript and searches the sequences against known protein sequence databases to clarify the coding and noncoding transcripts. We used the NCBI eukaryote protein database and set the e-value to $1e-10$ in our analysis. We translated each transcript into all three possible frames and used Pfam Scan to identify the occurrence of any of the known protein family domains documented in the Pfam database. Any transcript with a Pfam hit was excluded in the following steps. For the Pfam searches, default parameters of $-E\ 0.001$ $--domE\ 0.001$ were used. For simple sequence repeat (SSR) analysis SSRs in the transcriptome were identified using the MISA microsatellite finder. This tool allows the identification and localization of perfect microsatellites as well as compound microsatellites that are interrupted by a certain number of bases.

TMT-based and label-free quantitative proteomic analysis. TMT-based quantitative proteomics was conducted to study the changes in the expression of proteins in the eggs of *C. alburnus*-B and *C. alburnus*-A at four time points (unfertilized, and 0, 0.5 and 1 h postfertilization, $n=3$), and label-free quantitative proteomics was conducted to study the changes in the expression of proteins in the egg envelopes of *C. alburnus*-B and *C. alburnus*-A at 1 h after fertilization. Samples in lysis buffer (8 M urea, 1% Protease Inhibitor Cocktail) placed on ice were sonicated three times using a high-intensity ultrasonic processor (Scientz). The remaining debris was removed by centrifugation at 12,000 g at 4 °C for 10 min. Finally, the supernatant was collected, and the protein concentration was determined with a BCA kit according to the manufacturer's instructions. Subsequently, 100 µg of protein from each sample was reduced with dithiothreitol, alkylated with iodoacetamide, digested with trypsin and labelled with TMT reagents (Thermo, USA) (vitelline envelope samples for label-free quantitative proteomics do not require TMT labelling). The tryptic peptides were fractionated by high-pH reverse-phase HPLC using an Agilent 300Extend C18 column. The peptides were subjected to a nanospray ionisation (NSI) source, followed by tandem mass spectrometry (MS/MS) in a Q ExactiveTM Plus system (Thermo) coupled online to an ultraperformance liquid chromatography (UPLC) system. Proteins in fish eggs with a fold change > 1.2 or < 0.83 and $p < 0.05$ and proteins in egg envelope with a fold change > 1.5 or < 0.67 and $p < 0.05$ were considered differentially expressed. The functions of differentially expressed proteins were determined via GO and KEGG pathway analyses and functional enrichment analysis. The GO classification was performed using the UniProt-GOA database (<http://www.ebi.ac.uk/GOA/>), and InterProScan software was used to predict the subcellular localization of the proteins. For quality control purposes in TMT-based quantitative proteomics, the mass errors and lengths of all peptides were analysed. The distribution of the mass error was close to zero and mostly less than 10 ppm (Supplementary Figure S14a), which indicated that the mass accuracy of the MS data was sufficient. Additionally, the lengths of most peptides were distributed between 8 and 20 amino acid residues (Supplementary Figure S14b), which agrees with the properties of tryptic peptides, indicating that the results met the standard. For quality control purposes in the label-free quantitative proteomic analysis, the mass errors were mostly less than 10 ppm (Supplementary Figure S15a), and lengths of most peptides were distributed between 7 and 20 amino acid residues (Supplementary Figure S15b), which meant that the results met the standard.

Preparation of antibodies and Immunohistochemistry. Two kinds of anti-ZP antibodies were synthesized in rabbits for the *C. alburnus* ZP proteins (ZP2 and ZP3X1, differentially expressed proteins from Supplementary Table S3) and prepared by Dia-An Biotech, Inc., in Wuhan, China. Synthetic peptides designed from the amino acid sequences of the two ZP proteins were used as

