

Article

A Comparative Genomic and Transcriptomic Survey Provides Novel Insights into *N*-Acetylserotonin Methyltransferase (ASMT) in Fish

Kai Zhang ^{1,2}, Zhiqiang Ruan ², Jia Li ², Chao Bian ^{2,3}, Xinxin You ², Steven L. Coon ^{4,5,*} and Qiong Shi ^{1,2,3,*}

¹ BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China; zhangkai@genomics.cn

² Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, BGI, Shenzhen 518083, China; ruanzhiqiang@genomics.cn (Z.R.); lijia1@genomics.cn (J.L.); bianchao@genomics.cn (C.B.); youxinxin@genomics.cn (X.Y.)

³ BGI-Zhenjiang Institute of Hydrobiology, BGI Marine, Zhenjiang 212000, China

⁴ Molecular Genomics Core, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA

⁵ Molecular Genomics Laboratory, National Institutes of Health, Bethesda, MD 20892, USA

* Correspondence: coons@mail.nih.gov (S.L.C.); shiqiong@genomics.cn (Q.S.); Tel: +1-301-451-6622 (S.L.C.); +86-185-6627-9826 (Q.S.); Fax: +1-301-480-0378 (S.L.C.); +86-755-3630-7807 (Q.S.)

Received: 14 September 2017; Accepted: 30 September 2017; Published: 2 October 2017

Abstract: Melatonin is a multifunctional bioactive molecule that plays comprehensive physiological roles in all living organisms. *N*-acetylserotonin methyltransferase (ASMT, also known as hydroxyindole *O*-methyltransferase or HIOMT) is the final enzyme for biosynthesis of melatonin. Here, we performed a comparative genomic and transcriptomic survey to explore the ASMT family in fish. Two ASMT isotypes (ASMT1 and ASMT2) and a new ASMT-like (ASMTL) are all extracted from teleost genomes on the basis of phylogenetic and synteny analyses. We confirmed that C-terminal of the ASMTL proteins (ASMTL-ASMT) is homology to the full length of ASMT1 and ASMT2. Our results also demonstrate that the two ASMT isotypes and their distribution in teleosts seem to be the result of combinations of whole-genome duplication (WGD) and gene loss. Differences were also observed in tissue distribution and relative transcript abundances of ASMT1, ASMT2 and ASMTL through transcriptomic analysis. Protein sequence alignment and 3D structure prediction of ASMTs and ASMTL suggest differential roles for these ASMT genes. In summary, our current work provides novel insights into the ASMT genes in fish by combination of genomic and transcriptomic data.

Keywords: *N*-acetylserotonin methyltransferase; melatonin; phylogenetic analysis; comparative genomic and transcriptomic survey; fish

1. Introduction

Melatonin is a multifunctional bioactive molecule that regulates circadian rhythms and seasonal reproductive processes; it is mainly synthesized rhythmically and secreted by the pineal gland and retinae [1]. This rhythmic secretory pattern, highly conserved in all vertebrates, is determined by the precise circadian regulation of melatonin biosynthesis [2]. Melatonin biosynthesis from the amino acid tryptophan (Trp) involves four enzyme-catalyzed reactions [3]. Firstly, Trp is transformed into 5-hydroxytryptophan (5-HTp) by tryptophan hydroxylase (TPH, EC 1.14.16.4). Subsequently, 5-HTp is converted to 5-hydroxytryptamine (serotonin) by the

cytoplasmic enzyme dopa decarboxylase (DDC, also known as aromatic L-amino acid decarboxylase AADC, EC 4.1.1.28). The third step is the transformation of serotonin by aralkylamine *N*-acetyltransferase (AANAT, EC 2.3.1.87) to *N*-acetylserotonin (NAS). The last step is completed by transforming NAS to melatonin (*N*-acetyl-5-methoxytryptamine) with acetylserotonin *O*-methyltransferase (ASMT, also known as HIOMT, EC 2.1.1.4). As the final enzyme in the pathway of melatonin biosynthesis, ASMT is probably responsible for the seasonal variations in the melatonin secretion rhythm [4,5]. In addition, ASMT is considered a rate-limiting enzyme in melatonin synthesis since it provides an upper limit to the overall production rate once AANAT has been activated [6,7].

ASMT, identified in both animals and plants, belongs to the methyltransferase superfamily [8–10]. In plants, rice ASMT comprises three isoforms (*ASMT1*, *ASMT2*, and *ASMT3*) and all of the isoforms could encode ASMT activity, and overexpression of them could lead to the overproduction of melatonin in rice [8,11,12]. In mammals, the *ASMT* gene is located on the X chromosome [9,13]. The human *ASMT* gene possesses three possible isoforms, which have resulted from the alternative splicing of exons 6 and 7. Isoform 1 catalyzes the traditional transfer of a methyl group onto *N*-acetylserotonin, producing melatonin; however, isoforms 2 and 3 lack this enzyme activity [14]. In contrast to tetrapods, two *ASMT* genes have been reported in fish genomes, and may have been generated by the teleost-specific whole genome duplication (WGD) [15].

It was reported that the ASMT enzyme mainly exists in the retinae and pineal gland of the European sea bass [14]. *ASMT2* was detected in several peripheral tissues, including liver and gut in teleost [10]. The existence of melatonin synthesis in gut and liver of goldfish has been demonstrated [15]. Recent studies suggested high expression of gut *ASMT* in zebrafish (*Danio rerio*) and rainbow trout (*Oncorhynchus mykiss*) [16,17]. Findings on a tropical carp (*Catla catla*) recently indicated that abundant *ASMT* mRNAs were detected in the gut and the transcription level of *ASMT* in brain displayed a significant negative correlation with water temperature [18]. Moreover, the involvement of brain melatonin in the modulation of seasonal reproductive parameters through the putative hypothalamo-pituitary-gonadal axis is suggested by high expression of *ASMT* in the brain during the preparatory phase [19–21].

However, a previous study [22] reported a new gene, acetylserotonin methyltransferase-like (*ASMTL*) in animals, which is significantly homologous to the putative *ASMT* genes. In fact, *ASMTL* is the product of a gene fusion event, during which two different full-length genes (*maf* and *ASMT*) have been joined to form one single gene [22]. In contrast to those known *ASMT* genes, which are associated with the pathway of melatonin biosynthesis, the function of *ASMTL* is unclear as of yet. Interestingly, there are more clues to support the existence of *ASMTL* in sequenced fish genomes. For example, in the GenBank database, accession numbers NC_007120.7 (chromosome 9 of zebrafish) and NC_019879.1 (chromosome 21 of medaka) reveal the sequences for *ASMTL*.

Although two putative *ASMT* isoforms and *ASMTL* have been reported in fish, details of their evolution, 3D structures, tissue distribution and potential functions of these genes are still unknown. In addition, previous studies involved only a few bony fish, and no whole genome data were employed. With advances in next-generation sequencing capabilities, genomic and transcriptomic data of many fish species are now available, and with rapid developments in bioinformatics, this sequence information is more valuable. We can access fish *ASMT* genes (and their encoded protein sequences) from many published and unpublished whole genome sequences, including amphibious (mudskippers), cave-restricted or eyeless (*Sinocyclocheilus anshuiensis*), and tetraploid (*Sinocyclocheilus spp.*) fishes. This allows us to determine species-specific presence or absence of *ASMT* isoforms, as well as sequence differences across the species. In this report, we analyzed the protein sequence differences in fish *ASMT* isoforms and *ASMTL*, and constructed a phylogenetic tree with a synteny analysis. Furthermore, we also analyzed the exon-intron structures of these three genes, and obtained full-length *ASMT1*, *ASMT2*, and *ASMTL* in two representative mudskipper species by molecular cloning. Finally, transcriptome analysis was applied to determine the tissue distribution of *ASMT* genes.

