SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

November 2021 Vol.64 No.11: 1929–1948
https://doi.org/10.1007/s11427-020-1851-0

The new chimeric *chiron* genes evolved essential roles in zebrafish embryonic development by regulating NAD$^+$ levels

Chengchi Fang[1,3#], Xiaoni Gan[1], Chengjun Zhang[2] & Shunping He[1,3,4*]

[1]*The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China;*
[2]*Kunming Institute of Botany, Chinese Academy of Science, Kunming 650201, China;*
[3]*University of Chinese Academy of Sciences, Beijing 100039, China;*
[4]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China;*
[#]*Current address: Huazhong Agricultural University, Wuhan 430070, China*

Received August 18, 2020; accepted November 16, 2020; published online January 27, 2021

The origination of new genes is important for generating genetic novelties for adaptive evolution and biological diversity. However, their potential roles in embryonic development, evolutionary processes into ancient networks, and contributions to adaptive evolution remain poorly investigated. Here, we identified a novel chimeric gene family, the *chiron* family, and explored its genetic basis and functional evolution underlying the adaptive evolution of *Danioninae* fishes. The ancestral *chiron* gene originated through retroposition of *nampt* in *Danioninae* 48–54 million years ago (Mya) and expanded into five duplicates (*chiron1–5*) in zebrafish 1–4 Mya. The *chiron* genes (*chiron*s) likely originated in embryonic development and gradually extended their expression in the testis. Functional experiments showed that *chiron*s were essential for zebrafish embryo development. By integrating into the NAD$^+$ synthesis pathway, *chiron*s could directly catalyze the NAD$^+$ rate-limiting reaction and probably impact two energy metabolism genes (*nmnat1* and *naprt*) to be under positive selection in *Danioninae* fishes. Together, these results mainly demonstrated that the origin of new chimeric *chiron* genes may be involved in adaptive evolution by integrating and impacting the NAD$^+$ biosynthetic pathway. This coevolution may contribute to the physiological adaptation of *Danioninae* fishes to widespread and varied biomes in Southeast Asian.

new chimeric genes, essential function, NAD$^+$ rate-limiting enzyme, coevolution

Citation: Fang, C., Gan, X., Zhang, C., and He, S. (2021). The new chimeric *chiron* genes evolved essential roles in zebrafish embryonic development by regulating NAD$^+$ levels. Sci China Life Sci 64, 1929–1948. https://doi.org/10.1007/s11427-020-1851-0

INTRODUCTION

How new genes are retained and evolve new functions is one of the most fundamental questions in molecular evolution and species differentiation. The origination and evolution of new genes are important for generating genetic novelties for adaptive evolution and biological diversity (Long et al., 2003; Long et al., 2013). Many studies have focused on new genes created through duplications and *de novo* origination (Heinen et al., 2009; Long et al., 2013; Zhang et al., 2019a; Zhang et al., 2019b). Interestingly, these two processes represent two very different mechanisms responsible for sequence innovations. The duplication evolves from an ancestral sequence and accumulates mutations (Ohno et al., 1968), whereas the *de novo* origination creates novel protein-coding potential from scratch, which has been reported in recent years (Cai et al., 2008; Xie et al., 2012; Zhao et al., 2014; McLysaght and Hurst, 2016; Ruiz-Orera et al., 2018; Zhuang et al., 2019). However, another mechanism that can create great novelty in protein structures and consequently

*Corresponding author (email: clad@ihb.ac.cn)

functions is the recombination of existing coding sequences —exon shuffling (Gilbert, 1978; Long and Langley, 1993; Xia et al., 2016), which involves the shuffling of exons or corresponding domains leads to a chimeric gene. The hybrid structure of such a gene, with a new combination of exons and domains, generates novel molecular functions. Even single exon shuffling events constitute adequate, much simpler mechanisms than both the duplication-divergence and the *de novo* origination models. Many young chimeric genes, which are valuable for investigating gene evolution because they do not require accumulations of much sequence evolution and retain the basic features of their parental genes, have been reported in metazoans from *Drosophila* (Rogers and Hartl, 2012) to humans (Vinckenbosch et al., 2006). For example, Long et al. estimated that ~19%–28% of the exons in eukaryotes were involved in exon shuffling, implying that exon shuffling played an important role in evolution (Long et al., 1995). In mammals, more than 300 families of chimeras created by exon shuffling have been reported, although these often had ancient origins and were therefore very difficult to examine to determine the early stages of their evolution (Patthy, 2008). Furthermore, several distinct molecular mechanisms can create chimeric genes (Rogers and Hartl, 2012; Long et al., 2013), and those involved in the development of ancient chimeric genes are hardly discernable.

What are the roles of new chimeric genes in phenotypic evolution? How the genetic basis of the control of phenotypic and molecular functions has recently been shown to evolve in coding regions and in the regulatory elements of these regions (Weiss, 2005; Long et al., 2013). In particular, newly evolved genes play obvious roles in various critical functions and biological processes (Chen et al., 2013). For example, new genes created through various mechanisms of duplication have been reported to control centromere function in *Drosophila* (Ross et al., 2013), early development processes such as the paternal effect (Loppin et al., 2005), head-tail axis specification in the harlequin fly (Klomp et al., 2015), the developmental switch between different mouth structures in nematodes (Ragsdale et al., 2013), and viability and male fertility in *Drosophila* (Chen et al., 2010; Chen et al., 2012). In vertebrates, new genes of various ages have been found to directionally evolve with the U1–PAS axis in defining gene structure, suggesting functional evolution (Wu and Sharp, 2013), and a young noncoding RNA gene that appeared in mice (*Mus musculus)* was found to impact sperm motility and testis development (Heinen et al., 2009). These studies provide fresh insights into phenotypic and functional evolution from a new perspective of gene evolution inspiring new explorations into the mechanistic processes that determine structural and functional changes in genes involved in phenotypic development. However, the roles of new chimeric genes in the adaptive evolution of fish remain poorly understood.

The cyprinid subfamily Danioninae, including ~200 species, is one of the most species-rich subfamilies of the Cyprinidae. They are a widespread group and mainly concentrate in regions with highly diverse ecological environments extending from Africa to South/Southeast Asia (Tang et al., 2010). These fishes, living in tropical or subtropical regions, dwell within a wide variety of freshwater biotopes and can variously experience high temperatures, high levels of interspecies competition, and food resource pressure. The zebrafish (*Danio rerio*) and closely related species in the subfamily Leuciscinae (including grass carp) diverged within 54 Mya primarily in freshwater habitats and evolved significant diversity ranging from morphology to mating and trophic behaviors (Parichy, 2015). The genome sequences of these species show great genetic differentiation (Mayden et al., 2007), suggesting the rapid evolution of the genetic basis controlling their morphology, pigment patterns, and behaviors (Norton and Bally-Cuif, 2010). However, the molecular mechanisms underlying this extensive adaptability have not been well explored. Here, then, we investigate the role of new genes in the evolution of these species by genetically altering new genes in the model species. In doing so, we further explore the genetic basis of phenotypic or physiological adaption of this fish group in Southeast Asian freshwater environments.

To pinpoint evolutionary mechanisms and the phenotypic/physiological evolutionary consequences of new vertebrate genes, we identify a set of new chimeric genes in the zebrafish genome and detect a new chimeric retrogene family, named “the *chiron* family” (Chiron is a Greek mythological character with his front legs of the human and lower body of a horse). Phylogenetic analysis indicated that *chiron* genes (*chiron*s) were a Danioninae lineage-specific gene family. Interestingly, *chiron*s could integrate into the most essential biochemical pathways: the NAD^+ pathway by catalyzing the NAD^+ rate-limiting reaction. The NAD^+ pathway is so fundamental to control cell metabolism (Cantó et al., 2015). In a conventional expectation, the involved genes and pathways should not be changeable by new mutations including new genes and sequence evolution. But our data implied that new genes may reshape the essential biochemical pathways. Moreover, phenotypic effect analysis in zebrafish, experimentally developed by morpholino-mediated knockdown and CRISPR-Cas9 gene-editing experiments are all used to reveal the possible essential functions of zebrafish *chiron*s and their roles in adaptive evolution and embryo development. Furthermore, NAD^+ plays critical roles in cell survival, inflammatory responses, age-associated pathophysiology (Garten et al., 2015), and lifespan (Yoshida et al., 2019). The systematic improvement of the entire NAD^+ synthesis pathway by new *chiron*s may provide new insights into increasing the cells’ NAD^+ levels, thereby promoting the anti-aging and prolongation of the life span in the biomedical field.

RESULTS

The ways in which new genes are retained, integrate into ancestral networks, and evolve new functions remain some of the most significant areas of exploration in molecular evolution and species differentiation. To pinpoint this question, we used a similar method in our previous work (Fu et al., 2010) to detect chimeric retrogenes (Ensembl release 79). The main principle for searching the retrogenes is the identification of RNA-based duplications where one copy is a multiple exon coding gene (parental gene) whereas the second copy is a single exon coding gene (candidate retrogene). Occasionally, a portion of retrogenes could recruit novel coding sequences in flanking regions and then evolve into new chimerical retrogenes. By this method, we obtained more than 40 chimerical retrogenes (Table S1 in Supporting Information), and one of them caught our attention, which we named "*chiron*". First, the *chiron* gene has a complex exon-intron structure and it contains all the 11 coding exons of its parental gene, which is an ideal material for studying retrotransposition events. Moreover, the *chiron* is a new zebrafish-specific gene and has expanded into five duplicates (*chiron1–5*) in the zebrafish genome recently, which represents an excellent opportunity to investigate the evolutionary fate of new gene duplications. Third, its parental gene, Nicotinamide phosphoribosyltransferase (NAMPT), is a key rate-limiting enzyme in the salvage pathway of NAD^+ biosynthesis. By catalysing the NAD^+ generation, the Nampt protein plays critical roles in metabolism, cell survival, inflammatory responses, age-associated pathophysiology (Garten et al., 2015), and lifespan (Yoshida et al., 2019). Therefore, the zebrafish *nampt-chiron* gene pair provides a good opportunity for comprehensive analyses to determine the origin, sustainable evolution, and functions of new chimeric retrogenes.

Formation of the *chiron* genes in zebrafish

In zebrafish, the ancestral *chiron* gene was derived from the insertion of a retroposed sequence of the *nampt* gene in chromosome 4. It recruited upstream sequences to generate a new chimeric protein-coding structure in chromosome 11 or 18 (Figure 1A). A series of mutations, corresponding to the translation initiation site of *nampt*, occurred in *chiron* genomic structure (ATG (M)→TCA (S), marked with a red box, Figures S1 and S2 in Supporting Information), and a longer open reading frame (ORF) of *chiron* gene was formed. The ancestral *chiron* has subsequently been duplicated and expanded into five *chiron* genes with two copies (*chiron1* and *chiron2*) in a tandem array in chromosome 18 and another three copies (*chiron3*, *chiron4*, and *chiron5*) in a reverse tandem array in chromosome 11. However, the ORF of *chiron1* was interrupted by the insertion of an ~12.5-kb DNA fragment and then split into two loci (ENSDARG00000095263 and ENSDARG00000078738). This ~12.5-kb insertion DNA fragment contained many transposable elements, such as DNA2-1_DR, LTR, and SINE (Table S2 in Supporting Information). Besides, compared to the ORF of *chiron2–5*, the ORF of *chiron1* changed due to a premature stop codon (112 GAA→TAA, marked asterisk, Figure S2 in Supporting Information). Thus, *chiron1* likely evolved into a truncated gene because of the insertion of transposable elements and the formation of a premature stop codon.

Genomic structure of zebrafish *chiron*s

To detect the full length of the zebrafish *chiron*s mRNAs, we performed the 5′ RACE and 3′ RACE experiments. Based on the full-length cloning sequencing analysis (Table S3 in Supporting Information), we obtained another two variable transcript isoforms, which contained insertions or deletions at the amino acid sequence level when compared with the five annotated *chirons* transcripts (Figure S3 in Supporting Information). These transcripts shared high similarity but could be distinguished by the unique pattern of insertions and deletions in each copy. The full-length cloning sequences showed that zebrafish *chiron2–5* genes encode predicted proteins of 969 amino acid residues (aa), 1,005 aa, 1,005 aa, and 1,005 aa, respectively. Interestingly, the parental Nampt protein (~493 aa) functioned as a symmetrical homodimer (~986 aa), which was close to the predicted length of the Chiron proteins. Thus, whether the Chiron protein could self-fold into an intrinsic dimer to function as a phosphoribosyltransferase was still mysterious. Remarkably, the N-terminus of *chiron* protein contained three repeat fragments in a tandem array (red line, Figure S1 in Supporting Information), and each repeat fragment contained a short portion of the Nampt protein sequence (shown in the box, Figure S1 in Supporting Information). To investigate the origination of this 5′ flanking recruited structure, we used the tblastn algorithm to search its homologous sequence against all 63 well-assembled vertebrate genomes in Ensembl database and three other closely related species, *Cyprinus carpio* (common carp) (Xu et al., 2014), *Ctenopharyngodon idella* (grass carp) (Wang et al., 2015), and *Hypophthalmichthys molitrix* (silver carp), with the default parameters. However, no homologous sequences were detected in these groups. Thus, the 5′ flanking recruited structure was likely derived from the *de novo* material of a unique non-coding genomic sequence. Moreover, the 3′ RACE experiment showed that zebrafish five *chiron*s still contain short poly(A) tails in the genomic sequences (Figure S2 in Supporting Information), suggesting that they originated recently and the genetic characteristic of retrotransposition did not completely degrade.