the antigens. Their peptide sequences were as follows: ZP2: QNFNQRTGLKTDC and YQQGRKGVSSPPDPE; ZP3X1: SANGGGDDGDEDFRDGFV. Unfertilized eggs of *C. alburnus* were fixed in 4% paraformaldehyde in PBS at 4 °C overnight, dehydrated through a graded ethanol series, and embedded in paraffin. Five-micrometre sections were placed on the slides, and the samples were dried. After dewaxing and rehydration, the sections were subjected to antigen retrieval. After washing in PBS, the sections were incubated in a 1:100 dilution of 1% normal goat serum in PBS for 1 h to block nonspecific antibody staining and were then incubated with the primary antibody (anti-ZP2) diluted 1:100 in 1% normal goat serum in PBS at 4 °C overnight. After washing in PBS, the samples were incubated for 50 min at room temperature with a horseradish peroxidase (HRP)-conjugated secondary antibody. Cy3-TSA in diluted 1:2,000 was added, and the samples were incubated with TSA at room temperature for 10 min. After antigen retrieval, the sections were incubated in a 1:100 dilution of 1% normal goat serum in PBS for 1 h to block nonspecific antibody staining and were then incubated with the primary antibody (anti-ZP3X1) diluted 1:100 in 1% normal goat serum in PBS at 4 °C overnight. After washing in PBS, the samples were incubated for 50 min at room temperature with a fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Finally, the sections were observed under a light microscope capable of fluorescence detection and differential interference (Nikon Eclipse C1, Japan).

Functional verification tests. For these test, solutions of 100 nM bafilomycin A₁ (Cayman Chemical, USA), 1.5 μM cathepsin L inhibitor (Cayman Chemical, USA), 5 μM cyclosporine A (Cayman Chemical, USA) and 200 μM 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid disodium salt (DIDS, Sigma-Aldrich, Germany) were prepared in 0.05% dimethyl sulfoxide (DMSO, BioFroxx, Germany), to inhibit the activities of V-ATPase, cathepsin L, cyclophilin D and voltage-dependent anion-selective channel protein, respectively⁶⁵⁻⁶⁸. Mature zebrafish were treated with 10 μL of 0.05% DMSO or the inhibitor solution by intraperitoneal injection (n=3). After 17 h, zebrafish embryos (as semi-buoyant model, identical to other semi-buoyant eggs at the macroindex and biochemical levels, Supplementary Figure S16) were obtained by artificial insemination and then immersed in 0.05% DMSO or inhibitor solutions. Immediately thereafter, photos of eggs were taken with a stereomicroscope (MODEL C-BD230, Nikon, Japan) at 1, 5, 10, 15, 20, 25, 30 and 60 min postfertilization, and egg diameter was measured with ImageJ (k 1.45, National Institutes of Health, USA).

Monosaccharide analysis. An ICS 5000+ ion chromatography (IC) system (Thermo-Fisher Scientific, USA) was employed for monosaccharide analysis. The electrochemical detector was equipped with a gold working electrode and a pH/Ag/AgCl composite reference electrode. A CarboPac PA 10 guard column (50 mm × 4 mm) and CarboPac PA 10 (250 mm × 4 mm) separation column (Thermo-Fisher Scientific) were used for sugar separation. The system was operated in gradient mode as follows: 200 mmol L⁻¹ NaOH from -30 to -15.1 min, 18 mmol L⁻¹ NaOH from -15 to 0 min, and 18 mmol L⁻¹ NaOH from 0 to 20 min with a flow rate of 1.0 mL min⁻¹. The column and the cell of the pulsed amperometric detector were maintained at temperatures of 30 °C and 35 °C, respectively.

Statistical analyses. Statistical analyses of the data were performed by using SPSS package 19.0 (SPSS, Chicago, IL, USA). All values are presented as the mean ± standard error (SEM). The Kolmogorov-Smirnov test and Levene's test were employed to check the normality and homogeneity of variances in the data, respectively. One-way analysis of variance (ANOVA) and

Tukey's multiple comparison tests were applied to determine significant differences between the data of the unfertilized and fertilized groups. Significant differences were set at the $p < 0.05$ (*) and $p < 0.01$ (**).

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request and the source data underlying Fig. 2a-d, Supplementary Figs. 2-3, 7-8, 13, and 16 and Supplementary Table 1 are provided as a Source Data File.

Acknowledgements

This study was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31040000).

Author contributions

Y.W., F.C. and P.X. coordinated the manuscript writing, Y.W., F.C., J.H., L.C. and J.C. conceived and designed the experiments, Y.W., F.C., J.H., L. C., G.X., Y.Z. and Y.P. participated to the sampling campaign and the field work, Y.W., and F.C. performed the bioinformatics and statistical analyses.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Ping Xie and Jun Chen.

