



HAL
open science

The Timing of Timezyme Diversification in Vertebrates

Damien Cazaméa-Catalan, Laurence Besseau, Jack Falcón, Elodie Magnanou

► **To cite this version:**

Damien Cazaméa-Catalan, Laurence Besseau, Jack Falcón, Elodie Magnanou. The Timing of Timezyme Diversification in Vertebrates. PLoS ONE, 2014, 9 (12), pp.e112380. 10.1371/journal.pone.0112380 . hal-01345054

HAL Id: hal-01345054

<https://hal.sorbonne-universite.fr/hal-01345054>

Submitted on 13 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

The Timing of Timezyme Diversification in Vertebrates

Damien Cazaméa-Catalan^{1,2}, Laurence Besseau^{1,2}, Jack Falcón^{1,2}, Elodie Magnanou^{1,2*}

1. Sorbonne Universités, UPMC Univ Paris 06, UMR 7232, BIOM Biologie Intégrative des Organismes Marins, Observatoire Océanologique, Banyuls/Mer, France, 2. CNRS, UMR 7232, BIOM Biologie Intégrative des Organismes Marins, Observatoire Océanologique, Banyuls/Mer, France

*elodie.magnanou@obs-banyuls.fr

Abstract

All biological functions in vertebrates are synchronized with daily and seasonal changes in the environment by the time keeping hormone melatonin. Its nocturnal surge is primarily due to the rhythmic activity of the arylalkylamine *N*-acetyl transferase AANAT, which thus became the focus of many investigations regarding its evolution and function. Various vertebrate isoforms have been reported from cartilaginous fish to mammals but their origin has not been clearly established. Using phylogeny and synteny, we took advantage of the increasing number of available genomes in order to test whether the various rounds of vertebrate whole genome duplications were responsible for the diversification of AANAT. We highlight a gene secondary loss of the AANAT2 in the Sarcopterygii, revealing for the first time that the AAANAT1/2 duplication occurred before the divergence between Actinopterygii (bony fish) and Sarcopterygii (tetrapods, lobe-finned fish, and lungfish). We hypothesize the teleost-specific whole genome duplication (WGD) generated the appearance of the AANAT1a/1b and the AANAT2/2' paralogs, the 2' isoform being rapidly lost in the teleost common ancestor (ray-finned fish). We also demonstrate the secondary loss of the AANAT1a in a Paracantopterygii (Atlantic cod) and of the 1b in some Ostariophysi (zebrafish and cave fish). Salmonids present an even more diverse set of AANATs that may be due to their specific WGD followed by secondary losses. We propose that vertebrate AANAT diversity resulted from 3 rounds of WGD followed by previously uncharacterized secondary losses. Extant isoforms show subfunctionalized localizations, enzyme activities and affinities that have increased with time since their emergence.



OPEN ACCESS

Citation: Cazaméa-Catalan D, Besseau L, Falcón J, Magnanou E (2014) The Timing of Timezyme Diversification in Vertebrates. PLoS ONE 9(12): e112380. doi:10.1371/journal.pone.0112380

Editor: Yoav Gothliff, Tel Aviv University, Israel

Received: July 8, 2014

Accepted: October 15, 2014

Published: December 8, 2014

Copyright: © 2014 Cazaméa-Catalan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: Agence Nationale de la Recherche (grant number: ANR 07-BLAN-0097-01), JF. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Melatonin is a key hormonal regulator of the biological clock in vertebrates. In all vertebrates investigated so far blood melatonin levels are higher at night than during day and the amplitude of this rhythm varies with season, often reflecting the ambient temperature [1, 2]. Thus the daily and annual variations in the circulating levels of this hormone provide information about the time of day and season, allowing vertebrates to synchronize their metabolism, physiology and behavior to the cyclic variations of the environment [3, 4]. This circulating melatonin is produced by the pineal gland. Melatonin is also synthesized by the vertebrates' retina, where it exerts autocrine/paracrine functions [5]. Melatonin is synthesized from tryptophan in four enzymatic steps. The first is catalyzed by tryptophan hydroxylase (EC 1.14.16.4), producing 5-hydroxytryptophan, which is converted into serotonin by 5-hydroxytryptophan decarboxylase (EC 4.1.1.28). Serotonin is then acetylated by arylalkylamine *N*-acetyltransferase (AANAT, EC 2.3.1.87) to produce *N*-acetylserotonin. The last step, the methylation of *N*-acetylserotonin, is catalyzed by hydroxyindole-*O*-methyltransferase (HIOMT, EC 2.1.1.4.) and produces melatonin [1, 6]. AANAT is important in this pathway because it drives the rhythm of melatonin production, thus earning its nickname “The Timezyme” [7].