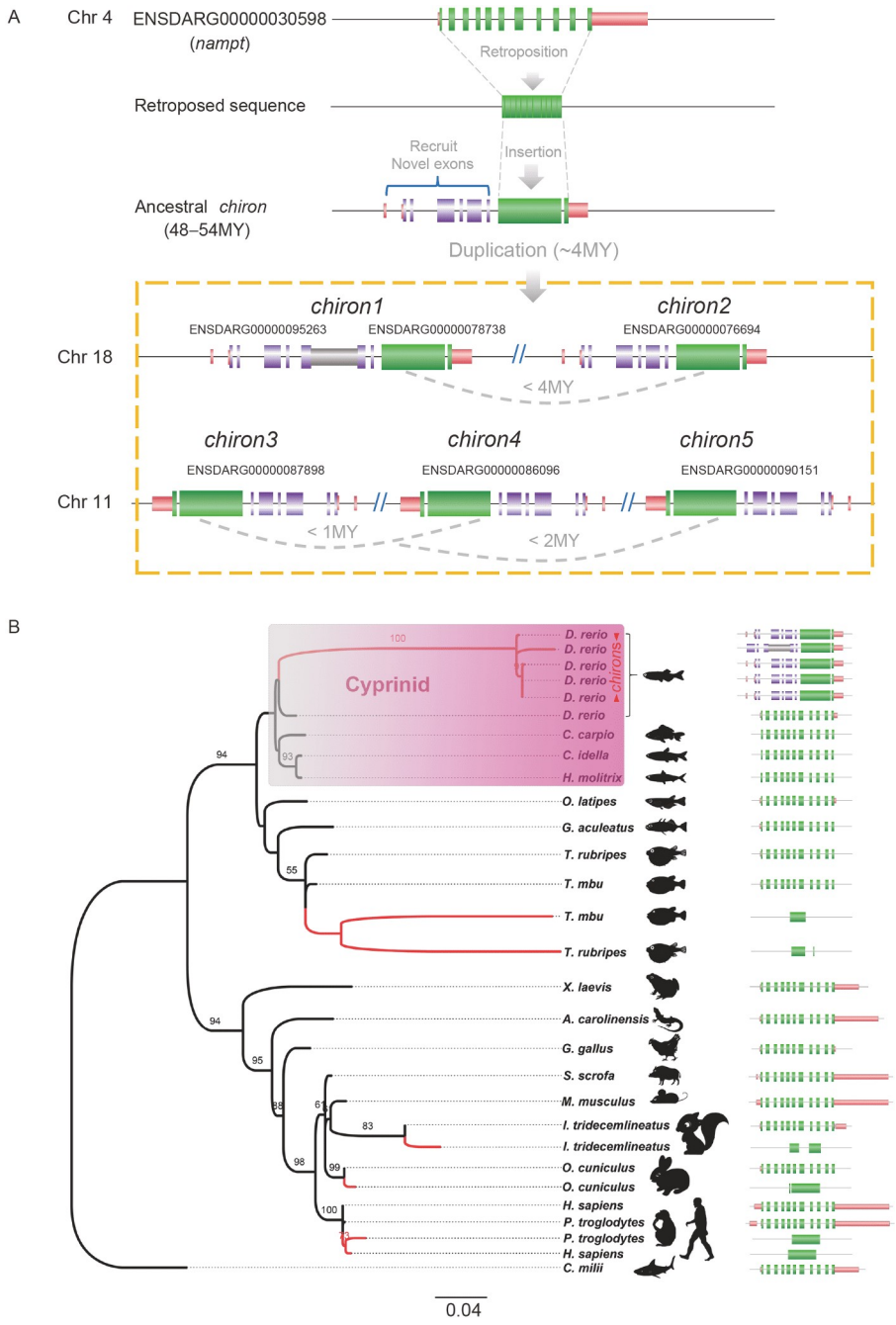

Figure 1 Origination of the *chiron* genes in zebrafish. A, Model for the origination of the *chiron* genes in zebrafish. The chimeric gene family, *chiron* family, was formed by the combined action of retroposition, *de novo* origination of 5′-flanking exons, and gene duplications. Red boxes, 5′ or 3′UTR regions; purple box, recruited protein-coding sequence; green boxes, exons; grey box, the DNA insertion region; MY, million years. The numbers above the dotted arrows indicate the estimated divergence time. B, The Maximum likelihood (ML) phylogenetic tree of *nampt* and retrocopies. The ML phylogenetic tree was performed with PhyML3.0 using the protein sequences of *nampt* and the retrocopies genes under the best fitting model (JTT+G) and bootstrap values 200 repetitions. The bootstrap support values (>50%) were above the branches. Branches corresponding to the retrocopies are in red lines. The pink box represents cyprinid fishes in Southeast Asian habitats. All species contained the parental gene *nampt*. For example, the parental gene *nampt* existed both in *Homo sapiens* (*H. sapiens*) and *Pan troglodytes* (*P. troglodytes*) but they showed very short branch length in the phylogenetic tree. Genomic structures of *nampt* homologous genes in different vertebrate species are listed. Red boxes, 5′ or 3′UTR regions; green boxes, protein-coding regions; purple box, the recruited protein-coding regions.

Finally, we identified four intact *chiron* duplicates (*chiron2–5*) and a fission *chiron1* duplicate in the zebrafish genome. Zebrafish Chiron2–5 proteins shared high identity with each other ranging from 92% to 99%. Previous structural and enzymological studies have indicated that some mutations in the mouse NAMPT protein could impair dimerization and attenuate enzymatic activity (Wang et al., 2006a). Multiple alignments of the zebrafish Chiron proteins

showed that they shared high conservation in important structural and catalytic positions with the Nampt protein (Figure S1 in Supporting Information). For example, the binding sites for nicotinamide, ribose, phosphate, and NMN (Asp16, Arg196, Arg311, Gly353, Asp354, and Gly383) were highly conserved. The important amino acid motifs surrounding the catalytic residues Tyr18, Phe193, Asp219, His247, Asp279, and Asp313 were also highly invariant (Figure 2A) (Burgos et al., 2009). Thus, long stretches of identical amino acids surrounding important structural and catalytic positions suggested that the zebrafish Chiron proteins may have been functionally conserved in their basic nicotinamide phosphoribosyl transferase enzymatic function during their evolutionary history.

The origination time and evolution of the *chiron* genes in vertebrates

The origination time of the *chiron*s in vertebrates was estimated by combining phylogenetic distribution, synonymous substitutions per synonymous site (*Ks*) calculations, and PCR-amplified experiments. First, we performed an extensive phylogenetic analysis of the *nampt*-retroposed sequences in vertebrates. The synteny-based gene-level dating strategy is very suitable to estimate the origin time for recently duplicated genes (Shao et al., 2019). It is helpful to identify new genes and works well with high-quality genomic data and good genome annotation, while the poor genome annotation, repetitive elements, and chromosome recombinations may limit the reliability of synteny. To avoid the undetectable sequence similarity caused by rapid evolution or incomplete genome, we used the tblastn algorithm that aligned Chiron protein sequences to all available high-quality genomic data in the Ensembl DNA database and three other closely related species (*Cyprinus carpio*, *Ctenopharyngodon idella*, and *Hypophthalmichthys molitrix*) using the *Callorhinchus milii* (elephant shark) genome as an outgroup. It is useful to search all *chiron* homologous sequences even if the genome contains only a small piece of homologous sequence. The distribution of *nampt*-retroposed homologous sequences showed that intact *nampt*-retroposed sequences were only detected in primates, *Oryctolagus cuniculus* (rabbit), and *Danio rerio*, while portions were detected in *Ictidomys tridecemlineatus* (squirrel), *Takifugu rubripes* (fugu), and *Tetraodon nigroviridis* (tetraodon) (Table S4 in Supporting Information). The phylogenetic tree of the *nampt*-retroposed sequence was helpful to estimate the initial divergence time of the *chiron*s. All homologous sequences were aligned with Clustal and a substitution model that best fit the data (JTT+G, G=0.566) was obtained with the ProtTest3. Then, we applied Phyml3.0 to construct a maximum likelihood (ML) gene tree. In the ML gene tree (Figure 1B), five *chiron* genes were only present in *Danio rerio* and clustered into an independent clade with long branches, suggesting that they were *Danio rerio*-specific genes and subjected to relatively high evolutionary rates.

Further, we used the *Ks* values between *nampt*-retroposed sequences and their parental genes to estimate the approximate origination time of retrocopies (Tables S5 and S6 in Supporting Information). Interestingly, we found that *nampt*-retroposed sequences in mammalian lineages (including primates, rabbits, and squirrels) originated less than 10 Mya (the synonymous substitution rate was approximately 2.1–3.7×10^{-9} mutations per site per year in mammalian genes (Hardison et al., 2003)). As we know, primates diverged from rabbits and squirrels approximately 90 Mya (Kumar and Hedges, 2011). That the expansion of these *nampt*-retroposed sequences in different mammals occurred so recently indicates that the *nampt*-retroposed sequences could originate simultaneously in different species as species-specific or lineage-specific genes. This phenomenon may be induced by a widespread burst of young retroposition events in mammals (Pan and Zhang, 2009), allowing the *nampt* genes to produce retrogenes simultaneously in different mammalian lineages. Meanwhile, the *nampt*-retroposed sequences appeared less than 40 Mya in fugu, less than 70 Mya in tetraodon, and less than 80 Mya in zebrafish ($5.7–6.4\times10^{-9}$ mutations per site per year in teleost genes (Wang et al., 2015)). As the zebrafish diverged from the fugu and the tetraodon ~230 Mya (Timetree), this result indicates that the *nampt*-retroposed sequences in fishes also originated independently in different species. In addition, *Ks* values showed that *chiron*s expanded into five duplicates (*chiron1–5*) specifically in zebrafish within 1–4 Mya, suggesting that they were new genes. Moreover, *nampt*-retroposed sequences are present independently in fishes and mammals at different evolutionary stages (in fishes <80 Mya, in mammals <10 Mya), suggesting an ongoing formation of *nampt* retrotransposition in vertebrates. This finding was consistent with the view, based on studies in plants and pigs, that retroposition events occurred continuously throughout the evolutionary history (Wang et al., 2006b; Fang et al., 2018).

New genes often experience rapid structural and sequence evolution (Long et al., 2013) and the substitution rates fluctuate in different species for a variety of factors such as environmental energy (Davies et al., 2004), physiological effects (Martin and Palumbi, 1993), generation times (Thomas et al., 2010) and sexual conflict (VanKuren and Long, 2018). Considering the absence of *chiron*s in *Ctenopharyngodon idella* and *Hypophthalmichthys molitrix* (~54 Mya), the origination time of about 68–76 Mya according to the *Ks* values was likely overestimated. Thus, we performed more detailed PCR-amplified and sequence experiments to confirm the existence of *chiron*s in 17 closely related zebrafish species. Primers were designed from the conserved exon regions of zebrafish *nampt* and *chiron*s (Table S3 in Sup-

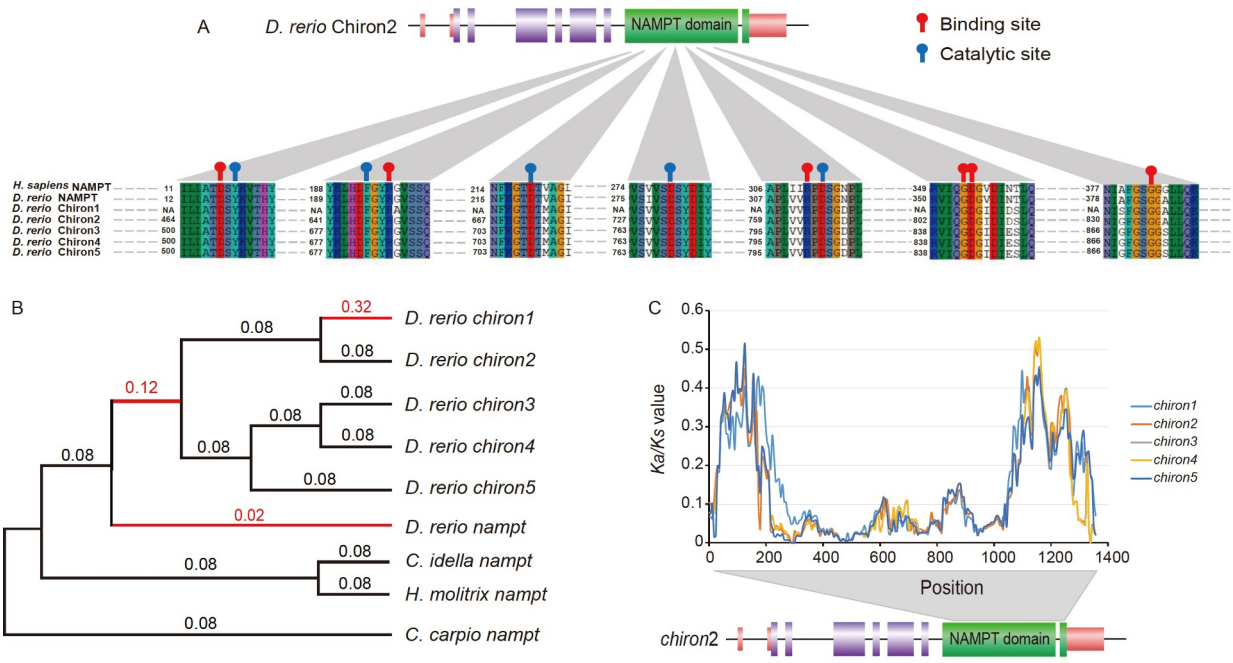

Figure 2 Functional region analysis of zebrafish nampt and chirons. A, Multiple sequence alignments of the zebrafish Chiron proteins. NA represents an unknown position because the first amino acid of *chiron*1 was unidentified. B, Gene tree of *nampt* and *chiron* genes used to test the roles of selection. The *nampt*-retroposed sequences of *chiron* genes were used to test the roles of selection. The ω values (*Ka*/*Ks*) of the optimal branch-specific model (using Method 1) are labeled above the lineages. The four-ratio model (4-RM) with four different ω ratios was the optimal branch model. C, Sliding window analysis of NAMPT domain between *nampt* and *chirons*. Each *nampt* and *chirons* pair was estimated with 120-bp windows and 6-bp slides, and ω values were calculated using the maximum likelihood method.

porting Information). The results showed that ~700-bp conserved fragments could be detected in the cyprinid subfamily Danioninae including the *Rasbora*, *Esomus*, and *Danio* lineages, but they were absent in out-groups including the *Cyprinus carpio*, *Ctenopharyngodon idella*, and *Hypophthalmichthys molitrix* (red line in Figure 3 and Figure S4 in Supporting Information), indicating that the ancestral *chiron* genes probably originated at the divergence between *Rasbora* and *Opsariichthys*. To reconfirm the absence of *chiron* genes in the outgroups, we also designed two other pairs of primers to amplify the conserved exon regions (Table S3 in Supporting Information). Likewise, no reliable PCR-amplified fragments could be detected. Together, these results suggested that *chiron* genes likely arose 48–54 Mya in the cyprinid subfamily Danioninae. Considering that the divergence time between *Danio* and *Microrasbora* was about 42 Mya, *chiron* gene only expanded into five duplicates in *Danio*-lineage within 1–4 Mya.