2. Results

2.1. Copy Number Variation and Phylogenetic Relationships

All of the *ASMT* sequences were collected from 40 vertebrate species (Tables 1 and S1). Previous studies reported that two *ASMT* genes are present in fish [18]. However, our present results confirmed that teleost genomes contain two or more *ASMT* genes. For example, many diploid teleosts possess two putative *ASMT* genes (*ASMT1* and *ASMT2*); however, the tetraploid *Sinocyclocheilus* fishes (Sg, Sr and Sa) possess two copies of the *ASMT2* gene (temporally named as *ASMT2a* and *ASMT2b*), although they displayed a high degree of conservation (over 93% similarity), reaching a total of three putative *ASMT* isotypes (Table 1). In contrast, only one *ASMT* gene was identified in the genomes of elephant shark (a cartilaginous fish) and Atlantic salmon. Moreover, our data demonstrated that teleost genomes possess only one *ASMTL* gene.

Table 1. Copy number of *ASMT1*, *ASMT2*, and *ASMTL* genes in sequenced fish genomes.

Common Name	Species Name	<i>ASMT1</i>	<i>ASMT2</i>	<i>ASMTL</i>
Golden-line fishes (Sa) ¹	<i>Sinocyclocheilus anshuiensis</i>	1	2	1
(Sg) ¹	<i>Sinocyclocheilus grahami</i>	1	2	1
(Sr) ¹	<i>Sinocyclocheilus rhinoceros</i>	1	2	1
Large yellow croaker	<i>Larimichthys crocea</i>	1	1	1
Northern snakehead	<i>Channa argus</i>	1	1	1
Giant-fin mudskipper	<i>Periophthalmus magnuspinnatus</i>	1	1	1
Blue-spotted mudskipper	<i>Boleophthalmus pectinirostris</i>	1	1	1
Lined seahorse	<i>Hippocampus erectus</i>	1	1	1
Kanglang fish	<i>Anabarilius grahami</i>	1	1	1
Black porgy	<i>Acanthopagrus schlegelii</i>	1	1	1
Southern platyfish	<i>Xiphophorus maculatus</i>	1	1	1
Bicolor damselfish	<i>Stegastes partitus</i>	1	1	1
Red-bellied piranha	<i>Pygocentrus nattereri</i>	1	1	1
Channel catfish	<i>Ictalurus punctatus</i>	1	1	1
Green arowana	<i>Scleropages formosus</i>	1	1	1
Golden arowana	<i>Scleropages formosus</i>	1	1	-
Red arowana	<i>Scleropages formosus</i>	1	1	-
Longjaw grenadier anchovy	<i>Coilia macrognathos</i>	1	1	-
Northern pike	<i>Esox lucius</i>	1	1	-
Atlantic salmon	<i>Salmo salar</i>	1	-	-
Elephant shark	<i>Callorhynchus milii</i>	1	-	-

-: unidentified. ¹ Tetraploid fishes.

In addition, we performed a phylogenetic analysis using all the protein sequences of vertebrate *ASMTs*. On the basis of the constructed phylogenetic tree (Figure 1), we observed that fish *ASMT* genes form three distinct clades with high internal similarity (70.2–95.6%). The protein sequences of *ASMTs* were also aligned to our unpublished transcriptome database of Fish-T1K (Transcriptomes of 1000 fishes [23]; <http://www.fisht1k.org/>). Interestingly, we found that homologous sequences of these genes were present in almost all sequenced fish species (data not shown), suggesting a wide existence of the three *ASMT* genes in fishes.

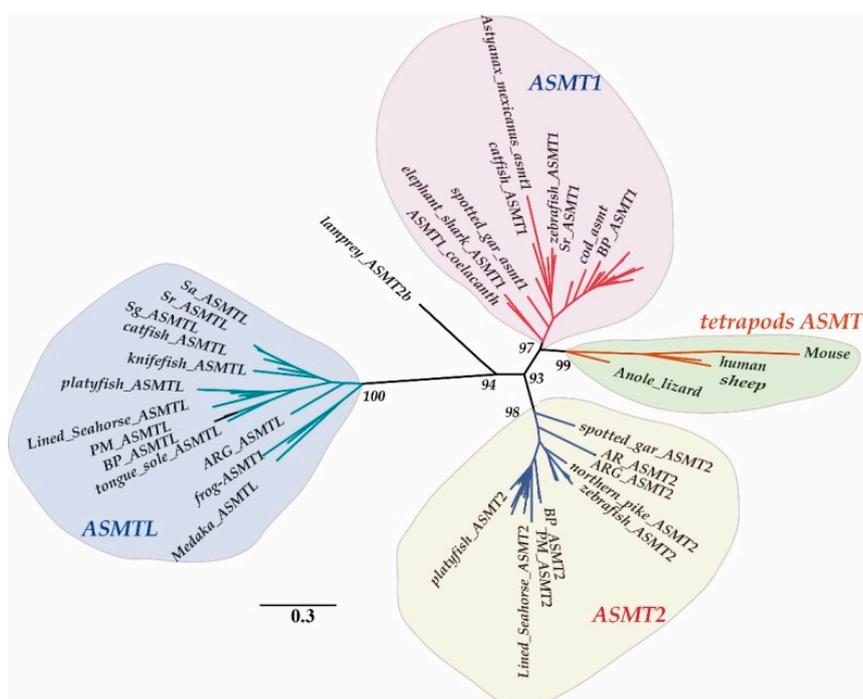


Figure 1. Clustering of ASMT genes. The phylogenetic analysis was performed by RAxML 8.0.17. Numbers on branches are bootstrap values. Scale bar indicates the rate of amino acid substitution per residue. Although the Mega tree is not shown here, it possesses the same tree topology. Different ASMT subfamilies are displayed with various colors: Pink, teleost ASMT1; Green, tetrapod ASMT; Yellow, teleost ASMT2; and Blue, teleost ASMTL. Abbreviations of fish species: PM, *Periophthalmus magnuspinnatus*; BP, *Boleophthalmus pectinirostris*; AR, Golden arowana; ARG, Green arowana; Sa, *Sinocyclocheilus anshuiensis*; Sg, *S. graham*; Sr, *S. rhinoceros*.