AANAT, also named VT-AANAT (for vertebrate AANAT), appeared in vertebrates following the duplication and divergence of an ancestral gene, NV-AANAT (non-vertebrate AANAT), after the Cephalochordate (amphioxus) – vertebrate split over half a billion years ago [8]. The emergence of VT-AANAT was followed by a dramatic acceleration of evolution that accompanied neofunctionalization after the duplication event. This scenario is consistent with the hypotheses that the advent of VT-AANAT contributed to the evolution of the pineal gland and lateral eyes from a common ancestral photodetector [8]. While NV-AANAT disappeared early in vertebrate evolution, the appearance of VT-AANAT marked a crucial step in the development of photodetection and time-keeping in vertebrates [8, 9].

Pineal melatonin secretion is nocturnal in all vertebrates. This conserved pattern masks a unique and dramatic evolution of the mechanisms that regulate its production, both at cellular and organismal levels - see details below and [10]. At the molecular level the regulation of *aanat* gene expression and AANAT enzyme activity are mediated through different pathways of light perception and transmission in tetrapods *vs.* teleosts (ray-finned fish). In addition, while all vertebrates from Cyclostomata (jawless fish) to tetrapods possess one single AANAT, Actinopterygii possess two, or even three, enzymes. Coon and Klein hypothesized that AANAT1 and AANAT2 duplicates arose at the base of the Actinopterygii (bony fish) clade [11]. Interestingly, while AANAT1 is more related to the AANAT from other vertebrates in terms of sequence, AANAT2 has no equivalent [11, 12]. Moreover, AANAT2 is specifically expressed in the pineal gland, while AANAT1 is more specifically expressed in the retina and some brain areas. Finally, some teleosts also express two AANAT1 genes, AANAT1a and

AANAT1b. It has been postulated that the emergence of these two forms arose at the split between Ostariophysi and Protacanthopterygii [11].

To date, our knowledge about the diversification of AANAT in teleosts is based on too few species, allowing no definitive conclusions. Here we used both phylogeny and synteny, taking advantage of the increasing number of available genomes, in order to disentangle the effects of repeated vertebrate WGD on AANAT diversification. We reexamine the current theory about AANAT diversification in vertebrates, paying particular attention to the timing of duplication of each AANAT isoform. We show new evidence for various gene secondary losses, giving a more precise scenario not only brings more insight into the evolutionary history of a key vertebrate enzyme, but also allows a better understanding of the crucial steps that may have driven the diversification of the melatonin signaling and synthesis pathways as a whole.

Material and Methods

Data Collection

NCBI, Ensembl and genomic databases were screened for publically available AANAT isoforms. Our search was exhaustive for Actinopterygians, but we selected only a subsample of the available AANATs, representative of Sarcopterygii taxa diversity. Cephalochordates, Chondrichthyes and Cyclostomata were also included. Sequences were manually assembled and reannotated whenever required, based on homology with known AANATs.

Assessments of Conserved Synteny

To examine patterns of conserved synteny, we annotated six genes found upstream and downstream of each *aanat* paralog in the stickleback genome (*Gasterosteus aculeatus*), used as a reference. Their synteny was compared to all other considered species; *i.e.* eight teleosts: the tetraodon (*Tetraodon nigroviridis*), medaka (*Oryzias latipes*), platyfish (*Xiphophorus maculatus*), tilapia (*Oreochromis niloticus*), Atlantic cod (*Gadus morhua*), zebrafish (*Danio rerio*), and Atlantic salmon (*Salmo salar*). We broaden this approach to one non-teleost Actinopterygii the spotted gar (*Lepisosteus oculatus*) and to four Sarcopterygii including the coelacanth (*Latimeria chalumnae*) and three tetrapods: the Western clawed-frog (*Xenopus tropicalis*), chicken (*Gallus gallus*) and human (*Homo sapiens*). Initial ortholog predictions were visualized using Genomicus v75.01, using its web page interface [13]. To identify additional unannotated genes lying upstream and downstream of the annotated AANAT genes, we also used the program Genscan [14]. The unannotated genes were compared with the non-redundant protein database using the Basic Local Alignment Search Tool (BLAST) [15].