Further, the selection pressures between *nampt* and the *chiron* genes were estimated based on the ω values (*Ka*/*Ks*) using a small phylogenetic tree for cyprinid fishes (Figure 2B). A dynamic programming procedure that searches for the optimal branch-specific model (Zhang et al., 2011a) was performed to detect selection pressures that affected only some branches. The result showed that the 4-RM model (4 ratio models) with four different ω ratios was the optimal branch model (Table S7 in Supporting Information). In the 4-RM model, the overall values of ω were significantly lower than 0.5 (P<0.05, Figure 2B). In a conservative test, that ω was significantly lower than 0.5 indicates functional constraint on both parental genes and retrogenes (Betrán et al., 2002). Thus, this result suggested that *nampt* and the *chiron* genes have been generally subjected to severe purifying selection. Moreover, *D. rerio nampt* was subjected to stronger functional constraints after its duplication (ω=0.02, Figure 2B). The ancestral function of nampt was likely partitioned into two functions with different levels of constrain. Even so, the zebrafish *chiron*s experienced a relatively rapid evolution after initial emergence when compared to background values (ω0=0.08, ω1=0.12, P<0.05). However, ENSDARG00000078738 (3′ end of Chiron1) exhibited a higher ω value (ω=0.32), which was consistent with the previous prediction that *chiron1* likely evolved into a pseudogene and thus was subsequently subjected to a relatively relaxed constraint.

Finally, to investigate whether the selective constraints affected all sites in the NAPRTase domain region (from 453 aa to 969 aa in Chiron2) over a long period, we performed the sliding window analysis (120-bp windows and 6-bp slides) to examine functional constraints between each *nampt* and *chiron* pair by using yn00 in PAML. All domain regions of *chirons* were subjected to strict purifying selection with ω values significantly less than 1 (P<0.05) and about 98%–100% of them was less than 0.5 (Figure 2C), implying

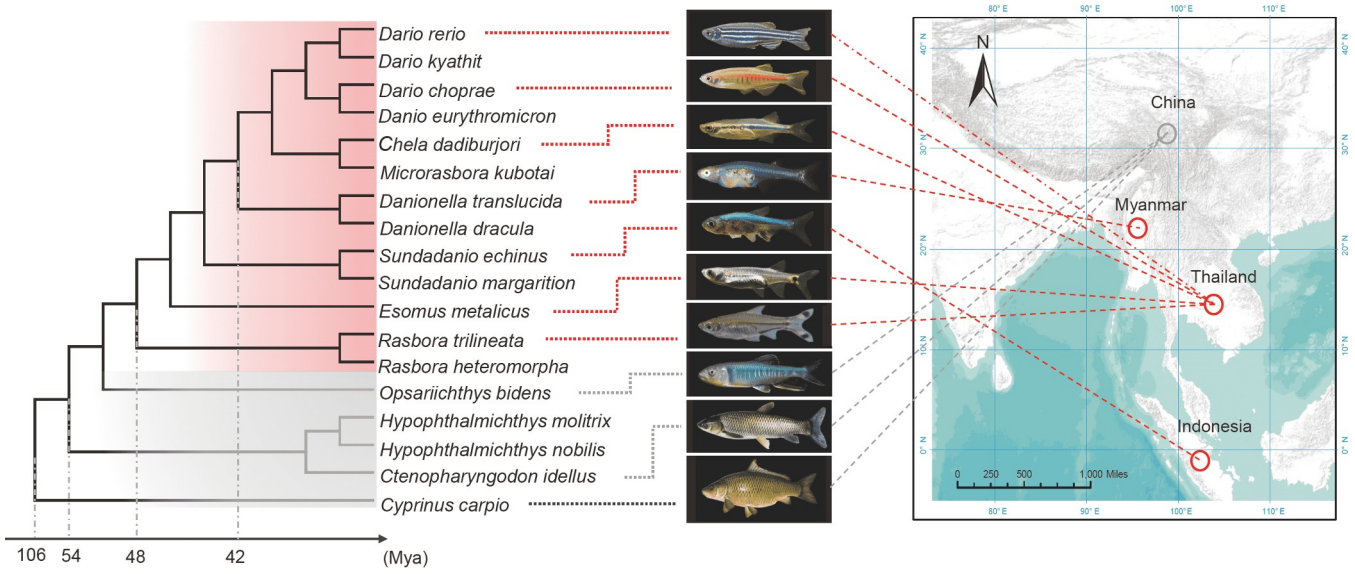

Figure 3 The origination time of *nampt*-retroposed sequences in cyprinid fishes. Red box: *nampt*-retroposed sequences can be detected by PCR-amplified experiments. Grey box: *nampt*-retroposed sequences are absent in these species. The circle in the right map indicates the living location of cyprinid fishes.

functional constraints throughout the zebrafish *chiron*s.

Expression analyses in adults and the embryo development

Gene expression profiles are ideal methods for assessing biological functional divergence of new genes (Assis and Bachtrog, 2013). We chose the zebrafish model to examine the expression profile between *nampt* and *chiron*s in adult tissues and the early embryo development stage. As the zebrafish *chiron*s shared high levels of identity in their transcript sequences, it was very difficult to identify the expression level of each *chiron* gene by PCR experiments. So, we used 53 transcriptome sequencing data of 15 zebrafish adult tissues and the unique mapping reads to assess the expression profiles (Table S8 in Supporting Information). As expected, the parental gene *nampt* was widely expressed in many tissues (orange line, Figure 4A). In vertebrates, *nampt* was a multifunctional gene in both physiology and pathophysiology and displayed a ubiquitous expression in several tissues and organs. Interestingly, the zebrafish *chiron*s were exclusively highly expressed in the testis (Figure 4A). The *chiron2* and *chiron5* genes were transcribed at higher levels, *chiron1* at a lower level, and *chiron3* and *chiron4* at an almost undetectable level. Besides, *chiron2* transcripts appeared to have a stronger expression level than that of *nampt* in the testis (P<0.01, Figure 4A).

In addition, we explored whether zebrafish *chiron*s could be expressed in early embryo development. As zebrafish *chiron*s shared high levels of identity with the parental *nampt* gene in the ORF region but varied in the 3′ UTR region, an antisense probe specific for identifying five *chiron*s by *in situ* hybridization was designed from a 3′ UTR cDNA fragment (Table S3 in Supporting Information). The antisense probe of the zebrafish *nampt* transcripts was derived from our previous work (Fang et al., 2015). The *nampt* and *chiron*s displayed similar expression patterns by *in situ* hybridization at early developmental stages. By 6 hours post-fertilization (hpf), we detected the weak expression levels of *nampt* and *chiron*s transcripts (Figure 4B1 and C1). By 12 hpf, *nampt* and *chiron*s transcripts were strongly expressed (Figure 4B2 and C2). By 24 hpf, *nampt* and *chiron*s transcripts were widely expressed across many tissues (Figure 4B3 and C3), with enhanced levels of expression in the anterior head region, including the brain, eyes, otic vesicles, and somite boundaries. By 48 hpf, *nampt* and *chiron*s transcripts were still predominantly expressed in the anterior head region, somite boundaries, and pectoral fins (Figure 4B4 and C4). The expression of *chiron*s transcripts in the body reduced by 72 hpf but could still be detected in the eyes, midbrain-hindbrain boundary, and pectoral fins. The *nampt* transcripts were still predominantly expressed in the anterior head region, eyes, and pectoral fin buds (Figure 4B5 and C5).

Regulating the NAD^+ levels in the salvage pathway

The parental gene *nampt*, a rate-limiting enzyme in the salvage pathway of NAD^+ biosynthesis, broadly exists in both prokaryotic and eukaryotic organisms. Maintenance of an adequate NAD^+ pool is essential for cell survival and function (Cantó et al., 2015). By regulating the NAD^+ level, the *nampt* gene can also increase the activity of NAD-dependent enzymes, such as sirtuins, poly(ADP-ribose) polymerases, and CD38/157. It thereby plays pleiotropic functions in cellular metabolism, mitochondrial biogenesis, and adaptive responses to inflammatory, oxidative, proteotoxic, and gen-

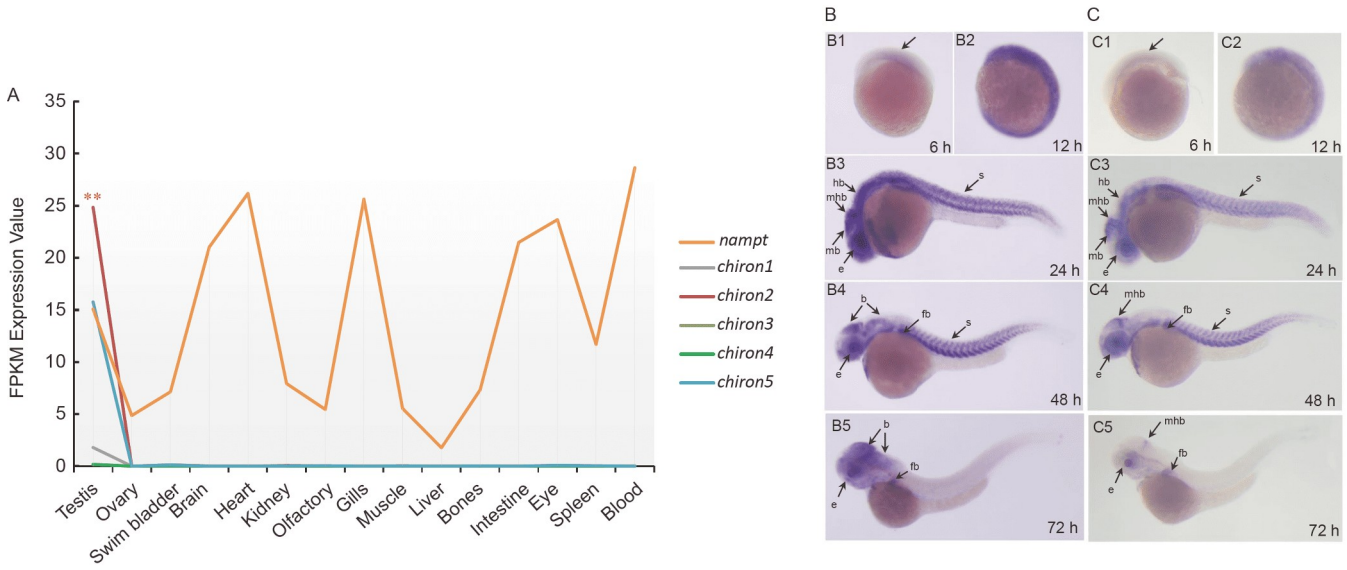

Figure 4 Expression of zebrafish *nampt* and *chirons* transcripts in adult tissues and embryo development. A, Expression levels of zebrafish *nampt* and *chirons* transcripts in 15 adult tissues. Relative gene expression was estimated using the FPKM expression values. The relative expression values are averaged from 53 sets of transcriptome data. The zebrafish *chirons* were specially expressed in the testis. B, Spatial-temporal expression of zebrafish *nampt* transcripts. C, Spatial-temporal expression of zebrafish *chirons* transcripts. e, eye; b, brain; mb, midbrain; hb, hindbrain; mhb, midbrain–hindbrain boundary; s, somite; fb, fin buds.

otoxic stress (Garten et al., 2015). The highly conserved protein structure with *nampt* implied that the zebrafish *chirons* were likely functionally conserved in the NAD^+ rate-limiting enzyme function. To investigate their roles in the NAD^+ rate-limiting enzyme, we transfected two zebrafish recombinant plasmids (pCMV-tag 2C-Nampt and pCMV-tag 2C-Chiron) and a control plasmid (pCMV-tag 2C) into HEK 293T cells respectively. The cultured cell samples were collected after 30 h for Western blot analysis and NAD^+/NADH assay. Western blot analysis showed that both the recombinant pCMV-tag 2C-Nampt and pCMV-tag 2C-Chiron proteins were expressed in the experimental groups, and the molecular weights were approximately 55 kD and 110 kD respectively (shown in the box, Figure 5), corresponding to previous predictions. Then, we detected the NAD^+ and NADH levels in both control and experimental groups and calculated the average NAD^+/NADH ratios. The average NAD^+/NADH ratios in the control, pCMV-tag2C-Nampt, and pCMV-tag2C-Chiron groups were about 1.11, 1.70, and 1.64, respectively (Figure 5). Compared to the control groups, the NAD^+/NADH ratios of the pCMV-tag2C-Nampt and pCMV-tag2C-Chiron groups increased significantly by 53.2% ($P<0.05$) and 47.8% ($P<0.05$), respectively. However, the NAD^+/NADH ratios showed no significant difference between these two experimental groups. The increase in the average NAD^+/NADH ratios in the two experimental groups was due to the upregulation of NAD^+ levels. Thus, these results suggested that zebrafish *chirons* were functionally conserved in basic NAD^+ rate-limiting enzymatic function and could regulate NAD^+ levels in the salvage pathway.