2.2. Synteny Data

Our synteny analysis showed that ASMT1 genes, and ASMT2 genes, across species share a conserved suite of genes bounding them on both sides, although some species may display gene loss (Figure 2). In general, 11 genes (*GYG2*, *CD99*, *ZBED1*, *DHRXS*, *akap17a*, *P2RY8*, *ASMTL*, *SLC25A6*, *IL3RA*, *CSF2RA*, and *CRLF2*) usually flank the mammalian ASMT on both sides (the shadowed area in Figure 2A). Four (*cps1*, *map2*, *myl1*, and *slc1za8*) and seven genes (*akap17a*, *SH3RF3*, *sept10*, *edar*, *chmp2ba*, *htr1fb*, and *epha3*) reside on the upstream and the downstream regions respectively of fish ASMT1 (Figure 2A). Four genes (*Fgfr3*, *letm1*, *whsc1*, and *jade1*) are located the upstream and four genes (*pgrmc2*, *larp1b*, *ahsa1b*, and *cdca4*) are positioned the downstream of fish ASMT2 (Figure 2B). Moreover, nine genes reside around fish ASMTL, in which five genes (*itgbl1*, *fgf14*, *cdc16*, *upf3a*, and *p2ry8*) and four genes (*shox*, *medl4*, *rpl8*, and *itga4*) are located on the upstream and the downstream of fish ASMTL, respectively (Figure 2C).

Interestingly, we observed that ASMT1 and ASMTL are sometimes located on the same chromosome or scaffold in vertebrates including fishes. For example, ASMT and ASMTL in humans reside on chromosome X, in zebrafish on chromosome 9, and in medaka on chromosome 21 (Figure 2A,C).

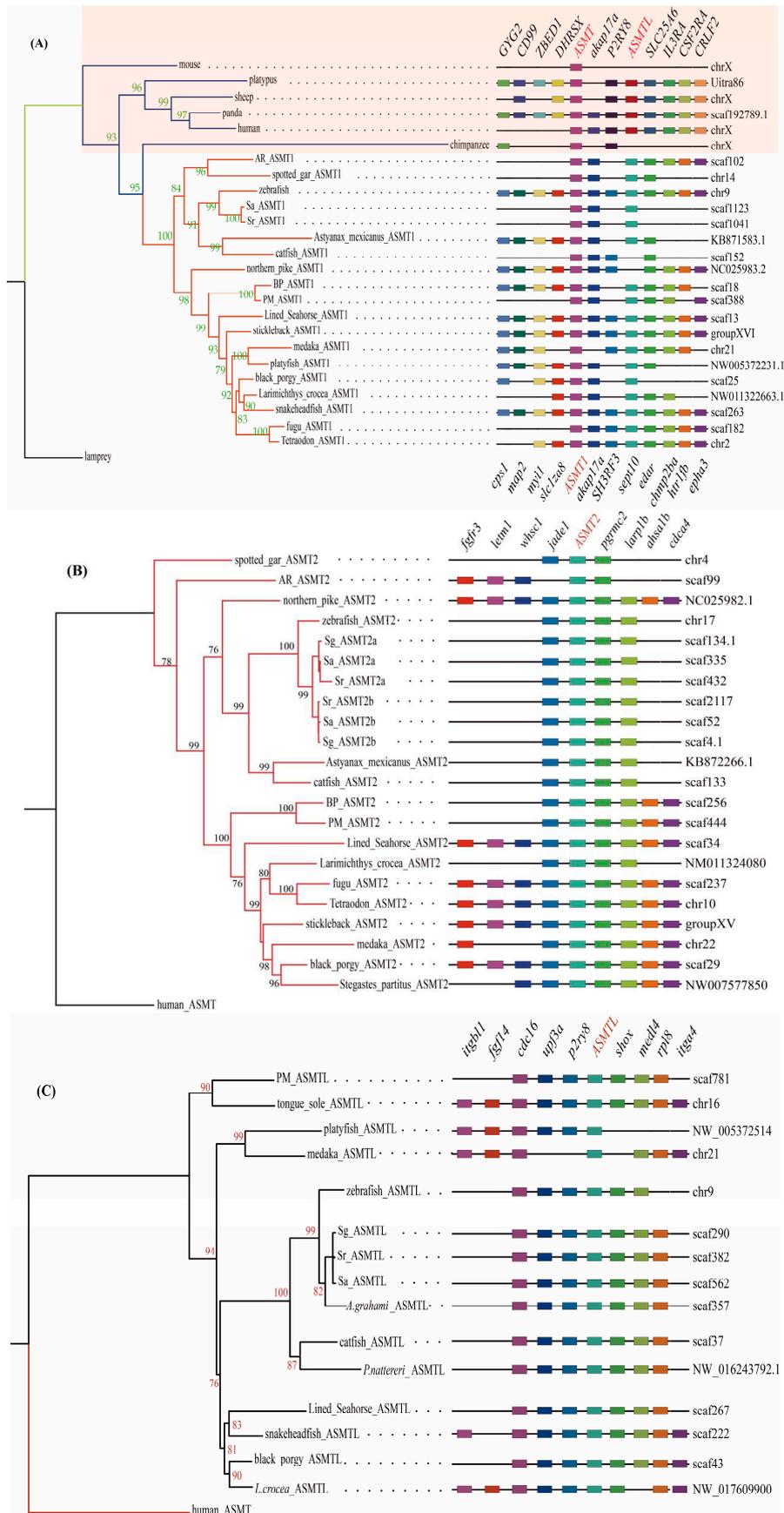


Figure 2. Phylogenetic trees and synteny of *ASMT1* (A), *ASMT2* (B), and *ASMTL* (C). These ML trees (on the left side of each figure) were constructed by FastTree v2.1.7. Bootstrap values are shown on branches. The right figures are synteny of the three ASMT genes.

2.3. Structural Analysis of ASMT1, ASMT2 and ASMTL

A previous study [22] reported that ASMTL has two different domains, with close homology to two different genes (multicopy associated filamentation (*maf*) and ASMT; Figure 3A), with the C-terminal (ASMTL-ASMT) region homologous to the full length of the putative ASMTs (Figure 3B). We obtained the gene structures of ASMT members by comparing the zebrafish coding sequences and the collected genomic DNA sequences. Obviously, ASMT1 and ASMT2 are, respectively, conserved, sharing an independently similar pattern of exon-intron structure (Figure 4A,B), while they are remarkably different from the ASMTL style (Figure 4C). Meanwhile, all of the exon-intron boundaries are conserved. For example, BP-ASMTL is composed of 13 exons, and exons 8–13 (ASMTL-ASMT) are homologous to the ASMT1 and ASMT2 sequences (Figures S1 and 3B). The Exon 7 (Figure S1) is the joining fragment between *maf* and ASMTL-ASMT (Figure 3A).