References

1. Nelson, J. S. et al. *Fishes of the world. Fifth edition.* (2016).
2. Chen, Y. Y. Fauna Sinica, Osteichthys: Cypriniformes (Part II) (Beijing: Science Press, 1998).
3. Wu, G. X. et al. The influence of mechanical and thermal forcing by the Tibetan Plateau on Asian climate. *J. Hydrometeorol.* **8**, 770-789, doi:10.1175/jhm609.1 (2007).
4. Duan, A. M. & Wu, G. X. Role of the Tibetan Plateau thermal forcing in the summer climate patterns over subtropical Asia. *Clim. Dyn.* **24**, 793-807, doi:10.1007/s00382-004-0488-8 (2005).
5. Yu, G. et al. High carbon dioxide uptake by subtropical forest ecosystems in the East Asian monsoon region. *Proc. Natl Acad. Sci. USA* **111**, 4910-4915, doi:10.1073/pnas.1317065111 (2014).
6. Zheng, H. Birth of the Yangtze River: age and tectonic-geomorphic implications. *Natl. Sci. Rev.* **2**, 438-453, doi:10.1093/nsr/nwv063 (2015).
7. Wang, X., Li, J. & He, S. Molecular evidence for the monophyly of East Asian groups of Cyprinidae (Teleostei : Cypriniformes) derived from the nuclear recombination activating gene 2 sequences. *Mol. Phylogenet. Evol.* **42**, 157-170, doi:10.1016/j.ympev.2006.06.014 (2007).
8. Xie, P. The past and present of the Yangtze River and its biodiversity. Wuhan, Changjiang press (2020).
9. Tao, W. et al. Phylogenomic Analysis Resolves the Formerly Intractable Adaptive Diversification of the Endemic Clade of East Asian Cyprinidae (Cypriniformes). *Plos One* **5**, doi:10.1371/journal.pone.0013508 (2010).
10. Cheng, P. L. Studies on phylogenetic pattern and macroevolutionary characteristics of the East Asian Endemic Group of Cyprinids. *P.H.D. Thesis.* University of Chinese Academy of

- Sciences (2015).
11. Yu, L. et al. Ecological Flow Assessment to Improve the Spawning Habitat for the Four Major Species of Carp of the Yangtze River: A Study on Habitat Suitability Based on Ultrasonic Telemetry. *Water* **10**, doi:10.3390/w10050600 (2018).
 12. Kolar, C. S. et al. Bigheaded carps: a biological synopsis and environmental risk assessment. *American Fisheries Society Special Publication* **33**, i-xiv, 1-204 (2007).
 13. Finn, R. N. & Fyhn, H. J. Requirement for amino acids in ontogeny of fish. *Aquac. Res.* **41**, 684-716, doi:10.1111/j.1365-2109.2009.02220.x (2010).
 14. Li, H. & Zhang, S. in *Oocytes: Maternal Information and Functions* Vol. 63 *Results and Problems in Cell Differentiation* (ed M. Kloc) 389-401 (2017).
 15. Rizzo, E., Sato, Y., Barreto, B. P. & Godinho, H. P. Adhesiveness and surface patterns of eggs in neotropical freshwater teleosts. *J. Fish Biol.* **61**, 615-632, doi:10.1006/jfbi.2002.2085 (2002).
 16. Riehl, R. & Patzner, R. A. Minireview: The modes of egg attachment in teleost fishes. *Ital. J. Zool. (Modena)* **65**, 415-420, doi:10.1080/11250009809386857 (1998).
 17. Mansour, N., Lahnsteiner, F. & Patzner, R. A. Physiological and biochemical investigations on egg stickiness in common carp. *Anim. Reprod. Sci.* **114**, 256-268, doi:10.1016/j.anireprosci.2008.09.005 (2009).
 18. Niksirat, H., Andersson, L., Golpour, A., Chupani, L. & James, P. Quantification of egg proteome changes during fertilization in sterlet *Acipenser ruthenus*. *Biochem. Biophys. Res. Commun.* **490**, 189-193, doi:10.1016/j.bbrc.2017.06.019 (2017).
 19. Sano, K. et al. Comparison of Egg Envelope Thickness in Teleosts and its Relationship to the Sites of ZP Protein Synthesis. *J. Exp. Zool. B. Mol. Dev. Evol.* **328**, 240-258, doi:10.1002/jez.b.22729 (2017).
 20. Yamagami, K., Hamazaki, T. S., Yasumasu, S., Masuda, K. & Iuchi, I. MOLECULAR AND CELLULAR BASIS OF FORMATION, HARDENING, AND BREAKDOWN OF THE EGG ENVELOPE IN FISH. *Int. Rev. Cytol.* **136**, 51-92, doi:10.1016/s0074-7696(08)62050-1 (1992).
 21. Hedrick, J. L. Anuran and pig egg zona pellucida glycoproteins in fertilization and early development. *Int. J. Dev. Biol.* **52**, 683-701, doi:10.1387/ijdb.082580jh (2008).
 22. Blackburn, D. G. Evolution of vertebrate viviparity and specializations for fetal nutrition: A quantitative and qualitative analysis. *J. Morphol.* **276**, 961-990, doi:10.1002/jmor.20272 (2015).
 23. Sauquet, H. et al. The ancestral flower of angiosperms and its early diversification. *Nat. Commun.* **8**, doi:10.1038/ncomms16047 (2017).
 24. Foster, C. S. P., Thompson, M. B., Van Dyke, J. U., Brandley, M. C. & Whittington, C. M. Emergence of an evolutionary innovation: Gene expression differences associated with the transition between oviparity and viviparity. *Mol. Ecol.* **29**, 1315-1327, doi:10.1111/mec.15409 (2020).
 25. Gao, W. et al. Genomic and transcriptomic investigations of the evolutionary transition from oviparity to viviparity. *Proc. Natl Acad. Sci. USA* **116**, 3646-3655, doi:10.1073/pnas.1816086116 (2019).
 26. Qi, P. et al. Assessing the genetic diversity and population structure of *Culter alburnus* in China based on mitochondrial 16S rRNA and COI gene sequences. *Biochem. Syst. Ecol.* **50**, 390-396 (2013).
 27. Guo, Z. T. et al. A major reorganization of Asian climate by the early Miocene. *Clim. Past.* **4**, 153-174, doi:10.5194/cp-4-153-2008 (2008).