Essential phenotypic effects on early zebrafish embryonic development

Considering that zebrafish *chirons* contained five duplicates, it seems infeasible to knock out all the copies at once by using TALEN or CRISPR-Cas9 gene-editing systems. We found that *chrions* shared highly conserved sequences in the ORF region, especially in translation initiation sites (Figure S2 in Supporting Information), so we used a morpholino antisense oligo to investigate the phenotypic effect of *chirons* in zebrafish. The translation-blocking morpholinos could specifically bind to the translation initiation sites of all the *chirons* mRNA molecules to block their translation but do not cause degradation of their mRNA targets. This method has been used extensively to study loss-of-function phenotypes during the early stages of zebrafish embryogenesis (Nasevicius and Ekker, 2000; Blum et al., 2015). The specific translation-blocking morpholinos were designed from Gene Tools (Philomath) and they could knock down the translation level of all *chiron* members but not impact *nampt* gene function (Figure 6A). First, we evaluated the efficiency and specificity of the *chirons* morpholinos (Chiron-MO). Two groups of mixtures (Chiron-MO/wild Chiron-GFP recombinant plasmid mixture (Chiron-GFP-WT) and Chiron-MO/mutant Chiron-GFP recombinant plasmid mixture (Chiron-GFP-MT)) were separately injected into the zebrafish embryos at the one-cell stage. The Chiron-MO (3 ng per embryo) successfully blocked the expression of Chiron-GFP-WT (Figure 6B1), suggesting that Chiron-MO morpholinos were efficient. Also, Chiron-MO did not block the expression of the mutant Chiron-GFP-MT (Figure 6B2),

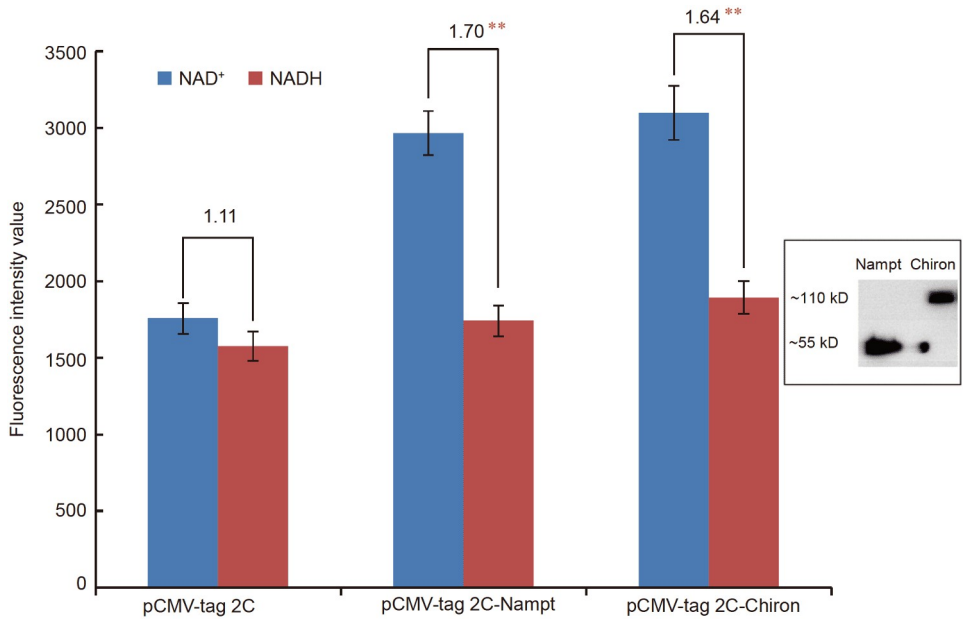

Figure 5 Average NAD^+/NADH levels and corresponding Western blot detection. The columns indicate the mean fluorescence intensity values. The error bars indicate standard errors. The average NAD^+/NADH ratios are above the line. **, $P<0.01$ compared to the control group. The corresponding Western blot detection is shown in the box, and the molecular weights of recombinant pCMV-tag 2C-Nampt and pCMV-tag 2C-Chiron proteins are approximately 55 kD and 110 kD, respectively.

suggesting that the Chiron-MO morpholinos were specific. Then the proper concentration of Chiron-MO and standard-MO were injected into zebrafish embryos at the one-cell stage separately. Our previous study showed that *nampt* played an essential role in zebrafish in early embryonic development (Fang et al., 2015). In this study, we also detected that embryos injected with Chiron-MO could display definite defects (Figure 6C). By 36 hpf, the Chiron-MO morphants showed significantly inflated hindbrain ventricles and abnormal brain intervals compared with those of the control group (embryo injected with standard-MO). At 72 hpf, the morphants showed slower development, including smaller eyes, shorter tails, narrower lateral, and hydrocephalus, which were consistent with the expression pattern of *chirons* RNAs. After 5 days post-fertilization (dpf), the morphants exhibited more abnormal defects and sank to the bottom without swimming, ultimately leading to embryonic lethality. To increase the credibility, we performed the mRNA rescue experiments. Two mixtures (Chiron-MO alone and Chiron-MO/mutant *chiron2* mRNA) were separately injected into embryos at the one-cell stage. Co-injection with ~300 pg of *chiron2* mRNA could partially rescue the mutant phenotype induced by Chiron-MO (Figure 6D). The reason was probably that we knocked down all *chirons* while only adding *chiron2* mRNA back. Even though, the morphant rate was significantly reduced in mRNA rescue groups, demonstrating a direct relationship between the morpholino-mediated knockdown of *chiron*s and the lethal effect of the mutant phenotype. Thus, the morpholino-mediated knockdown data showed that *chiron*s likely played an essential role in zebrafish embryo development.

To further verify the essential function of zebrafish *chiron*s, we also performed a CRISPR-Cas9 gene-editing experiment in zebrafish. Although this experiment could not ideally knock out five *chiron* genes at once, the random removal of *chiron* duplicates at least would increase the understanding of their function at the dose level. The CRISPR sgRNA was designed to target the conserved NAmPRTase domain of zebrafish *chiron*s in exon 8 (Figure 6A), which could specifically target all *chiron* duplicates but not the *nampt* (Figure S2 in Supporting Information, blue box). Different suitable amounts of *chiron*s sgRNA or non-targeting control sgRNA (60–80 pg/cell) were mixed with Cas9 mRNA and then were co-injected into one-cell-stage wild type embryos respectively. Mutation screens of F1 embryos by Sanger sequence showed that *chiron*s sgRNA was valid and the mutation efficiency was about 40% (Figure S5A in Supporting Information). We sequenced 60 clones for F1 embryos and identified 29 mutations (deletions and insertions), indicating that *chirons* were knocked out (Figure S5B in Supporting Information). Then, we measured the survival rates in both control and experimental groups. The multiple copies of *chiron* genes may influence total mutation efficiency. Even though, the survival rates of the sgRNA-*chiron*s groups significantly decreased when compared to the sgRNA-control group by 48 pfh ($P<0.05$, Figure 6E). Moreover, injection with 80-pg sgRNA-*chiron*s could induce a lower survival rate than that of 60-pg sgRNA-*chiron*s ($P<0.05$), suggesting a dose-dependence manner (Figure 6E). Thus, this CRISPR-Cas9 assay was consistent with the

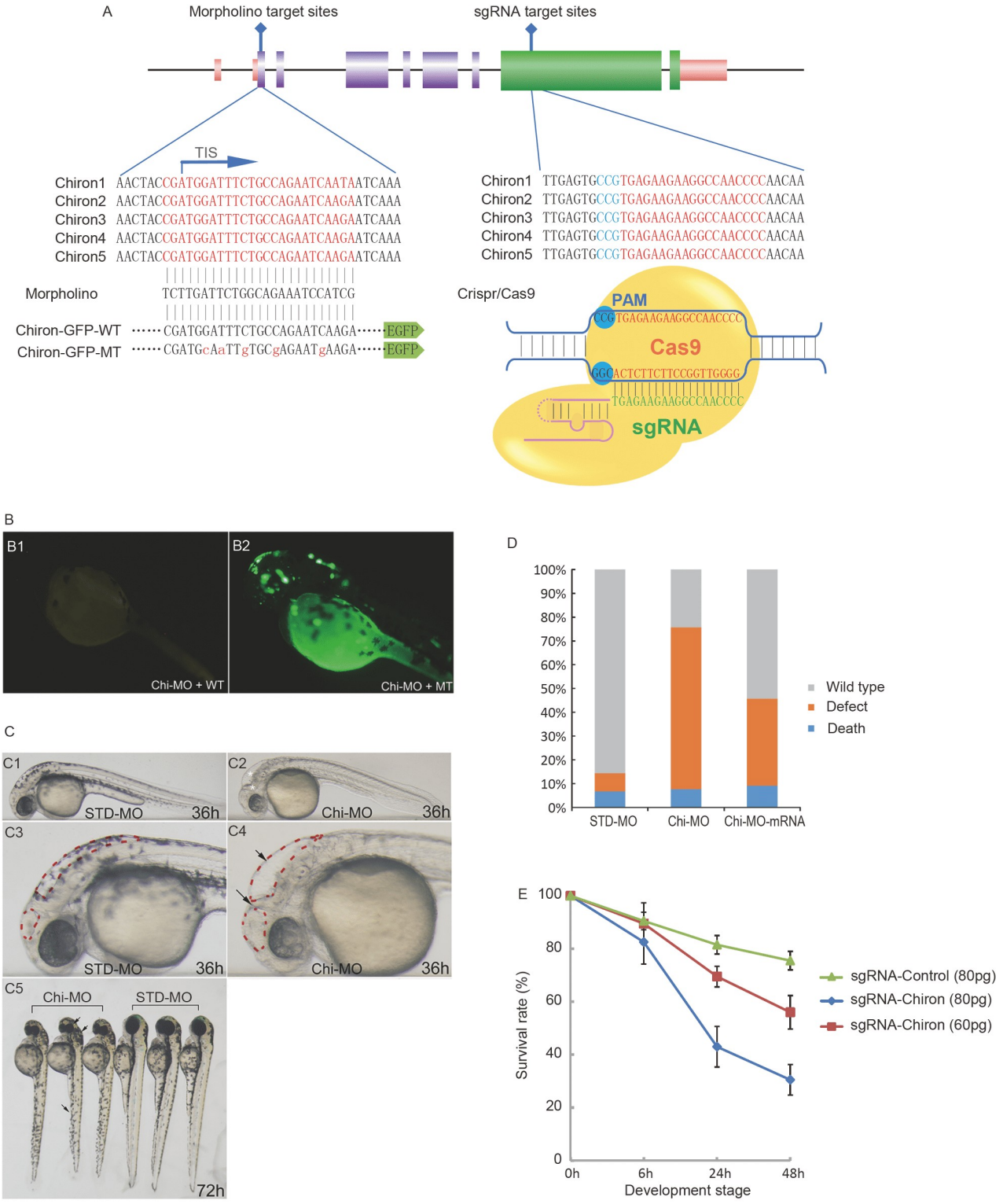

Figure 6 The essential function of *chiron* genes during zebrafish embryo development. A, Morpholinos and sgRNA target sites. Red sequences indicate morpholinos or sgRNA target sites; TIS, translation initiation sites. B, Efficiency and specificity of the *chiron*s morpholino. B1, Chi-MO+WT: embryos were injected with Chiron-MO and a wild Chiron-GFP-WT fusion protein expression vector. B2, Chi-MO+MT: embryos were injected with Chiron-MO and a mutant Chiron-GFP-MT fusion protein expression vector. C, Morphology of the *chiron*s knockdown embryos. C1 and C3, morphants of STD-MO at 36 hpf. C2 and C4, morphants of Chi-MO at 36 hpf. C5, morphants of Chi-MO and the STD-MO at 72 hpf. D, mRNA rescue experiments. Zebrafish embryos were injected with morpholinos and *chiron2* mRNA and scored at 72 hpf: STD-MO (3 ng), Chi-MO (3 ng) and Chi-MO (3 ng) plus Chi-mRNA (300 pg). Grey box: wild type indicates embryos with no obvious defects; orange box: defect type indicates embryos with a reduced head and curvature of body shape; blue box: dead embryos at 24 hpf. E, Survival rates of *chiron*s CRISPR-Cas9 gene-editing embryos.

morpholino-mediated knockdown experiments, demonstrating that zebrafish *chiron*s may play an essential function during the early embryo development.

How the emergence and evolution of *chiron*s are correlated with other members of the NAD^+ synthesis pathway

As we know, the rate of NAD^+ synthesis is largely determined by the first step in the salvage pathway that converts NAM to NMN by NAMPT. The NAD^+ molecule is mainly produced from one of three of the following different pathways (Denu, 2007; Verdin, 2015): (1) the *de novo* synthesis pathway—building the niacinamide molecule from scratch starting from L-tryptophan (Trp), (2) Preiss-Handler pathway—producing it from niacin (nicotinic acid (NA)), or (3) the salvage pathway—synthesizing it from NAM, NR and NMN (Figure 7A). However, no matter which of these methods is used to produce NAD^+ the first time, the salvage pathway is the dominant way to re-produce NAD^+ after it is involved in the cell metabolic process (Denu, 2007). The salvage pathway is a stable and vital hub NAD^+ biosynthetic route which is widely present and expanded in eukaryotes. However, the *chiron* genes could integrate into this ancestral hub network to acquire their corresponding biological roles by regulating the NAD^+ levels. Thus, it is valuable to inspect whether *chiron* genes could impact or even reshape the up- and downstream genetic evolution of the NAD^+ biosynthetic pathway after their emergence in the cyprinid subfamily Danioninae fishes.