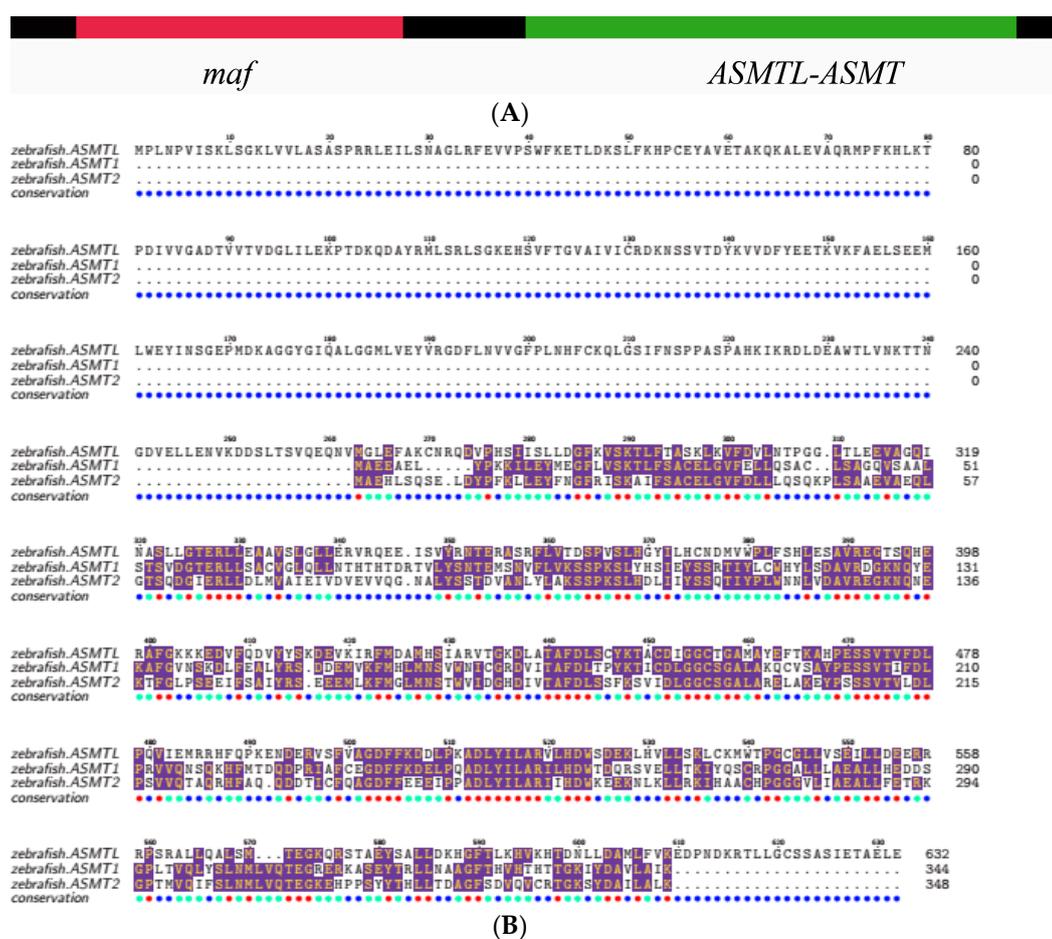


Figure 3. Alignment of ASMT protein sequences of zebrafish. (A) ASMTL has two different domains, with homology to two different genes (*maf* and ASMTL-ASMT; modified from [22]). (B) Similarity of the C-terminal of ASMTL (ASMTL-ASMT) to ASMT1 and ASMT2. See more details about the exon boundary in Figure S1.

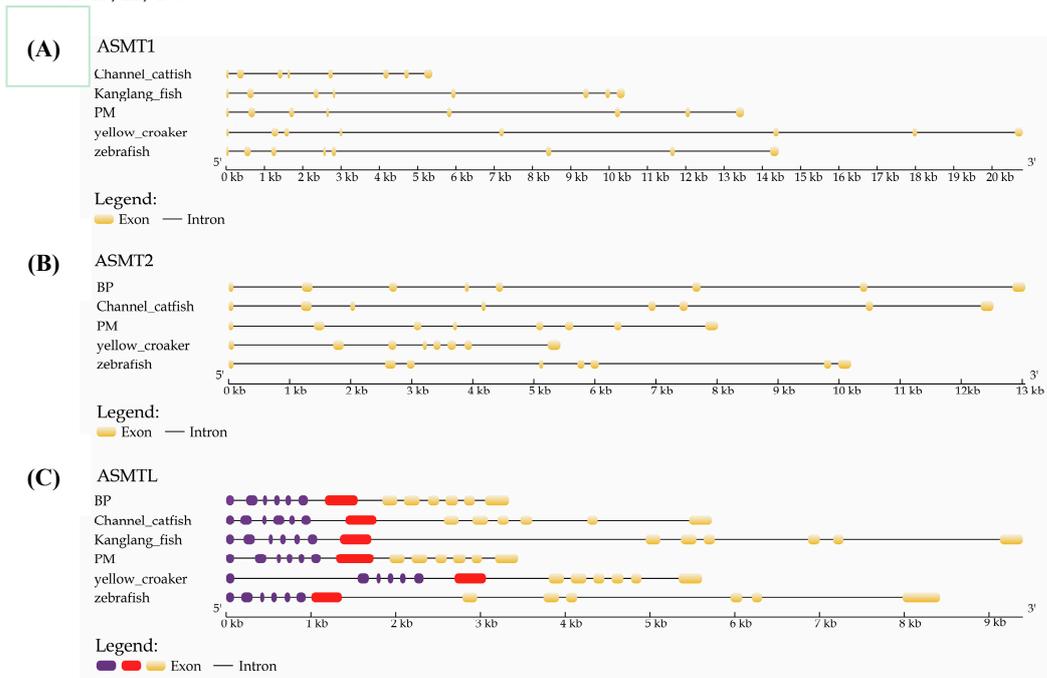


Figure 4. Exon-intron structures of *ASMT1* (A), *ASMT2* (B) and *ASMTL* (C). Boxes in purple, red and yellow stand for the exons, while gray lines represent the introns. Note that the yellow boxes in (C) *ASMTL* are homologous to *ASMT1* and *ASMT2*.

2.4. The Structure of ASMT Proteins

Representative ASMT protein sequences were aligned, and many residues were found to be well conserved (Figure 5). Previous studies reported that the polypeptide chain of ASMT comprises a C-terminal domain (a typical feature different from other SAM-dependent O-methyltransferases) and an N-terminal domain (intertwining several helices with another monomer to form the physiologically active dimer) [14]. Many residues around the SAM (S-adenosyl methionine) binding site are strictly conserved, such as the residues F156, G187, G235, D236, F237, F238, W257, D210, and R252. They may be helpful to maintain the H-bonds to the SAM moiety. At the NAS binding site, some residues are also well conserved between the sequences of *ASMT1* and *ASMT2*, including Y108, F156, H255, D256, N302, M303, Q306, E311, and Y338. However, for these residues, some variants were present in the *ASMTL*-*ASMT* sequences. For example, Y108H, N302S and E311Q are among the obvious differences. Interestingly, previous findings revealed that some other variants can lead to reduced ASMT activity, such as E61Q and P243L [14], which were also identified in our current work (Figure 5).