28. Sun, X. J. & Wang, P. X. How old is the Asian monsoon system? Palaeobotanical records from China. *Palaeogeogr. Palaeocl.* **222**, 181-222, doi:10.1016/j.palaeo.2005.03.005 (2005).
29. Finn, R. N. The maturational disassembly and differential proteolysis of paralogous vitellogenins in a marine pelagophil teleost: A conserved mechanism of oocyte hydration. *Biol. Reprod.* **76**, 936-948, doi:10.1095/biolreprod.106.055772 (2007).
30. Finn, R. N., Ostby, G. C., Norberg, B. & Fyhn, H. J. In vivo oocyte hydration in Atlantic halibut (*Hippoglossus hippoglossus*); proteolytic liberation of free amino acids, and ion transport, are driving forces for osmotic water influx. *J. Exp. Biol.* **205**, 211-224 (2002).
31. Skoblina, M. N. Hydration of oocytes in teleost fishes. *Russ J. Dev. Biol.* **41**, 1-12, doi:10.1134/s1062360410010017 (2010).
32. Jia, Y. D., Niu, H. X., Meng, Z., Liu, X. F. & Lei, J. L. Biochemical composition of the ovarian fluid and its effects on the fertilization capacity of turbot *Scophthalmus maximus* during the spawning season. *J. Fish Biol.* **86**, 1612-1620, doi:10.1111/jfb.12676 (2015).
33. Shoshan-Barmatz, V. & Gincel, D. The voltage-dependent anion channel - Characterization, modulation, and role in mitochondrial function in cell life and death. *Cell Biochem. Biophys.* **39**, 279-292, doi:10.1385/cbb:39:3:279 (2003).
34. Bernardi, P., Rasola, A., Forte, M. & Lippe, G. THE MITOCHONDRIAL PERMEABILITY TRANSITION PORE: CHANNEL FORMATION BY F-ATP SYNTHASE, INTEGRATION IN SIGNAL TRANSDUCTION, AND ROLE IN PATHOPHYSIOLOGY. *Physiol. Rev.* **95**, 1111-1155, doi:10.1152/physrev.00001.2015 (2015).
35. Fabra, M., Raldua, D., Power, D. M., Deen, P. M. T. & Cerda, J. Marine fish egg hydration is aquaporin-mediated. *Science* **307**, 545-545, doi:10.1126/science.1106305 (2005).
36. Favre, A. et al. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biol. Rev.* **90**, 236-253, doi:10.1111/brv.12107 (2015).
37. Iuchi, I., Ha, C. R., Sugiyama, H. & Nomura, K. Analysis of chorion hardening of eggs of rainbow trout, *Oncorhynchus mykiss*. *Dev. Growth. Differ.* **38**, 299-306 (1996).
38. Chang, Y. S., Wang, Y. W. & Huang, F. L. Cross-linking of ZP2 and ZP3 by transglutaminase is required for the formation of the outer layer of fertilization envelope of carp egg. *Mol. Reprod. Devel.* **63**, 237-244, doi:10.1002/mrd.10174 (2002).
39. Welch, M. D. & Mullins, R. D. Cellular control of actin nucleation. *Annu. Rev. Cell. Dev. Biol.* **18**, 247-288, doi:10.1146/annurev.cellbio.18.040202.112133 (2002).
40. Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**, 453-465, doi:10.1016/s0092-8674(03)00120-x (2003).
41. Nuckolls, G. H., Romer, L. H. & Burridge, K. MICROINJECTION OF ANTIBODIES AGAINST TALIN INHIBITS THE SPREADING AND MIGRATION OF FIBROBLASTS. *J. Cell Sci.* **102**, 753-762 (1992).
42. Dos Remedios, C. G. et al. Actin binding proteins: Regulation of cytoskeletal microfilaments. *Physiol. Rev.* **83**, 433-473, doi:10.1152/physrev.00026.2002 (2003).
43. Cohen, I., Youngbandala, L., Blankenberg, T. A., Siefring, G. E. & Brunerlorand, J. FIBRINOLIGASE-CATALYZED CROSS-LINKING OF MYOSIN FROM PLATELET AND SKELETAL-MUSCLE. *Arch. Biochem. Biophys.* **192**, 100-111, doi:10.1016/0003-9861(79)90075-4 (1979).
44. Cohen, I., Blankenberg, T. A., Borden, D., Kahn, D. R. & Veis, A. FACTOR-XIIIA-CATALYZED CROSS-LINKING OF PLATELET AND MUSCLE ACTIN - REGULATION BY