To avoid mutation saturation, we used tblastn to search for the homologous genes of three NAD^+ pathways against the genomes of *Danio rerio* and 10 other closely related species that diverged within 100 Mya (*Anabarilius grahami*, *Carassius auratus*, *Cyprinus carpio*, *Danionella dracula*, *Danionella translucida*, *Hypophthalmichthys molitrix*, *Hypophthalmichthys nobilis*, *Oxygymnocypris stewartii*, *Leuciscus waleckii*, and *Pimephales promelas*). But in these tblastn results, we did not find any QAPRT homologs in the zebrafish genome. QAPRT could connect the *de novo* synthesis pathway and the Preiss-Handler pathway by transforming quinolinic acid (QA) into NAMN (Figure 7A). We, therefore, recovered the other six genes that were involved in the Preiss-Handler pathway (*nadsyn*, *naprt*, *nmnat1*, *nmnat2*, and *nmnat3*) and salvage pathway *(nmnat1*, *nmnat2*, *nmnat3*, and *nampt*) to investigate their selection pressures and positively selected sites in cyprinid subfamily Danioninae fishes (Figure 7B, Table S9 in Supporting Information).

Interestingly, we detected that three genes (*nmnat1*, *naprt*, and *nampt*) displayed significantly lower *Ka*/*Ks* ratios in the cyprinid subfamily Danioninae (*branch* model in PAML with a lower *Ka*/*Ks* ratio, $P<0.05$, Figure 7A, red genes, Table S10 in Supporting Information). All lower *Ka*/*Ks* values suggested that these genes in Preiss-Handler pathway and salvage pathway were likely subjected to more functional constraints after the insertion of *chiron* genes. Furthermore, since overall strong negative selection could mask positive selection at a few individual codon sites, we used *branch-site* models to test the potential for positive selection at specific sites in the subfamily Danioninae. Interestingly, we found that two protein genes, Nmnat-1 (corresponding to human *NMNAT1* S135, Figure 7B) and Naprt (corresponding to human *NAPRT* G459, Figure S6 in Supporting Information), have been subjected to positive selection in the subfamily Danioninae (Table S11 in Supporting Information). Then, we conducted three-dimensional (3D) structure simulations to examine the possible effects of these positively selected sites on the enzyme structure by using Phyre2 (Kelley et al., 2015). According to the crystal structure of NMNAT-1 (1gzu.1.A), S135 was located within an unresolved external loop (Figure 7C2 and C3, red site in the grey circle), which was well suited to mediate the interactions of NMNAT-1 with other proteins, such as the nuclear import proteins, phosphorylated kinase or PARP-1(Garavaglia et al., 2002; Berger et al., 2007). In the human NMNAT-1 protein, this loop also contained a nuclear localization sequence (amino acids 123–129) and the well-identified and important phosphorylation site of the residing serine, S136 (Figure 7C2 and C3, orange site). NMNAT-1 was a specific activator of PARP-1, and phosphorylation at S136 could reduce the interaction of NMNAT-1 and PARP-1 to regulate PARP-1-dependent functions (Berger et al., 2007). Besides, Danioninae Nmnat-1 also showed a rapid accumulation of mutations in this region with high variability and a 5-aa deletion, implying some adaptive variation.

In the Danioninae Naprt protein, a positive selection position G459 was also detected (Figure S6 in Supporting Information). The Naprt protein could convert Nicotinic acid (niacin or NA) into NA mononucleotide (NaMN), which would then be converted into NA adenine dinucleotide (NaAD) and finally to NAD^+. The G459 site is located at the bottom of the open-faced sandwich domain and on the enzyme surface (Figure S6 in Supporting Information). Although no evidence of its involvement in catalysis or substrate binding exists, the structural variation site G459 in the protein surface may fulfill unknown functions of an extracellular form of hNaPRTase (Marletta et al., 2015).

Considering that *chiron* genes have duplicated several times especially in zebrafish 1–4 Mya, we consequently investigated the impact of *chiron* genes expansion in the zebrafish NAD^+ biosynthetic pathway. As we speculated, a positive selection site (corresponding to human *NAMPT* C151) was detected in zebrafish Nampt (Figure 7D1, Table S12 in Supporting Information). The C151 site, located in the interface (Figure 7D3, marked in yellow) of the Nampt

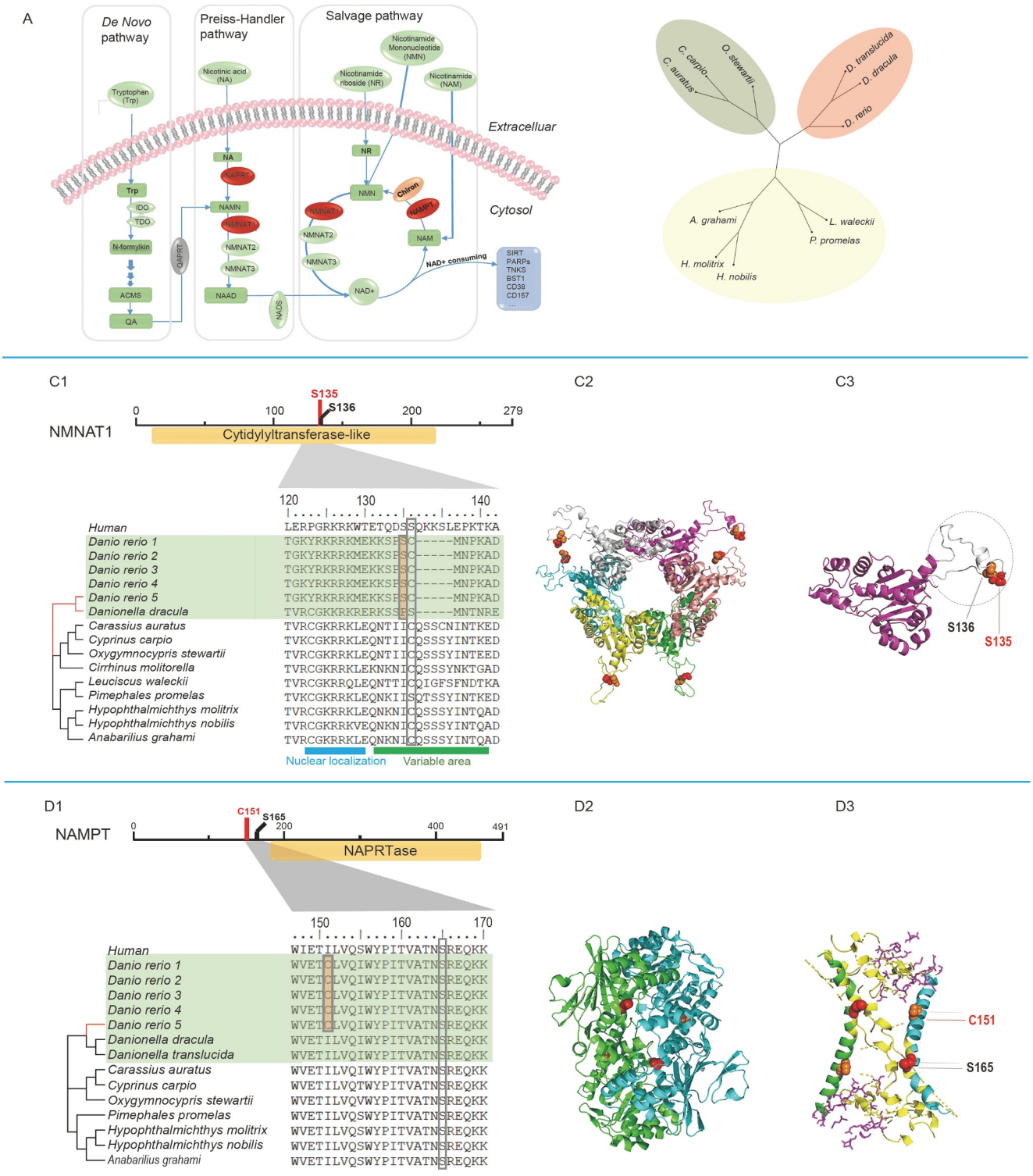

Figure 7 How the emergence and evolution of *chirons* are correlated with other members of the NAD$^+$ synthesis pathway. A, Three NAD$^+$ synthesis pathway network. The NAD$^+$ molecule is mainly produced from one of three of the following different pathways: the *de novo* synthesis pathway, Preiss-Handler pathway or the salvage pathway. The red gene represents positively selected genes. QAPRT is not present in *Danio rerio*. The *chiron* gene family inserts into NAD$^+$ salvage pathway. B, Phylogeny tree for selection pressure test. The red circle represents the cyprinid subfamily Danioninae. C, Sequence alignment and 3D structure of NMNAT1 hexamer. Each of the six subunits in the hexamer is colored differently. The red box and red sphere indicate the positively selected site in Danioninae (S135). The orange sphere indicates the important phosphorylation site resides in the human (S136). D, Sequence alignment and 3D structure of NAMPT dimer. Each of the six subunits in the hexamer is colored differently. The red box and red sphere indicate the positively selected site in *Danio rerio* (C151). The orange sphere indicates the important phosphorylation site resides in the human (S165). The interface is marked in yellow and the pocket of active sites of the opposite dimer marked in the purple stick.

protein structure, may impact the dimer cohesion (Figure 7D2, red site). Moreover, the C151 site also lies at the same α-helix structure as S165 site (Figure 7D2, orange site), where the S165F/Y mutation could impact the PRPP binding site and stimulate PARP-1 by unwinding a helix in the Nampt protein active site 380GRS motif (Wang et al., 2014). Although the C151 was not close to this 380GRS motif, it was spatially close to the pocket of active sites in the opposite dimer (Figure 7D3, marked by the purple stick), which may influence the interaction of Nampt with other proteins.

To validate whether these positive selection sites were fixed in zebrafish, we re-analyzed four other resequenced *Danio rerio* genomes (GCA_000767325.1, GCA_001487895.1, GCA_001483285.1, and GCA_000002035.4). We found that all positive selection sites (S135 in Danioninae Nmnat-1, G459 in Danioninae Naprt, and C152 in zebrafish Nampt) and nearby amino acid sites had highly shared identities in these individuals (Figure 7C1 and 7D1, Figure S6 in Supporting Information, *Danio rerio* 1–5), indicating a stable genetic basis.

DISCUSSION

The origination and subsequent evolution of novel genes with novel functions are considered a major contributor to genetic and phenotypic diversity in organisms (Long et al., 2013). Understanding how new genes are formed and subsequently evolve to become critical components of genetic systems is crucial to recognizing the genetic basis for the origin and evolution of novel phenotypes and biological diversity. Previous data detected the widespread presence of chimeric genes as genomic novelties in various metazoan organisms, from fruit flies, fish, to mammals (Rogers and Hartl, 2012; Liu et al., 2018). The tissue-specific expression patterns and physiological roles of these genes suggested their novel functions for the adaptive evolution of the organisms. However, the direct evidence remains limited for the protein functions of these chimeric genes, with a few exceptions (e.g., a systematic study of the origin, evolution, and function of *Jingwei* (Zhang et al., 2004; Zhang et al., 2005; Zhang et al., 2010)), although the RNA transcripts were usually identified from the new genes. Our analyses have revealed a lineage-specific chimeric gene family, the *chiron* family, which was generated by a recent insertion of retroposed sequences into a new genome locus from the parental copy of *nampt*, with subsequent duplications. Our study provides the first evidence that the chimeric proteins translated from *chiron* genes have essential functions in zebrafish development, as demonstrated by our morpholino translation-blocking and CRISPR-Cas9 gene-editing data.

The expression analyses of *chirons* showed that they were expressed in the testis and embryogenesis. In the testis, widespread demethylation of CpG in promoter regions could increase transcriptionally active chromatin states, leading to the promiscuous transcription of functional genes and pseudogenes (Kaessmann, 2010). Thus, many new genes originated and evolved novel functions in the testis due to the testis tolerance of spurious transcriptional initiation and the subsequent influence of natural selection (Kaessmann, 2010). Nevertheless, the significantly higher expression level of *chiron2* in the testis implied that zebrafish *chiron*s were likely expressed in the testis as functional chimeric genes. Moreover, during the early embryo development stages after the midblastula transition (MBT), post-fertilization chromatin states are organized by dynamic changes in epigenetic modifications, including histone modifications and DNA methylation, and consequently, widespread promiscuous transcription is initiated. A similar pattern in the testis may facilitate the origination and expression of new genes (Xu et al., 2018). Whether they first evolved expression in the testis or embryo needs a further comparative analysis with the ancestral expression status of *chiron* reconstructed. Unlike DNA duplication, however, the RNA-based duplication generally retained only its parental ORF region but lost the corresponding promoter region. Thus, this widespread distribution of *chiron*s transcripts in embryonic development implies that new genes may evolve complicated promoters and be recruited for functions in embryonic development. Together, these results may imply that zebrafish *chiron*s likely originated in embryonic development and gradually evolved extensive expression patterns in the testes, indicating that their expression breadth was increasing and that they were in the progress of evolving broader functions.

Retrotransposition has been described as having an essential role in the origination of new genes, and several models of molecular evolution can account for the functional roles of gene duplication, such as the evolution of redundancy (Engel et al., 1970), subfunctionalization (Force et al., 1999), neofunctionalization (Lynch et al., 2001) and Innovation–Amplification–Divergence (IAD) model (Bergthorsson et al., 2007). Our previous important protein sites and NAD^+ rate-limiting enzyme analysis suggest that *chiron* is likely maintained as a result of gene redundancy. However, *nampt* is a multifunctional gene, and complete redundancy over long evolutionary periods is unlikely because mutational pressure will ultimately result in the non-functionalization of *chiron*. Combining selection analyses, expression patterns, and survival fitness, we propose that the retention of *chiron* may be a consequence of IAD model. This hypothesis is based on several pieces of evidence, as follows. There is an abundance of regulatory elements within the *nampt* 5′ flanking region (Ognjanovic et al., 2001), which enables widespread protein distribution in different development stages and tissues and diverse roles in physiology and pathophysiology (Yoshino et al., 2011; Dahl et al., 2012).