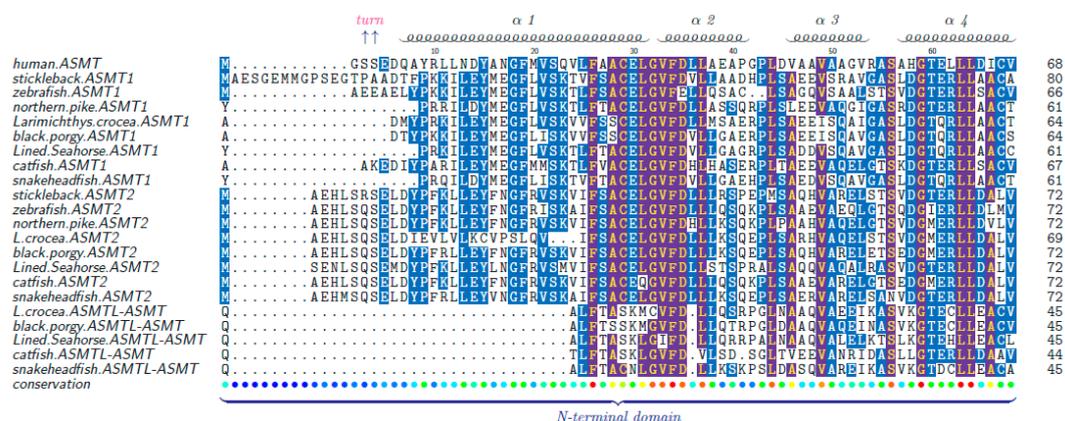


Figure 5. Cont.

2.5. Transcription and Cloning of ASMTs

Transcriptome data were analyzed to localize ASMT transcripts in different tissues of three representative *Sinocyclocheilus* fishes (Sg, *Sinocyclocheilus graham*; Sr, *S. rhinoceros*; and, Sa, *S. anshuiensis*) and two representative mudskippers (BP, *Boleophthalmus pectinirostris*; PM, *Periophthalmus magnuspinnatus*). RPKM (Reads Per Kilobase Transcriptome per Million mapped reads) values (Table 2) were used to quantify gene transcription levels.

Table 2. Reads Per Kilobase Transcriptome per Million (RPKM) of ASMT transcripts in three *Sinocyclocheilus* fishes.

Tissue	Gene	Sg	Sr	Sa
Eye	ASMT1	30.57	13.50	1.62
	ASMT2a	18.14	3.15	0.90
	ASMT2b	0.00	0.00	8.57
	ASMTL	0.00	3.35	122.11
Skin	ASMT1	0.09	0.07	0.19
	ASMT2a	7.36	2.71	2.07
	ASMT2b	0.00	0.35	5.02
	ASMTL	0.00	4.55	37.71
Liver	ASMT1	0.00	0.00	0.00
	ASMT2a	1.30	0.96	1.59
	ASMT2b	0.00	0.11	5.46
	ASMTL	0.00	71.22	118.35
Gonad	ASMT1	0.00	0.00	0.42
	ASMT2a	0.00	0.48	0.21
	ASMT2b	0.00	0.05	34.99
	ASMTL	0.00	17.71	14.78

In the *Sinocyclocheilus* fishes, transcripts of four ASMT genes (ASMT1, ASMT2a, ASMT2b, and ASMTL) were determined in the analyzed tissues (Table 2). Our results demonstrated that ASMT1 was the one most highly transcribed in the eyes, with the transcription level decreasing in the following order: Sg > Sr > Sa. Its transcription pattern is consistent with the habitat conditions of these fishes, since Sg is surface-dwelling, Sr is semi-cave dwelling and Sa is cave-restricted [24]. However, ASMTL was widely distributed in the examined tissues (eye, skin, liver and gonad) of Sr and Sa, although with differential transcription levels; whereas no ASMTL mRNA was detectable in Sg. Transcription of ASMT2a was observed in these *Sinocyclocheilus* fishes while ASMT2b was only highly transcribed in Sa. Meanwhile, we compared the transcriptome data from multiple tissues (liver, muscle, skin, gill and brain) of two mudskippers (Table 3). We observed that ASMT2 and ASMTL displayed a wide distribution in the tissues of BP and PM, whereas ASMT1 mRNA was only detectable in the brain (including the pineal gland) and the liver, with lower expression in the liver as compared to the brain.

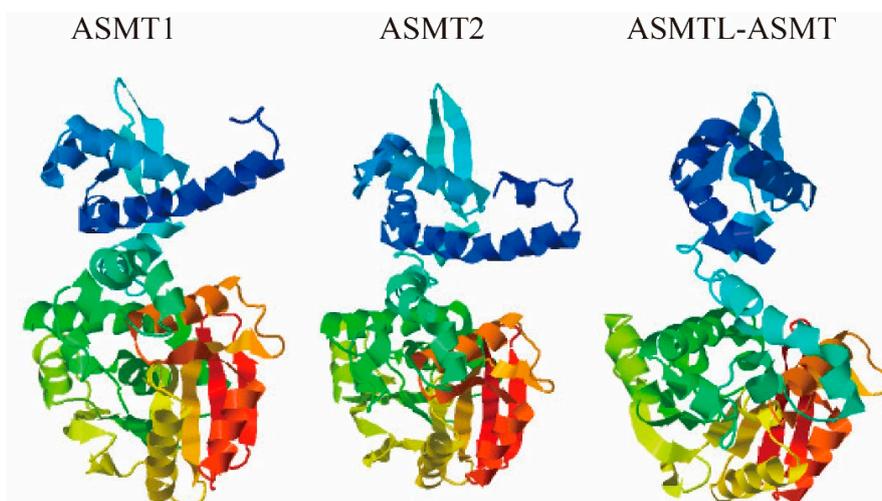
Molecular cloning confirmed the cDNA sequences for ASMT1, ASMT2, and ASMTL genes derived from the genome of BP (corresponding accession numbers: MF787797, MF285271, and MF787798). As expected, BP-ASMT1, BP-ASMT2, and BP-ASMTL encode proteins with 344, 347 and 605 residues respectively.

Table 3. RPKMs of *ASMT* transcripts in two mudskippers.

Tissue	Gene	PM	BP
Brain	<i>ASMT1</i>	1.31	0.98
	<i>ASMT2</i>	8.47	2.43
	<i>ASMTL</i>	2.05	4.51
Gill	<i>ASMT1</i>	0.00	0.00
	<i>ASMT2</i>	2.43	2.43
	<i>ASMTL</i>	2.76	1.54
Liver	<i>ASMT1</i>	0.00	0.37
	<i>ASMT2</i>	3.09	1.73
	<i>ASMTL</i>	7.15	1.63
Muscle	<i>ASMT1</i>	0.00	0.00
	<i>ASMT2</i>	5.72	1.07
	<i>ASMTL</i>	3.01	1.64
Skin	<i>ASMT1</i>	0.00	0.00
	<i>ASMT2</i>	6.47	2.37
	<i>ASMTL</i>	4.93	0.85

2.6. Predicted Three-Dimensional (3D) Structures of Fish *ASMT*s

The I-TASSER Suite was employed to predict 3D structures and potential functions of fish *ASMT* proteins. Three 3D atomic models with high accuracy for *ASMT1*, *ASMT2* and *ASMTL-ASMT* in BP were generated (Figure 6), and their corresponding C-score values were calculated to be 1.58, 1.56, and 1.28, respectively. Similar structures suggested that all the three *ASMT* proteins may possess conserved ligand-binding sites to realize similar functions. We subsequently matched these predicted models to all available structures in the public PDB library, and interestingly, the best hit with the closest structure similarity occurred to the human *ASMT* protein (PDB ID: 4a6dA). The overall fold of the predicted fish *ASMT* dimers is very similar to the conditions in bacteria and plant *O*-methyltransferases (*O*-MTs), such as isoflavone *O*-MT and chalcone *O*-MT from alfalfa and isoflavone *O*-MT from barrel medic [14], which are SAM-dependent methyltransferases associated with secondary metabolism in alfalfa [25]. However, previous reports [14,26] suggested that the *ASMT* structures are in a “closed” configuration (similar to Figure 6), while chalcone *O*-MT from alfalfa and isoflavone *O*-MT from barrel medic are in an “open” construction.