- NUCLEOTIDES. *Biochim. Biophys. Acta.* **628**, 365-375, doi:10.1016/0304-4165(80)90386-4 (1980).
45. Mosher, D. F., Schad, P. E. & Vann, J. M. CROSS-LINKING OF COLLAGEN AND FIBRONECTIN BY FACTOR-XIIIA - LOCALIZATION OF PARTICIPATING GLUTAMINYL RESIDUES TO A TRYPTIC FRAGMENT OF FIBRONECTIN. *J. Biol. Chem.* **255**, 1181-1188 (1980).
 46. Chang, Y. S. & Huang, F. L. Fibroin-like substance is a major component of the outer layer of fertilization envelope via which carp egg adheres to the substratum. *Mol. Reprod. Dev.* **62**, 397-406, doi:10.1002/mrd.10125 (2002).
 47. Portela, I. B., Martinez-Zorzano, V. S., Molist-Perez, I. & Garcia, P. M. Ultrastructure and Glycoconjugate Pattern of the Foot Epithelium of the Abalone *Haliotis tuberculata* (Linnaeus, 1758) (Gastropoda, Haliotidae). *Sci. World J.*, doi:10.1100/2012/960159 (2012).
 48. Hogfors-Ronnholm, E., Norrgard, J. & Wiklund, T. Adhesion of smooth and rough phenotypes of *Flavobacterium psychrophilum* to polystyrene surfaces. *J. Fish Dis.* **38**, 429-437, doi:10.1111/jfd.12250 (2015).
 49. Akimov, S. S., Krylov, D., Fleischman, L. F. & Belkin, A. M. Tissue transglutaminase is an integrin-binding adhesion coreceptor for fibronectin. *J. Cell Biol.* **148**, 825-838, doi:10.1083/jcb.148.4.825 (2000).
 50. Akimov, S. S. & Belkin, A. M. Cell surface tissue transglutaminase is involved in adhesion and migration of monocytic cells on fibronectin. *Blood* **98**, 1567-1576, doi:10.1182/blood.V98.5.1567 (2001).
 51. Peramo, A., Meads, M. B., Dalton, W. S. & Matthews, W. G. Static adhesion of cancer cells to glass surfaces coated with glycosaminoglycans. *Colloids. Surf. B Biointerfaces.* **67**, 140-144, doi:10.1016/j.colsurfb.2008.07.019 (2008).
 52. Miya, M., Satoh, T. R. & Nishida, M. The phylogenetic position of toadfishes (order Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. *Biol. J. Linn. Soc.* **85**, 289-306, doi:10.1111/j.1095-8312.2005.00483.x (2005).
 53. Zhang, D. et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **20**, 348-355, doi:10.1111/1755-0998.13096 (2020).
 54. Kalyaanamoorthy, S., Bui Quang, M., Wong, T. K. F., von Haeseler, A. & Jermin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **14**, 587-+, doi:10.1038/nmeth.4285 (2017).
 55. Lam-Tung, N., Schmidt, H. A., von Haeseler, A. & Bui Quang, M. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).
 56. Guindon, S. et al. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**, 307-321, doi:10.1093/sysbio/syq010 (2010).
 57. Ronquist, F. et al. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **61**, 539-542, doi:10.1093/sysbio/sys029 (2012).
 58. Yu, Y., Harris, A. J., Blair, C. & He, X. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Mol. Phylogenet. Evol.* **87**, 46-49, doi:10.1016/j.ympev.2015.03.008