However, unlike the DNA duplication, RNA-based retrotransposition only contains the parental ORF region but lost the promoter region. The 5′ recruited segment of *chiron* originates *de novo* from noncoding regions of DNA, driving its divergent expression patterns with *nampt* both in early embryo development and adult tissues. This result indicates that ancestral *chiron* has a weak secondary activity (innovation). Then, the ancestral *chiron* has experienced the amplification process and expanded into five duplicates (*chiron1–5*) specifically in zebrafish 1–4 Mya. The *chirons* are set to accumulate mutations that provide enzymatic divergence of different copies. Considering the insertion of transposable elements, the formation of a premature stop codon and a lower expression level, *chiron1* is likely either subject to neofunctionalization due to fission or pseudogenization. The *chiron*2 evolves into an altered gene express localization in the testis. Selection may have maintained *chiron*3–5 due to the gene dosage effect, which needs further investigation. Ultimately, in the *nampt-chirons* duplication, nampt may encode the parental activity and *chirons* may provide an improved role for fitness.

The translation-blocking morpholino was an important method to study loss-of-function phenotypes during the early stages of zebrafish embryogenesis (Nasevicius and Ekker, 2000; Blum et al., 2015). Injection of Chiron-MO resulted in abnormal morphants including slower development, smaller eyes, shorter tails, narrower lateral, and hydrocephalus, which were similar to the effects of MOs that were mediated through p53-induced apoptosis. One possibility is that unspecified off-target effects of MOs result in the induction of a p53-dependent cell apoptosis pathway (Robu et al., 2007). Nevertheless, the coinjection (wild Chiron-GFP or mutant Chiron-GFP) and RNA rescue experiments indicated the efficiency and specificity of the *chiron*s morpholinos. The other one is that *chiron* could regulate p53 and functionally induced apoptosis. The enzyme activity experiments showed that *chiron*s were functionally conserved in basic NAD^+ rate-limiting enzymatic function and could regulate NAD^+ levels. NAD^+ could activate NAD^+-dependent deacetylase SIRT1 activity, which can directly inhibit the transcription activity of p53 (Pfister et al., 2014; Yang et al., 2017). Suppression of SIRT1 activity could increase p53 acetylation and induce cardiomyocyte apoptosis (Mu et al., 2014). In addition, p53 proteins also required NAD^+ as a co-factor. NAD^+ could bind to p53 tetramers and change the conformation to prevent its binding to DNA, which will reduce p53-mediated transcriptional activity and suppress the expression of its downstream proteins (McLure et al., 2004). Thus, *chiron* may catalyze the NAD^+ levels, regulate p53 expression, and functionally modulate cell apoptosis. Considering that CRISPR-Cas9 assay was consistent with the morpholino-mediated knockdown experiments, we suggested that zebrafish *chiron*s may play an essential function during the early embryo development. In traditional evolution-developmental biology, embryonic development is based on a conserved genetic basis and any change would have many adverse consequences (Roux and Robinson-Rechavi, 2008; Domazet-Lošo and Tautz, 2010; Kalinka et al., 2010; Piasecka et al., 2013). However, by RNA interference (RNAi), Chen et al. proved that new genes could play lethal roles in diverse early development stages in *Drosophila* (Chen et al., 2010). In human, Zhang et al. also identified many new genes which showed high expression levels in fetal brain development, indicating their important roles in early development (Zhang et al., 2011b). In *Xenopus*, many protein-coding genes were significantly enriched for expression in early developmental stages and a young orphan gene (Fog2) acted as an essential functional factor in morphological development (Xu et al., 2018). Together with our evidence in fish, we proposed that new genes could play a crucial role in embryonic development which may enrich the evolution-developmental theory.

Previous studies have shown that individual new genes can participate in local ancestral networks and acquire important functions in fruit flies (Zhang et al., 2010; Chen et al., 2012), yeast (Capra et al., 2010; Li et al., 2010), mammals (Zhang et al., 2015; Berto and Nowick, 2018; Zu et al., 2019), and plants (Matsuno et al., 2009; Weng et al., 2012; Zhang et al., 2020). However, whether these new genes could regulate the essential biochemical pathways is rarely studied. The NAD^+ pathway is a fundamentally important and conserved pathway that is widely present in both prokaryotic and eukaryotic organisms to control cell metabolism and cell survival (Dong et al., 2009; Dahl et al., 2012; Garten et al., 2015). In a conventional expectation, this conserved pathway should not be changeable by new mutations including new genes and sequence evolution, due to the fundamentally important function of the pathway. Here, we found that new *chiron* genes could integrate into the ancestral NAD^+ biosynthetic pathway and evolved into an essential gene by regulating NAD^+ levels. Although this study does not explain the detailed mechanism, it is great to raise a new and good scientific problem for future research about how new genes functionally impact fundamentally important and conserved pathways. In addition, it is intriguing to ask whether new genes could impact the evolution of other genes in the ancient hub networks after their integration. Our results showed that *chiron*s were subjected to rapid evolution and strictly functional constraints. The *chiron*s are not only functionally integrated into the NAD^+ biosynthetic pathway, but also probably impact the evolution of the entire network of both the up- and downstream genes. Positive selection at special sites followed by functional constraints on *naprt*, *nmnat1*, and *nampt* genes may have systematically improved the efficiency of the entire NAD^+ biosynthetic pathway.

The cyprinid subfamily Danioninae, containing approxi-

mately 50 genera and more than 300 recognized species (Tang et al., 2010), is one of the most diverse groups and is mainly distributed throughout ecological environments from India to Southeast Asia. This group displays a broad array of significant phenotypic diversity ranging from morphology to mating and foraging behaviors (Fang 2003). This group also includes an important model organism, the zebrafish, which thus provides an excellent system for mechanistically analyzing the genetics and evolution of this fish group. Previous reports showed that the wild zebrafish could survive at elevations of 8–1,576 m and in a range of water conditions, including temperatures ranging from 12 to 39°C, pH levels of 5.9–9.8, and salinities of 0.01–0.8 (Parichy, 2015). This vast variation indicates a huge degree of adaptation to diverse environments. However, the molecular mechanism underlying the Danioninae adaptation has been seldom understood. The phenotypic diversity in species closely related to zebrafish warrants investigation into the mechanism of molecular evolution and its genetic basis through comparative studies among related species. Interestingly, we identified a Danioninae-specific novel chimeric gene family, the *chiron* family, which could regulate NAD^+ levels in the NAD^+ salvage pathway and reshape the entire NAD+ biosynthesis pathway. NAD^+ is a vital cofactor with a wide range of effects in energy metabolism, Sirtuins/PARPs activation, mitochondrial fitness, and so on (Garten et al., 2015). One of the most prominent functions of the NAD^+ is the regulation of the activity of metabolic pathways, such as glycolysis in the cytoplasm and the TCA cycle/oxidative phosphorylation in the mitochondrion. NAD^+, as a coenzyme, could influence metabolic efficiency in most metabolic pathways. For example, the increment of NAD^+ levels was involved in the improvement of mitochondrial function under stress, protection against dietary, and age-related diabetes (Johnson and Imai, 2018). In addition, NAD^+ levels increased in response to energy stresses, such as glucose deprivation, fasting, caloric restriction, and exercise (Garten et al., 2015). Most Sirtuins are activated by NAD^+ levels during energy deficits and periods of reduced carbohydrate-based energy, which would trigger cellular adaptations and improve metabolic efficiency (Boutant and Cantó, 2014). Further, NAD^+-mediated CD41 T-cell differentiation dramatically regulated innate and adaptive immune responses through marrow-derived mast cells (MCs) and significantly affected physiologic functions of protection against bacterial infection (Biefer et al., 2018). Thus, by regulating the NAD^+ levels, the *chiron* genes may represent a physiologically important homeostatic mechanism, which could constantly maintain an adequate cellular NAD^+ pool and further increase the adaption of Danioninae tribes to diverse ecological environments, especially in the critical period of food shortage and energy deficiency. Thus, we propose that this phenotypic and physiological adaptive evolution may improve the survival and prosperity of subfamily Danioninae fishes in Southeast Asian habitats.

In summary, we identified a novel chimeric gene family, the *chiron* family, and explored the genetic basis underlying the adaptive evolution of subfamily Danioninae fishes. The *chiron* genes played an essential role in zebrafish embryo development. By regulating NAD^+ levels, the rapid evolution of *chiron* genes may impact the evolution of the NAD^+ pathway and enhance their effects related to metabolic efficiency, energy stresses, and innate and adaptive immunity in subfamily Danioninae fishes. These changes likely contributed to the phenotypic and physiological adaptive evolution, survival, and prosperity of subfamily Danioninae fishes in Southeast Asian habitats. Furthermore, NAD^+ played critical roles in cell survival, age-associated pathophysiology (Garten et al., 2015), and lifespan (Yoshida et al., 2019). The systematic improvement of the entire NAD^+ synthesis pathway by new genes could also provide new insights into how the cell NAD^+ level may be increased, thereby promoting the prolongation of the life span in the bio-medical field.

MATERIALS AND METHODS

Identification of chimeric genes

We adapted a bioinformatics approach previously used in zebrafish to identify chimeric genes (Fu et al., 2010). The zebrafish genome sequence and annotated protein datasets were downloaded from Ensembl release 79. A series of chimeric retrogenes identified by a bioinformatics study were considered as candidates for the study of adaptation and the development of novel functions. Here, we focused on the chimeric gene family, *chiron* family. To investigate the distribution of *chiron* duplicates in other vertebrates, we used tblastn against the Ensembl DNA database in all 63 vertebrate species and three other closely related species (*Cyprinus carpio*, *Ctenopharyngodon idella*, and *Hypophthalmichthys molitrix*) with the default parameters. Moreover, we analyzed the genomic structure (exon/intron) of candidate sequences to determine whether these genes were chimeric structures.

Cloning of zebrafish *chiron*s and genomic organization

We performed the zebrafish experiments following the National Research Council Guide for the Care and Use of Laboratory Animals and it has been approved by the Ethics Committee of Huazhong Agricultural University, Wuhan city, Hubei province, China. We used a rapid amplification of cDNA ends (RACE) assay and reverse transcription (RT)-PCR to characterize the zebrafish *chiron*s structures. Total RNA was extracted from zebrafish testis and head using

TRIzol (Invitrogen, USA). RACE experiments were performed using a FirstChoice® RLM-RACE Kit (Ambion, USA). The full-length cDNAs of intact zebrafish *chiron*s were PCR-amplified. All PCR-amplified products were cloned into the pMD18-T vector (TaKaRa, Japan) and sequenced. Additionally, the genomic organization (exon/intron structures) of each gene was manually checked by cDNA-to-genome comparison. The RepeatMasker program was used to identify transposable elements within DNA sequences.

Phylogenetic analysis

Multiple CDS sequence alignments were generated using the Clustal program using the default parameters in MEGA 5.0. The nucleotides with ambiguity were manually checked and deleted to optimize the alignment. The *nampt* genes were selected as the out-group to analyze the relationship among the *chiron* genes. The ProtTest3.0 (Darriba et al., 2011) was used to obtain a substitution model that best fit the data. The maximum likelihood (ML) gene tree was performed with PhyML3.0 (Guindon et al., 2010) under the bootstrap values of 200 repetitions.

Evolution of the *nampt* and *chiron* genes

The divergence time between *nampt*-retroposed sequences and their parental copies was estimated by a phylogenetic tree and *Ks* values. The *Ka* and *Ks* substitution rates and *Ka*/*Ks* ratios were calculated using the codeml and yn00 programs implemented in the PAML v4.3 package (Yang, 2007). The *Ks* value was then used to estimate the divergence time between duplicates. The nonsynonymous to synonymous substitution rate ratio (ω=*Ka*/*Ks*) provided a method to detect selection pressure at the protein level, with ω<1, ω=1, and ω>1 indicating purifying selection, neutral evolution and positive selection, respectively (Yang and Nielsen, 2002). We performed the dynamic programming procedure to search for the optimal branch-specific model to analyze the lineage-specific selection pressures (Zhang et al., 2011a). Besides, we performed a sliding windows analysis (with 120-bp windows and 6-bp slides) to estimate the *Ka*/*Ks* ratio between each *nampt*-*chiron* gene pair. To inspect whether *chiron* genes could impact the genetic evolution of the NAD^+ biosynthetic pathway, we performed *branch* model (model=2; NSsites=0) and *branch-site* model (model=2; NSsites=2) in codeml to detect the selection pressures. The maximum likelihood method was used to examine whether the ω ratio was significantly different. The positive selection sites were detected by Bayes Empirical Bayes analysis with a posterior probability >95%. These positively selected sites were then mapped onto the three-dimensional structure of proteins (PDB) using Pymol (http://www.pymol.org/).

Expression patterns

We used high-throughput expression data to analyze the relative gene expression levels of each *chiron* gene and *nampt* in different tissues of zebrafish. A total of 53 transcriptome datasets, including data from 15 different tissues (Table S8 in Supporting Information), were downloaded from the Sequence Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra). Raw reads were trimmed for adapter contamination; the A-tails had a minimum quality score of 20 and a minimum length of 20 bp using Tim Galore. Trimmed reads were then aligned to the zebrafish reference genome (Ensembl 79: Zv9) using TopHat 2.0.14 (Kim et al., 2013). Then, the unique aligned reads were used to estimate differential gene expression using Cufflinks v.2.2.1 (Trapnell et al., 2010) software. The expression profiles of *nampt* and the *chiron* duplicates were estimated by the FPKM (fragments per kilobase of exon model per million mapped reads) value, which was calculated by Cuffdiff software using the transcriptome data (Mortazavi et al., 2008).