**Figure 6.** Predicted 3D models of *ASMT1*, *ASMT2* and *ASMTL-ASMT* from BP.

3. Discussion

In this paper, we investigated many aspects of the fish *ASMT* family, and provide new information relating to the diversity, structural difference and tissue distribution of *ASMTs* in fish, from an integrated view of genomic and transcriptomic levels. The phylogenetic analysis, along with the synteny comparison, amino acid sequence alignment and protein structure prediction, revealed the presence of two putative *ASMT* isotypes (*ASMT1* and *ASMT2*) and a new *ASMTL* in teleost genomes. Meanwhile, we cloned the full-length cDNAs of *ASMT1*, *ASMT2* and *ASMTL* in the representative mudskipper BP.

Previous studies confirmed that all teleosts have gone through at least three rounds (Rs) of whole-genome duplication (WGD), with 1R and 2R before divergence of ray-finned fishes from jawed vertebrates [27–29]; some fish families even have undergone a fourth WGD [30,31]. According to our present results, we propose that variations in the number of *ASMT* genes in different fishes may be a result of WGD and gene loss, as in the case of arylalkylamine *N*-acetyltransferase (AANAT) [32,33], which is another important rate-limiting enzyme for melatonin biosynthesis. Tetrapods have only one *ASMT* gene, whereas the examined diploid fishes possess three isoforms and tetraploid teleosts like the *Sinocyclocheilus* fishes have even more copies of *ASMT* genes. The ratio of *ASMT* gene number between diploids and tetraploids is not always 1:2, due to possible gene loss as occurred in fish AANATs [32]. Moreover, the synteny regions of *ASMT1*, *ASMT2*, and *ASMTL* were well conserved across species, respectively. Interestingly, our data demonstrated that *ASMT1*, and *ASMTL* are located on the same chromosome, especially in mammals *ASMT1* and *ASMTL* are very close in localization. The *ASMT1* and *ASMT2* shared a conserved exon-intron pattern and the exons 8–13 of *ASMTL* are homologous to *ASMT1* and *ASMT2*.

In general, the WGD duplicates can escape the fate of gene loss when they obtain subfunctionalization or neofunctionalization [34,35]. For *ASMTs*, both putative isotypes (*ASMT1* and *ASMT2*) are present in most fishes, indicating that some subfunctionalization or neofunctionalization has occurred in these species. All of the *ASMT* enzymes and *ASMTL*-*ASMT* displayed strict conservation around the SAM-binding site, while differences were found at the NAS-binding site (no more than 73% identity). These differences may influence the stability and substrate selectivity of these enzymes. Altogether, these results suggest that *ASMT1*, *ASMT2* and *ASMTL* may play differential functions in teleosts.

For a better understanding of the respective roles of *ASMTs*, we investigated their tissue distribution in several representative fish species. Our data indicated that *ASMT2* and *ASMTL* showed a wide transcriptional distribution, while *ASMT1* transcript was only localized in the eyes and the brain (not dissected from the pineal gland in our work), which are the main sources of melatonin biosynthesis. The wide presence of *ASMT2* transcript in fish peripheral organs is consistent with previous findings in goldfish [15,36,37], which proposed an earlier existence of the melatonin synthesizing system in peripheral tissues. A previous study also suggested a ubiquitous expression of *ASMTL* [22], which is similar to our present report. Conversely, the transcription of *ASMT1* was much more limited. As presented in our work, the eyes (including retinae) possess the highest transcription level of *ASMT1* among the tissues investigated in the three *Sinocyclocheilus* species, and only the brain (including the pineal gland) contained abundant *ASMT1* transcripts in the two mudskipper species. In agreement with a previous study in sea bass [10], our results may support the idea that *ASMT1* is mainly present in retinae and the pineal gland in fish for melatonin synthesis. It is interesting to note that the level of *ASMT1* transcription distinctly decreased in the following order: Sg > Sr > Sa, which is possibly associated with the regressive features of cave-restricted fish (such as eye degeneration and lack of circadian rhythms in Sa) [24]. In contrast, the transcription levels of *ASMTL* and *ASMT2b* were very high in the examined tissues of Sa as compared to those in Sg and Sr, suggesting that they may contribute to the compensatory evolution [24] and *ASMTL* may have a substrate other than NAS.

Generally speaking, protein structure determines protein's function [38]. In order to support the possible function differences among the two putative *ASMT* isotypes and *ASMTL*-*ASMT* of

ASMTL, we aligned their protein sequences and predicted related 3D structures (Figures 5 and 6). We observed that these residues around the SAM binding sites are highly conserved among teleost ASMT1 and ASMT2, with only a few variations being noticed between them (such as N17E, I208V and R291S). Previous studies reported that these kinds of variations might play a strong impact on the ASMT activity [14]. A previous finding also suggested that ASMTL may be not associated with the melatonin pathway. Although the exact function of ASMTL is unclear, remarkable conservation of the putative catalytic domain for SAM binding support a methyltransferase activity of this enzyme. The notable difference distinguishing ASMTL-ASMT from ASMT1 and ASMT2 was located at the NAS binding sites, and some variations occurred (such as Y108H, N302S, E311Q), suggesting that ASMTL may have new function(s) in term of substrate selectivity.

4. Materials and Methods

4.1. Acquisition of ASMT and ASMTL for Nucleotide and Protein Sequences

A total of 40 vertebrate species, including 33 fish species were examined for our present research. These data were obtained by two ways. First, published ASMT and ASMTL sequences were downloaded from the public databases GenBank and Ensembl (Table S1). Second, unreported ASMT and ASMTL sequences from 21 fish species were derived from whole genome data generated by us and our collaborators. In detail, potential homology-based ASMT genes were retrieved from fish genomes using tBLASTn [39] with an e-value of 10^{-5} . The BLAST results were subsequently processed by a perl script to obtain the best hit of each alignment. Finally, GeneWise v2.2.0 [40] was employed to predict the ASMT and ASMTL genes from the best hits.

4.2. Sequence Alignment and Phylogenetic Analysis

Nucleotide and protein sequences of all of the collected ASMT and ASMTL genes were used for further phylogenetic analysis. In brief, MAFFT software [41] was employed to align protein sequences of ASMTs and a Maximum Likelihood (ML) phylogenetic analysis was performed using RAxML 8.0.17 [42,43]. Additionally, ML phylogenetic trees of the three ASMT isotypes were constructed using their corresponding coding sequences by FastTree v2.1.7 [44]. We also downloaded from the public Protein Data Bank (PDB), a protein model of human ASMT (PDB code: 4A6D) for comparing structural differences among the fish ASMTs.