- (2015).
59. Newton, M. A. & Raftery, A. E. APPROXIMATE BAYESIAN-INFERENCE WITH THE WEIGHTED LIKELIHOOD BOOTSTRAP. *J. R. Stat. Soc. Series B Stat. Methodol.* **56**, 3-48 (1994).
 60. Xie, W., Lewis, P. O., Fan, Y., Kuo, L. & Chen, M.-H. Improving Marginal Likelihood Estimation for Bayesian Phylogenetic Model Selection. *Syst. Biol.* **60**, 150-160, doi:10.1093/sysbio/syq085 (2011).
 61. Chen, G. J., Fang, F. & Chang, M. M. A new cyprinid closely related to cultrins plus xenocyprinins from the mid-Tertiary of South China. *J. Vertebr. Paleontol.* **25**, 492-501, doi:10.1671/0272-4634(2005)025[0492:Ancert]2.0.Co;2 (2005).
 62. Li, C. K. et al. THE ARAGONIAN VERTEBRATE FAUNA OF XIACAOWAN, JIANGSU. *Vertebrata Palasiatica* **21**, 313-327 (1983).
 63. Liu, H. T. & Su, T. T. Pliocene fishes from Yushe Basin, Shansi. *Vertebrata Palasiatica* **6**, 1-47 (1962).
 64. Lonning, S. & Davenport, J. THE SWELLING EGG OF THE LONG ROUGH DAB, HIPPOGLOSSOIDES-PLATESSOIDES-LIMANDOIDES (BLOCH). *J. Fish Biol.* **17**, 359-378, doi:10.1111/j.1095-8649.1980.tb02771.x (1980).
 65. Selman, K., Wallace, R. A. & Cerda, J. Bafilomycin A1 inhibits proteolytic cleavage and hydration but not yolk crystal disassembly or meiosis during maturation of sea bass oocytes. *J. Exp. Zool.* **290**, 265-278, doi:10.1002/jez.1057 (2001).
 66. Zhang, T., Rawson, D. M., Tosti, L. & Carnevali, O. Cathepsin activities and membrane integrity of zebrafish (*Danio rerio*) oocytes after freezing to -196 degrees C using controlled slow cooling. *Cryobiology* **56**, 138-143, doi:10.1016/j.cryobiol.2008.01.002 (2008).
 67. Robinson, B. L. et al. Cyclosporine exacerbates ketamine toxicity in zebrafish: Mechanistic studies on drug-drug interaction. *J. Appl. Toxicol.* **37**, 1438-1447, doi:10.1002/jat.3488 (2017).
 68. Lin, Y.-H. et al. Anion Exchanger 1b in Stereocilia Is Required for the Functioning of Mechanotransducer Channels in Lateral-Line Hair Cells of Zebrafish. *Plos One* **10**, doi:10.1371/journal.pone.0117041 (2015).