To visualize zebrafish *chiron*s' spatial-temporal expression patterns during early embryo development, we performed *in situ* hybridization experiments using digoxigenin-labeled antisense RNA as previously described (Chitramuthu and Bennett, 2013). The zebrafish embryos were fixed at appropriate developmental stages from 6 to 72 hpf in 4% PFA overnight. The antisense RNA probe of the *chiron* duplicates for *in situ* hybridization was synthesized from gene-specific 3′ UTR cDNA fragments (Ensembl: ENSDART-00000136950, nucleotides 2854–3251). The antisense probe of the parental gene *nampt* was derived from our previous work.

NAD^+/NADH level determination and Western blot

To evaluate whether Chiron proteins maintained the rate-limiting enzyme in the salvage pathway for NAD^+ formation, we measured NAD^+ and NADH levels in the human embryonic kidney (HEK) 293T cells. The entire coding regions of zebrafish *nampt* and *chiron2* were obtained by RT-PCR with the appropriate primers (Table S3 in Supporting Information). The PCR-amplified products were digested with restriction enzymes (*Bam*H I and *Hin*d III) and then cloned into the corresponding site of pCMV-Tag 2C (Agilent, USA). After sequencing, the constructs were transfected into HEK 293T cells using VigoFect (Vigorous, Beijing, China), and the cells were incubated under standard experimental conditions. The plasmid pCMV-Tag 2C was used as a control group, and each sample was repeated three times.

The cultured cell samples were collected after 30 h and washed with cold PBS. Each sample was lysed with NAD^+/NADH lysis buffer, followed by centrifugation at 14,000 r min^{-1} for 5 min. The supernatants were collected for

the NAD^+/NADH assay and Western blot analysis. We used a Fluorimetric NAD^+/NADH Ratio Assay Kit (AAT Bioquest, USA) to detect the NAD^+/NADH levels in the control and experimental groups. The fluorescence increase of each sample was monitored with a fluorescence plate reader at Ex/Em=540/590 nm. Recombinant proteins from the supernatants were boiled in 5× protein loading buffer (Beyotime, Shanghai, China) and visualized by SDS-PAGE and Western blot.

Knockdown and rescue experiments

Antisense translation-blocking morpholinos (MOs) targeting the start codon of all the zebrafish *chiron*s (Chiron_MO, TCTTGATTCTGGCAGAAATCCATCG, blue box in Figure S2 in Supporting Information) and the standard control (STD) morpholino (STD_MO, CCTCTTACCTCAGTTACAATTTATA) were designed from Gene Tools (Philomath, USA). The morpholinos were resuspended in 1 Danieau medium (58 mmol L^{-1} NaCl, 0.7 mmol L^{-1} KCl, 0.4 mmol L^{-1} $MgSO_4$, 0.6 mmol L^{-1} $Ca(NO_3)_2$, and 5.0 mmol L^{-1} HEPES, pH 7.6). To test the efficiency of Chiron-MO, 516-bp zebrafish *chiron* cDNA fragments, including the target region of the respective MO, were subcloned into pCAEGFP-N1 (Clontech, USA) to construct wild-type *chiron* recombinant plasmids (Chiron-GFP-WT). To validate the specificity of Chiron-MO, the zebrafish GFP-tagged mutated *chiron* recombinant plasmids (Chiron-GFP-MT) were also restructured by PCR using a forward primer with five mismatched nucleotides (5′-ATCGAAGCTTCCG ATGcAaTTgTGCgAGAATgA-3′, with mismatched nucleotides in the lowercase). Different amounts of MOs were injected into the yolks at the one-cell stage to obtain a specific defect. For the rescue experiments, quenching by the Chiron-MO was avoided; the five mismatched nucleotides were included in the forward primer and were inserted into the full-length zebrafish *chiron2* cDNA (5′-ATCGAAGCTTATGaAcTTCTGtCAagATCAAGA-3′, with mismatched nucleotides in the lowercase, Table S3 in Supporting Information). The sequence was subcloned into the Psp64 poly (A) vector (Promega, USA). The full length 5′ capped sense zebrafish *chiron2* mRNA was synthesized using the Amptcap SP6 High Yield message maker kit (Ambition, USA). After mixing with Chiron-MO, different amounts of synthetic mRNA were co-injected into the zebrafish yolk at the one-cell stage to obtain an optimal rescue effect.

CRISPR-cas9 gene-editing experiment

CRISPR guide RNAs were designed using ZiFiT (http://zifit.partners.org/ZiFiT/CSquare9Nuclease.aspx). CRISPR targets the Nicotinamide phosphoribosyltransferase activity region of all zebrafish *chiron* duplicates in exon 8 (target sequence: GGGGTTGGCCTTCTTCTCA, Figure S2 in Supporting Information). Double strand DNA for specific sgRNA synthesis was PCR-amplified using pMD 19-T vector as the temple. After gel extraction, sgRNA was synthesized using T7 RNA polymerase. A non-targeting sgRNA (CGAGTTAGAACGTAGGAGTC) was designed as a negative non-targeting control sgRNA. The zebrafish optimized Cas9 mRNA was synthesized *in vitro* from pT3 Cas9-UTRglobin vector. To test the efficiency, *chiron*s sgRNA (80 pg) mixed with Cas9 mRNA (500 pg) were co-injected into yolk at the one-cell stage. The genomic DNA of zebrafish embryos at 50 hpf was extracted, and the genomic region surrounding the CRISPR target site was PCR-amplified for sequencing. To confirm the essential function of zebrafish *chiron*s, different amounts of *chiron*s sgRNA or control sgRNA mixed with Cas9 mRNA were co-injected into one-cell-stage wild-type embryos respectively, and each sample was repeated three times.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This project is supported by the Pilot projects (XDB13020100), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000) and the National Natural Science Foundation of China (NSFC, 31601859). We would like to thank Prof. Manyuan Long (The University of Chicago), Prof. Wuhan Xiao (Institute of Hydrobiology, CAS), Prof. Z. Yin (Institute of Hydrobiology, CAS), Prof. Yonghua Sun (Institute of Hydrobiology, CAS), Prof. Shuhong Zhao (Huazhong Agricultural University), and Prof. Changchun Li (Huazhong Agricultural University) for providing experiment platform and technical assistance. We also thank Dr. Liandong Yang and Dr. Zaixuan Zhong (Institute of Hydrobiology, CAS) for providing the transcriptome sequencing data. We also thank Dr. Dylan Sosa and Emily Mortola (The University of Chicago) for refining the style and language.*

References

Assis, R., and Bachtrog, D. (2013). Neofunctionalization of young duplicate genes in *Drosophila*. Proc Natl Acad Sci USA 110, 17409–17414.

Berger, F., Lau, C., and Ziegler, M. (2007). Regulation of poly(ADP-ribose) polymerase 1 activity by the phosphorylation state of the nuclear NAD biosynthetic enzyme NMN adenylyl transferase 1. Proc Natl Acad Sci USA 104, 3765–3770.

Bergthorsson, U., Andersson, D.I., and Roth, J.R. (2007). Ohno's dilemma: Evolution of new genes under continuous selection. Proc Natl Acad Sci USA 104, 17004–17009.

Berto, S., and Nowick, K. (2018). Species-specific changes in a primate transcription factor network provide insights into the molecular evolution of the primate prefrontal cortex. Genome Biol Evol 10, 2023–2036.

Betrán, E., Thornton, K., and Long, M. (2002). Retroposed new genes out of the X in *Drosophila*. Genome Res 12, 1854–1859.

Biefer, H.R.C., Heinbokel, T., Uehara, H., Camacho, V., Minami, K., Nian, Y., Koduru, S., El Fatimy, R., Ghiran, I., Trachtenberg, A.J., et al. (2018). Mast cells regulate $CD4^+$ T-cell differentiation in the absence of antigen presentation. J Allergy Clin Immunol 142, 1894–1908.e7.

Blum, M., De Robertis, E.M., Wallingford, J.B., and Niehrs, C. (2015). Morpholinos: antisense and sensibility. Dev Cell 35, 145–149.

Boutant, M., and Cantó, C. (2014). SIRT1 metabolic actions: Integrating recent advances from mouse models. Mol Metab 3, 5–18.
Burgos, E.S., Ho, M.C., Almo, S.C., and Schramm, V.L. (2009). A phosphoenzyme mimic, overlapping catalytic sites and reaction coordinate motion for human NAMPT. Proc Natl Acad Sci USA 106, 13748–13753.
Cai, J., Zhao, R., Jiang, H., and Wang, W. (2008). *De Novo* origination of a new protein-coding gene in *Saccharomyces cerevisiae*. Genetics 179, 487–496.
Cantó, C., Menzies, K.J., and Auwerx, J. (2015). NAD^+ metabolism and the control of energy homeostasis: a balancing act between mitochondria and the nucleus. Cell Metab 22, 31–53.
Capra, J.A., Pollard, K.S., and Singh, M. (2010). Novel genes exhibit distinct patterns of function acquisition and network integration. Genome Biol 11, R127.
Chen, S., Ni, X., Krinsky, B.H., Zhang, Y.E., Vibranovski, M.D., White, K. P., and Long, M. (2012). Reshaping of global gene expression networks and sex-biased gene expression by integration of a young gene. EMBO J 31, 2798–2809.
Chen, S., Zhang, Y.E., and Long, M. (2010). New genes in *Drosophila* quickly become essential. Science 330, 1682–1685.
Chen, S., Krinsky, B.H., and Long, M. (2013). New genes as drivers of phenotypic evolution. Nat Rev Genet 14, 645–660.
Chitramuthu, B.P., and Bennett, H.P.J. (2013). High resolution whole mount *in situ* hybridization within zebrafish embryos to study gene expression and function. J Vis Exp, doi: 10.3791/50644.
Dahl, T.B., Holm, S., Aukrust, P., and Halvorsen, B. (2012). Visfatin/NAMPT: a multifaceted molecule with diverse roles in physiology and pathophysiology. Annu Rev Nutr 32, 229–243.
Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165.
Davies, T.J., Savolainen, V., Chase, M.W., Moat, J., and Barraclough, T.G. (2004). Environmental energy and evolutionary rates in flowering plants. Proc R Soc Lond B 271, 2195–2200.
Denu, J.M. (2007). Vitamins and aging: pathways to NAD^+ synthesis. Cell 129, 453–454.
Domazet-Lošo, T., and Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature 468, 815–818.
Dong, W.R., Xiang, L.X., and Shao, J.Z. (2009). Pre-B cell colony-enhancing factor in lower vertebrates: first evidence of this cytokine being involved in antioxidant activity by reconstruction of a novel NAD salvage pathway in *E. coli*. Int J Biochem Cell Biol 41, 1127–1137.
Engel, W., Hof, J.O., and Wolf, U. (1970). Gene duplication by polyploid evolution: the isoenzyme of the sorbitol dehydrogenase in herring- and salmon-like fishes (Isospondyli). Humangenetik 9, 157–163.
Fang, C., Guan, L., Zhong, Z., Gan, X., and He, S. (2015). Analysis of the nicotinamide phosphoribosyltransferase family provides insight into vertebrate adaptation to different oxygen levels during the water-to-land transition. FEBS J 282, 2858–2878.
Fang, C., Zou, C., Fu, Y., Li, J., Li, Y., Ma, Y., Zhao, S., and Li, C. (2018). DNA methylation changes and evolution of RNA-based duplication in *Sus scrofa*: based on a two-step strategy. Epigenomics 10, 199–218.
Fang, F. (2003). Phylogenetic analysis of the Asian cyprinid genus *Danio* (Teleostei, Cyprinidae). Copeia, 714–728.
Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531–1545.
Fu, B., Chen, M., Zou, M., Long, M., and He, S. (2010). The rapid generation of chimerical genes expanding protein diversity in zebrafish. BMC Genomics 11, 657.
Garavaglia, S., D'Angelo, I., Emanuelli, M., Carnevali, F., Pierella, F., Magni, G., and Rizzi, M. (2002). Structure of human NMN adenylyltransferase. J Biol Chem 277, 8524–8530.
Garten, A., Schuster, S., Penke, M., Gorski, T., de Giorgis, T., and Kiess, W. (2015). Physiological and pathophysiological roles of NAMPT and NAD metabolism. Nat Rev Endocrinol 11, 535–546.
Gilbert, W. (1978). Why genes in pieces? Nature 271, 501.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59, 307–321.
Hardison, R.C., Roskin, K.M., Yang, S., Diekhans, M., Kent, W.J., Weber, R., Elnitski, L., Li, J., O'Connor, M., Kolbe, D., et al. (2003). Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. Genome Res 13, 13–26.
Heinen, T.J.A.J., Staubach, F., Häming, D., and Tautz, D. (2009). Emergence of a new gene from an intergenic region. Curr Biol 19, 1527–1531.
Johnson, S., and Imai, S.I. (2018). NAD^+ biosynthesis, aging, and disease. F1000Res 7, 132.
Kaessmann, H. (2010). Origins, evolution, and phenotypic impact of new genes. Genome Res 20, 1313–1326.
Kalinka, A.T., Varga, K.M., Gerrard, D.T., Preibisch, S., Corcoran, D.L., Jarrells, J., Ohler, U., Bergman, C.M., and Tomancak, P. (2010). Gene expression divergence recapitulates the developmental hourglass model. Nature 468, 811–814.
Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845–858.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14, R36.
Klomp, J., Athy, D., Kwan, C.W., Bloch, N.I., Sandmann, T., Lemke, S., and Schmidt-Ott, U. (2015). A cysteine-clamp gene drives embryo polarity in the midge *Chironomus*. Science 348, 1040–1042.
Kumar, S., and Hedges, S.B. (2011). TimeTree2: species divergence times on the iPhone. Bioinformatics 27, 2023–2024.
Li, D., Dong, Y., Jiang, Y., Jiang, H., Cai, J., and Wang, W. (2010). A de novo originated gene depresses budding yeast mating pathway and is repressed by the protein encoded by its antisense strand. Cell Res 20, 408–420.
Liu, Q., Qi, Y., Liang, Q., Xu, X., Hu, F., Wang, J., Xiao, J., Wang, S., Li, W., Tao, M., et al. (2018). The chimeric genes in the hybrid lineage of Carassius auratus cuvieri (♀)×Carassius auratus red var. (♂). Sci China Life Sci 61, 1079–1089.
Long, M., Betrán, E., Thornton, K., and Wang, W. (2003). The origin of new genes: Glimpses from the young and old. Nat Rev Genet 4, 865–875.
Long, M., and Langley, C.H. (1993). Natural selection and the origin of jingwei, a chimeric processed functional gene in *Drosophila*. Science 260, 91–95.
Long, M., Rosenberg, C., and Gilbert, W. (1995). Intron phase correlations and the evolution of the intron/exon structure of genes. Proc Natl Acad Sci USA 92, 12495–12499.
Long, M., VanKuren, N.W., Chen, S., and Vibranovski, M.D. (2013). New gene evolution: little did we know. Annu Rev Genet 47, 307–333.
Loppin, B., Lepetit, D., Dorus, S., Couble, P., and Karr, T.L. (2005). Origin and neofunctionalization of a *Drosophila* paternal effect gene essential for zygote viability. Curr Biol 15, 87–93.
Lynch, M., O'Hely, M., Walsh, B., and Force, A. (2001). The probability of preservation of a newly arisen gene duplicate. Genetics 159, 1789–1804.
Marletta, A.S., Massarotti, A., Orsomando, G., Magni, G., Rizzi, M., and Garavaglia, S. (2015). Crystal structure of human nicotinic acid phosphoribosyltransferase. FEBS Open Bio 5, 419–428.
Martin, A.P., and Palumbi, S.R. (1993). Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci USA 90, 4087–4091.
Matsuno, M., Compagnon, V., Schoch, G.A., Schmitt, M., Debayle, D., Bassard, J.E., Pollet, B., Hehn, A., Heintz, D., Ullmann, P., et al. (2009). Evolution of a novel phenolic pathway for pollen development. Science