4.3. Analyses of Conserved Synteny and Gene Structures

For evaluating the conservation of ASMT genes, we checked several genes residing in the upstream and the downstream regions of each ASMT paralog. Related genomic data were obtained from GenBank and our lab as mentioned above. The stickleback (*Gasterosteus aculeatus*) genome [45] was used as the reference standard for searching any ASMT upstream and downstream regions. The genome assemblies of different fish species were searched using BLAST software, and the best hit was selected using a Perl script. GeneWise v2.2.0 [40] was employed to predict the ASMT gene structures.

4.4. Molecular Cloning of Mudskipper ASMT1, ASMT2 and ASMTL Transcripts

Total RNA from the eyes of blue-spotted mudskipper (*Boleophthalmus pectinirostris*, BP) [46] was extracted with TRIzol reagent (Invitrogen, Carlsbad, CA, USA). It was subsequently reverse-transcribed using the M-MuLV First Strand cDNA Synthesis Kit (Sangon, Shanghai, China) in 25- μ L reactions, which were carried out for 50 min at 42 °C, followed by 15 min at 72 °C to inactivate the reverse transcriptase. These synthesized cDNAs were amplified using Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) with specific primer pairs (Table S2) on a ABI 9700 thermal cycler (Life Technologies, Carlsbad, CA, USA) under the following cycling conditions: initial denaturation at 94 °C for 5 min; then 35 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s; and, final extension at 72 °C for 10 min. The amplified products

were resolved in a 1.5% agarose gel, purified using a SanPrep Column DNA Gel Extraction Kit (Sangon, Shanghai, China), and sub-cloned into a pGEM-T Easy vector (Promega, Madison, WI, USA). Several positive clones of *E. coli* DH5 α cells were picked for subsequent purification and sequencing validation.

4.5. Acquisition of Transcriptomic Data and Quantification of ASMT Transcripts

To investigate the tissue distribution of the ASMT genes, we selected two representative fish groups, tetraploid *Sinocyclocheilus* species and diploid mudskippers, for transcriptome analysis. Related transcriptomic data of four tissues (eye, skin, liver, and ovary) from three *Sinocyclocheilus* species that were previously generated by our lab [24] and have been deposited at NCBI Sequence Read Archive (SRA; accession numbers: *S. graham* (Sg), SRS1179797 to SRS1179800; *S. rhinoceros* (Sr), SRS1179996 to SRS1179999; and *S. anshuiensis* (Sa), SRS1180000 to SRS1180003). In addition, we previously produced transcriptomic data of five tissues (gill, skin, liver, muscle, and brain) from two representative mudskippers, BP and PM [46]. The Cufflink program 2.1.1 [47] with the core parameters (-FDR 0.05, -geometric-norm TRUE, -compatible-hits-norm TRUE) was employed to calculate the RPKM values of each ASMT gene, which are comparative parameters to quantify the relative transcription levels.

4.6. Tertiary Structure and Functional Prediction of Each ASMT Protein

As described previously [48], I-TASSER was employed to predict the tertiary structure and function of each ASMT. The confidence of models is quantitatively measured by a C-score that is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. A C-score is typically in the range of [-5, 2], where a high C-score supports the corresponding model with a high confidence.

5. Conclusions

In summary, we provide some novel insights into the fish ASMT family by a combination of comparative genomic and transcriptomic studies. We confirmed the existence of ASMT1, ASMT2 and ASMTL in fish, and reported variations in the number of genes and the tissue distribution between species. We demonstrated that the C-terminal part of ASMTL (the ASMTL-ASMT region) is homology to the full length of ASMT1, and ASMT2, which is consistent with a previous report that two genes (*maf* and ASMT) were joined to form the ASMTL by gene fusion and duplication events. Our results also offer solid evidence to support previous findings that ASMT1 is preferentially expressed in fish retinae and pineal gland, while ASMT2 and ASMTL were mainly expressed in the fish peripheral tissues (such as liver, gut, skin, and gonad).

Supplementary Materials: The following are available online. Figure S1. Alignment of ASMT exon sequences from zebrafish. Table S1. The accession numbers for ASMT sequences. Table S2. Primer sequences used for molecular cloning of ASMTs and ASMTL.

Acknowledgements: This work was supported by National Natural Science Foundation of China (No. 31370047), Shenzhen Special Program for Development of Emerging Strategic Industries (No. JSGG20170412153411369), Shenzhen Dapeng Special Program for Industrial Development (No. KY20170204), the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development at the USA National Institutes of Health, and Zhenjiang Leading Talent Program for Innovation and Entrepreneurship.

Author Contributions: Qiong Shi and Kai Zhang conceived and designed the experiments. Kai Zhang performed the experiments. Kai Zhang, Steven L. Coon and Qiong Shi prepared the manuscript. Zhiqiang Ruan, Jia Li, Xinxin You and Chao Bian participated in discussions and provided suggestions.

Conflicts of Interest: The authors have declared that they have no conflicts of interest.

References

1. Falcón, J.; Besseau, L.; Sauzet, S. Melatonin effects on the hypothalamopituitary axis in fish. *Trends Endocrinol. Metab.* **2007**, *18*, 81–88.
2. Klein, D.C. Arylalkylamine *N*-acetyltransferase: “the Timezyme”. *J. Biol. Chem.* **2007**, *282*, 4233–4237.
3. Klein, D.C.; Coon, S.L.; Roseboom, P.H. The melatonin rhythm-generating enzyme: Molecular regulation of serotonin *N*-acetyltransferase in the pineal gland. *Recent Prog. Horm. Res.* **1997**, *52*, 307–358.
4. Falcon, J. Cellular circadian clocks in the pineal. *Prog. Neurobiol.* **1999**, *58*, 121–162.
5. Falcon, J.; Besseau, L.; Boeuf, G. Molecular and cellular regulation of pineal organ responses. *Fish Physiol.* **2006**, *25*, 243–306.
6. Liu, T.; Borjigin, J. *N*-acetyltransferase is not the rate-limiting enzyme of melatonin synthesis at night. *J. Pineal Res.* **2005**, *39*, 91–96.
7. Reiter, R.J.; Tan, D.; Terron, M.P. Melatonin and its metabolites: New findings regarding their production and their radical scavenging actions. *Acta Biochim. Pol.* **2007**, *54*, 1.
8. Park, S.; Byeon, Y.; Back, K. Functional analyses of three ASMT gene family members in rice plants. *J. Pineal Res.* **2013**, *55*, 409–415.
9. Wang, L.; Li, J.; Ruan, Y. Sequencing ASMT identifies rare mutations in Chinese Han patients with autism. *PLoS ONE* **2013**, *8*, e53727.
10. Paulin, C.H.; Cazaméa-Catalan, D.; Zilberman-Peled, B. Subfunctionalization of arylalkylamine *N*-acetyltransferases in the sea bass *Dicentrarchus labrax*: Two-ones for one two. *J. Pineal Res.* **2015**, *59*, 354–364.
11. Kang, K.; Kong, K.; Park, S. Molecular cloning of a plant *N*-acetylserotonin methyltransferase and its expression characteristics in rice. *J. Pineal Res.* **2011**, *50*, 304–309.
12. Park, S.; Byeon, Y.; Kim, Y.S. Kinetic analysis of purified recombinant rice *N*-acetylserotonin methyltransferase and peak melatonin production in etiolated rice shoots. *J. Pineal Res.* **2013**, *54*, 139–144.
13. Rodriguez, I.R.; Mazuruk, K.; Schoen, T.J. Structural analysis of the human hydroxyindole-*O*-methyltransferase gene. Presence of two distinct promoters. *J. Biol. Chem.* **1994**, *269*, 31969–31977.
14. Botros, H.G.; Legrand, P.; Pagan, C. Crystal structure and functional mapping of human ASMT, the last enzyme of the melatonin synthesis pathway. *J. Pineal Res.* **2013**, *54*, 46–57.
15. Velarde, E.; Cerdá-Reverter, J.M.; Alonso-Gómez, A.L. Melatonin-synthesizing enzymes in pineal, retina, liver, and gut of the goldfish (*Carassius*): mRNA expression pattern and regulation of daily rhythms by lighting conditions. *Chronobiol. Int.* **2010**, *27*, 1178–1201.
16. Khan, Z.A.; Yumnamcha, T.; Rajiv, C. Melatonin biosynthesizing enzyme genes and clock genes in ovary and whole brain of zebrafish (*Danio rerio*): Differential expression and a possible interplay. *Gen. Comp. Endocrinol.* **2016**, *233*, 16–31.
17. Muñoz-Pérez, J.L.; López-Patiño, M.A.; Álvarez-Otero, R. Characterization of melatonin synthesis in the gastrointestinal tract of rainbow trout (*Oncorhynchus mykiss*): Distribution, relation with serotonin, daily rhythms and photoperiod regulation. *J. Comp. Physiol. B* **2016**, *186*, 471–484.
18. Sanjita Devi, H.; Rajiv, C.; Mondal, G. Melatonin bio-synthesizing enzyme genes (*Tph1*, *Aanat1*, *Aanat2*, and *Hiomt*) and their temporal pattern of expression in brain and gut of a tropical carp in natural environmental conditions. *Cogent Biol.* **2016**, *2*, 1230337.
19. Mukherjee, S.; Maitra, S.K. Gut melatonin in vertebrates: Chronobiology and physiology. *Front. Endocrinol.* **2015**, *6*, 112.
20. Velarde, E.; Delgado, M.J.; Alonso-Gómez, A.L. Serotonin-induced contraction in isolated intestine from a teleost fish (*Carassius auratus*): Characterization and interactions with melatonin. *Neurogastroenterol. Motil.* **2010**, *22*, e364–e373.
21. Velarde, E.; Haque, R.; Iuvone, P.M. Circadian clock genes of goldfish, *Carassius auratus*: cDNA cloning and rhythmic expression of period and cryptochrome transcripts in retina, liver, and gut. *J. Biol. Rhythm.* **2009**, *24*, 104–113.
22. Ried, K.; Rao, E.; Schiebel, K. Gene duplications as a recurrent theme in the evolution of the human pseudoautosomal region 1: Isolation of the gene ASMTL. *Hum. Mol. Genet.* **1998**, *7*, 1771–1778.
23. Sun, Y.; Huang, Y.; Li, X. Fish-T1K (Transcriptomes of 1,000 Fishes) Project: Large-scale transcriptome data for fish evolution studies. *GigaScience* **2016**, *5*, 18.
24. Yang, J.; Chen, X.; Bai, J. The *Sinocyclocheilus* cavefish genome provides insights into cave adaptation. *BMC Biol.* **2016**, *14*, 1.

25. Zubieta, C.; He, X.Z.; Dixon, R.A. Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. *Nat. Struct. Mol. Biol.* **2001**, *8*, 271–279.
26. Liu, C.J.; Deavours, B.E.; Richard, S.B. Structural basis for dual functionality of isoflavonoid O-methyltransferases in the evolution of plant defense responses. *Plant Cell* **2006**, *18*, 3656–3669.
27. Jaillon, O.; Aury, J.M.; Brunet, F. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **2004**, *431*, 946–957.
28. Kasahara, M.; Naruse, K.; Sasaki, S. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **2007**, *447*, 714–719.
29. Nakatani, Y.; Takeda, H.; Kohara, Y. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* **2007**, *17*, 1254–1265.
30. Mayfield-Jones, D.; Washburn, J.D.; Arias, T. Watching the grin fade: Tracing the effects of polyploidy on different evolutionary time scales. *Semin. Cell Dev. Biol.* **2013**, *24*, 320–331.
31. Li, J.; You, X.; Bian, C. Molecular evolution of aralkylamine N-acetyltransferase in fish: A genomic survey. *Int. J. Mol. Sci.* **2015**, *17*, 51.
32. Davidson, W.S.; Koop, B.F.; Jones, S.J.M. Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol.* **2010**, *11*, 403.
33. Zilberman-Peled, B.; Bransburg-Zabary, S.; Klein, D.C. Molecular evolution of multiple arylalkylamine N-acetyltransferase (AANAT) in fish. *Mar. Drugs* **2011**, *9*, 906–921.
34. Force, A.; Lynch, M.; Pickett, F.B. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **1999**, *151*, 1531–1545.
35. Lien, S.; Koop, B.F.; Sandve, S.R. The Atlantic salmon genome provides insights into rediploidization. *Nature* **2016**, *533*, 200–205.
36. Chowdhury, I.; Maitra, S.K.; Verster, G.C. Melatonin in the Promotion of Health. In *Melatonin Time Line: From Discovery to Therapy*; Taylor and Francis: Boca Raton, FL, USA, 2012; pp. 1–60.
37. Jimenez-Jorge, S.; Guerrero, J.M.; Jimenez-Caliani, A.J. Evidence for melatonin synthesis in the rat brain during development. *J. Pineal Res.* **2007**, *42*, 240–246.
38. Skolnick, J.; Fetrow, J.S. From genes to protein structure and function: Novel applications of computational approaches in the genomic era. *Trends Biotechnol.* **2000**, *18*, 34–39.
39. Pevsner, J. Basic Local Alignment Search Tool (BLAST). In *Bioinformatics and Functional Genomics*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2009; pp. 100–138.
40. Birney, E.; Clamp, M.; Durbin, R. GeneWise and Genomewise. *Genome Res.* **2004**, *14*, 988–995.
41. Yamada, K.D.; Tomii, K.; Katoh, K. Application of the MAFFT sequence alignment program to large data-reexamination of the usefulness of chained guide trees. *Bioinformatics* **2016**, *32*, 3246–3251.
42. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690.
43. Stamatakis, A.; Hoover, P.; Rougemont, J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* **2008**, *57*, 758–771.
44. Price, M.N.; Dehal, P.; Arkin, A.P. FastTree: Computing Large Minimum-Evolution Trees with Profiles instead of a Distance Matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650.
45. Jones, F.C.; Grabherr, M.G.; Chan, Y.F. The genomic basis of adaptive evolution in three spine sticklebacks. *Nature* **2012**, *484*, 55–61.
46. You, X.; Bian, C.; Zan, Q. Mudskipper genomes provide insights into the terrestrial adaptation of amphibious fishes. *Nat. Commun.* **2014**, *5*, 5594.
47. Trapnell, C.; Williams, B.A.; Pertea, G. Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat. Biotechnol.* **2010**, *28*, 511–515.
48. Yang, J.; Yan, R.; Roy, A. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8.

Sample Availability: Samples of the compounds are available from the authors.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).