325, 1688–1692.
Mayden, R.L., Tang, K.L., Conway, K.W., Freyhof, J., Chamberlain, S., Haskins, M., Schneider, L., Sudkamp, M., Wood, R.M., Agnew, M., et al. (2007). Phylogenetic relationships of *Danio* within the order Cypriniformes: a framework for comparative and evolutionary studies of a model species. J Exp Zool 308B, 642–654.
McLure, K.G., Takagi, M., and Kastan, M.B. (2004). NAD^+ modulates p53 DNA binding specificity and function. Mol Cell Biol 24, 9958–9967.
McLysaght, A., and Hurst, L.D. (2016). Open questions in the study of *de novo* genes: what, how and why. Nat Rev Genet 17, 567–578.
Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5, 621–628.
Mu, W.L., Zhang, Q.J., Tang, X.Q., Fu, W.Y., Zheng, W., Lu, Y.B., Li, H. L., Wei, Y.S., Li, L., She, Z.G., et al. (2014). Overexpression of a dominant-negative mutant of SIRT1 in mouse heart causes cardiomyocyte apoptosis and early-onset heart failure. Sci China Life Sci 57, 915–924.
Nasevicius, A., and Ekker, S.C. (2000). Effective targeted gene 'knockdown' in zebrafish. Nat Genet 26, 216–220.
Norton, W., and Bally-Cuif, L. (2010). Adult zebrafish as a model organism for behavioural genetics. BMC Neurosci 11, 90.
Ognjanovic, S., Bao, S., Yamamoto, S.Y., Garibay-Tupas, J., Samal, B., and Bryant-Greenwood, G.D. (2001). Genomic organization of the gene coding for human pre-B-cell colony enhancing factor and expression in human fetal membranes. J Mol Endocrinol 26, 107–117.
Ohno, S., Wolf, U., and Atkin, N.B. (1968). Evolution from fish to mammals by gene duplication. Hereditas 59, 169–187.
Pan, D., and Zhang, L. (2009). Burst of young retrogenes and independent retrogene formation in mammals. PLoS ONE 4, e5040.
Parichy, D.M. (2015). Advancing biology through a deeper understanding of zebrafish ecology and evolution. eLife 4, e05635.
Patthy, L. (2008). Protein Evolution, 2nd ed. (Oxford: Wiley-Blackwell).
Pfister, N.T., Yoh, K.E., and Prives, C. (2014). p53, DNA damage, and NAD^+ homeostasis. Cell Cycle 13, 1661–1662.
Piasecka, B., Lichocki, P., Moretti, S., Bergmann, S., and Robinson-Rechavi, M. (2013). The hourglass and the early conservation models—co-existing patterns of developmental constraints in vertebrates. PLoS Genet 9, e1003476.
Ragsdale, E.J., Müller, M.R., Rödelsperger, C., and Sommer, R.J. (2013). A developmental switch coupled to the evolution of plasticity acts through a sulfatase. Cell 155, 922–933.
Robu, M.E., Larson, J.D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S.A., and Ekker, S.C. (2007). p53 activation by knockdown technologies. PLoS Genet 3, e78.
Rogers, R.L., and Hartl, D.L. (2012). Chimeric genes as a source of rapid evolution in *Drosophila melanogaster*. Mol Biol Evol 29, 517–529.
Ross, B.D., Rosin, L., Thomae, A.W., Hiatt, M.A., Vermaak, D., de la Cruz, A.F.A., Imhof, A., Mellone, B.G., and Malik, H.S. (2013). Stepwise evolution of essential centromere function in a *Drosophila* neogene. Science 340, 1211–1214.
Roux, J., and Robinson-Rechavi, M. (2008). Developmental constraints on vertebrate genome evolution. PLoS Genet 4, e1000311.
Ruiz-Orera, J., Verdaguer-Grau, P., Villanueva-Cañas, J.L., Messeguer, X., and Albà, M.M. (2018). Translation of neutrally evolving peptides provides a basis for *de novo* gene evolution. Nat Ecol Evol 2, 890–896.
Shao, Y., Chen, C., Shen, H., He, B.Z., Yu, D., Jiang, S., Zhao, S., Gao, Z., Zhu, Z., Chen, X., et al. (2019). GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. Genome Res 29, 682–696.
Tang, K.L., Agnew, M.K., Hirt, M.V., Sado, T., Schneider, L.M., Freyhof, J., Sulaiman, Z., Swartz, E., Vidthayanon, C., Miya, M., et al. (2010). Systematics of the subfamily Danioninae (Teleostei: Cypriniformes: Cyprinidae). Mol Phylogenet Evol 57, 189–214.
Thomas, J.A., Welch, J.J., Lanfear, R., and Bromham, L. (2010). A generation time effect on the rate of molecular evolution in invertebrates. Mol Biol Evol 27, 1173–1180.
Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28, 511–515.
VanKuren, N.W., and Long, M. (2018). Gene duplicates resolving sexual conflict rapidly evolved essential gametogenesis functions. Nat Ecol Evol 2, 705–712.
Verdin, E. (2015). NAD^+ in aging, metabolism, and neurodegeneration. Science 350, 1208–1213.
Vinckenbosch, N., Dupanloup, I., and Kaessmann, H. (2006). Evolutionary fate of retroposed gene copies in the human genome. Proc Natl Acad Sci USA 103, 3220–3225.
Wang, T., Zhang, X., Bheda, P., Revollo, J.R., Imai, S., and Wolberger, C. (2006a). Structure of Nampt/PBEF/visfatin, a mammalian NAD^+ biosynthetic enzyme. Nat Struct Mol Biol 13, 661–662.
Wang, W., Elkins, K., Oh, A., Ho, Y.C., Wu, J., Li, H., Xiao, Y., Kwong, M., Coons, M., Brillantes, B., et al. (2014). Structural basis for resistance to diverse classes of NAMPT inhibitors. PLoS ONE 9, e109366.
Wang, W., Zheng, H., Fan, C., Li, J., Shi, J., Cai, Z., Zhang, G., Liu, D., Zhang, J., Vang, S., et al. (2006b). High rate of chimeric gene origination by retroposition in plant genomes. Plant Cell 18, 1791–1802.
Wang, Y., Lu, Y., Zhang, Y., Ning, Z., Li, Y., Zhao, Q., Lu, H., Huang, R., Xia, X., Feng, Q., et al. (2015). The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation. Nat Genet 47, 625–631.
Weiss, K.M. (2005). The phenogenetic logic of life. Nat Rev Genet 6, 36–45.
Weng, J.K., Li, Y., Mo, H., and Chapple, C. (2012). Assembly of an evolutionarily new pathway forα-pyrone biosynthesis in *Arabidopsis*. Science 337, 960–964.
Wu, X., and Sharp, P.A. (2013). Divergent transcription: a driving force for new gene origination? Cell 155, 990–996.
Xia, S., Wang, Z., Zhang, H., Hu, K., Zhang, Z., Qin, M., Dun, X., Yi, B., Wen, J., Ma, C., et al. (2016). Altered transcription and neofunctionalization of duplicated genes rescue the harmful effects of a chimeric gene in *Brassica napus*. Plant Cell 28, 2060–2078.
Xie, C., Zhang, Y.E., Chen, J.Y., Liu, C.J., Zhou, W.Z., Li, Y., Zhang, M., Zhang, R., Wei, L., and Li, C.Y. (2012). Hominoid-specific *de novo* protein-coding genes originating from long non-coding RNAs. PLoS Genet 8, e1002942.
Xu, H.B., Li, Y.X., Li, Y., Otecko, N.O., Zhang, Y.P., Mao, B., and Wu, D. D. (2018). Origin of new genes after zygotic genome activation in vertebrate. J Mol Cell Biol 10, 139–146.
Xu, P., Zhang, X., Wang, X., Li, J., Liu, G., Kuang, Y., Xu, J., Zheng, X., Ren, L., Wang, G., et al. (2014). Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. Nat Genet 46, 1212–1219.
Yang, L., Ma, X., He, Y., Yuan, C., Chen, Q., Li, G., and Chen, X. (2017). Sirtuin 5: a review of structure, known inhibitors and clues for developing new inhibitors. Sci China Life Sci 60, 249–256.
Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24, 1586–1591.
Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol 19, 908–917.
Yoshida, M., Satoh, A., Lin, J.B., Mills, K.F., Sasaki, Y., Rensing, N., Wong, M., Apte, R.S., and Imai, S.I. (2019). Extracellular vesicle-contained eNAMPT delays aging and extends lifespan in mice. Cell Metab 30, 329–342.e5.
Yoshino, J., Mills, K.F., Yoon, M.J., and Imai, S. (2011). Nicotinamide mononucleotide, a key NAD^+ intermediate, treats the pathophysiology of diet- and age-induced diabetes in mice. Cell Metab 14, 528–536.
Zhang, C., Wang, J., Xie, W., Zhou, G., Long, M., and Zhang, Q. (2011a). Dynamic programming procedure for searching optimal models to estimate substitution rates based on the maximum-likelihood method.

Proc Natl Acad Sci USA 108, 7860–7865.
Zhang, J., Dean, A.M., Brunet, F., and Long, M. (2004). Evolving protein functional diversity in new genes of *Drosophila*. Proc Natl Acad Sci USA 101, 16246–16250.
Zhang, J., Long, M., and Li, L. (2005). Translational effects of differential codon usage among intragenic domains of new genes in *Drosophila*. Biochim Biophys Acta Gene Struct Expr 1728, 135–142.
Zhang, J., Yang, H., Long, M., Li, L., and Dean, A.M. (2010). Evolution of enzymatic activities of testis-specific short-chain dehydrogenase/reductase in *Drosophila*. J Mol Evol 71, 241–249.
Zhang, L., Ren, Y., Yang, T., Li, G., Chen, J., Gschwend, A.R., Yu, Y., Hou, G., Zi, J., Zhou, R., et al. (2019a). Rapid evolution of protein diversity by *de novo* origination in *Oryza*. Nat Ecol Evol 3, 679–690.
Zhang, W., Gao, Y., Long, M., and Shen, B. (2019b). Origination and evolution of orphan genes and *de novo* genes in the genome of *Caenorhabditis elegans*. Sci China Life Sci 62, 579–593.
Zhang, W., Landback, P., Gschwend, A.R., Shen, B., and Long, M. (2015). New genes drive the evolution of gene interaction networks in the human and mouse genomes. Genome Biol 16, 202.
Zhang, Y.E., Landback, P., Vibranovski, M.D., and Long, M. (2011b). Accelerated recruitment of new brain development genes into the human genome. PLoS Biol 9, e1001179.
Zhang, Z., Fan, Y., Xiong, J., Guo, X., Hu, K., Wang, Z., Gao, J., Wen, J., Yi, B., Shen, J., et al. (2020). Two young genes reshape a novel interaction network in *Brassica napus*. New Phytol 225, 530–545.
Zhao, L., Saelao, P., Jones, C.D., and Begun, D.J. (2014). Origin and spread of *de novo* genes in *Drosophila melanogaster* populations. Science 343, 769–772.
Zhuang, X., Yang, C., Murphy, K.R., and Cheng, C.H.C. (2019). Molecular mechanism and history of non-sense to sense evolution of antifreeze glycoprotein gene in northern gadids. Proc Natl Acad Sci USA 116, 4400–4405.
Zu, J., Gu, Y., Li, Y., Li, C., Zhang, W., Zhang, Y.E., Lee, U.J., Zhang, L., and Long, M. (2019). Topological evolution of coexpression networks by new gene integration maintains the hierarchical and modular structures in human ancestors. Sci China Life Sci 62, 594–608.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-020-1851-0. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.