The Atlantic salmon genome was visualized in Genomicus. The salmon genomic data (AGKD00000000.1 [16] was loaded from NCBI (<http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AGKD01>) into Geneious Pro 5.4.6 [17], using the

Add Sequence Database tool. Contigs containing AANAT isoforms were searched with the Sequence search Tool. Surrounding genes were identified using geneScan and Blast as described previously.

We also screened the genome of the species of interest in search of putative paralogs of AANATs neighboring genes.

To facilitate comparisons, we report the order of genes in the same orientation as they appear in the Stickleback genome assembly, regardless of how they are found in the Ensembl database.

Sequence Alignment

Amino-acid sequences of known or predicted AANATs were aligned using ClustalW [18] displayed in MEGA 5 [19], and manually adjusted. The alignment was then imported in Geneious Pro 5.4.6 [17].

Phylogenetic Analyses

Before characterizing the genomic organization of the AANAT genes in nine Actinopterygii (eight teleosts and the spotted gar) and four Sarcopterygii, we performed phylogenetic analyses to reconstruct the duplicative history of the AANAT various isoforms. For this analysis, we added all available Actinopterygii AANAT genes and included sequences from representative of the tetrapods and Chondrichthyes clades for comparisons. We extended our analysis to NV-AANAT forms from Chondrichthyes and Cephalochordates.

Some AANAT isoforms, especially the VT and NV subfamilies, present a high degree of divergence for some regions between investigated species. Thus we selected conserved blocks of the alignment using Gblocks [20], giving a matrix of 167 amino acids. The amino acid alignment was analyzed by maximum likelihood (ML). The model of amino acid substitution was selected in Prottest webserver [21]. The best model based on the Akaike Information Criterion (AIC) was JTT+G ($\alpha = 0.76$) and was implemented in PhyML displayed in Geneious for further ML reconstructions [22]. Node support was assessed with the bootstrap technique, using 1,000 replicates. Bayesian phylogenetic analyses were also performed on this amino acid matrix using MrBayes 3.0 [23] implemented in Geneious. The Bayesian analysis was performed with the Metropolis-coupled Markov chain Monte Carlo algorithm using a Mixed Model. The tree-space was explored by using four chains run during 4 million generations and saving every 100th tree. The first 25,000 trees were discarded (Burnin') based on preliminary runs that allowed identification of the point when the chains became stationary. A consensus was built with the remaining trees. Bayesian probabilities were used to evaluate branch support.

Results

AANAT Isoforms in vertebrates

A total of 90 sequences were selected for this study, 67 isoforms were found for Actinopterygii, 14 were selected to represent Sarcopterygii taxonomic diversity, 8 for Chondrichthyes and Cyclostomata, and one for Cephalochordates. Sequence sources and their annotations are provided in [Table 1](#).

Chondrichthyes are characterized by the presence of two AANAT families: the vertebrate VT- and the NV-AANAT ([Fig. 1](#)). Only the VT form was reported in the different lamprey genomes available (Cyclostomata). The amphioxus *Branchiostoma floridae* possesses seven NV-AANATs. The beta form was selected based on its high degree of similarity with Chondrichthyes NV-AANAT, and was used as an outgroup in the phylogeny. Actinopterygii also possess two enzyme subfamilies: the AANAT1 and the AANAT2. Within this clade Acanthopterygii possessed up to three forms: the AANAT1a, AANAT1b and AANAT2. Note that *aanat1b* was not found in any Ostariophysi, including the two species with available genomes (*Danio rerio* and *Cyprinus carpio*). In the same way, no *aanat1a* was ever reported in Paracanthopterygii, neither could we locate any *aanat1a* in the *Gadus morhua* genome. Up to five *aanat* genes were identified in the salmonid *Salmo salar* genome including two versions of the *aanat2* and *aanat1b*, and one version of the *aanat1a* gene. Only the AANAT1 has been documented in Sarcopterygii that encompass tetrapods and Actinistia (coelacanth).

Phylogenetic Relationships among AANAT Genes

Amino acid based Bayesian and ML reconstructions were both performed using the beta NV-AANAT from *B. floridae* as an outgroup. Both techniques gave the same tree topology, showing only very marginal differences, and we chose to feature the Bayesian topology ([Fig. 2](#)) including node support provided by the Bayesian and the ML reconstructions (posterior probabilities and bootstrap values respectively). NV- and VT-AANATs are segregated in two robustly supported groups. The overall pattern of VT-AANAT structure shows three well supported clades (posterior probabilities and bootstrap resampling). The first one was constituted of the VT-AANAT form of Chondrichthyes, the second the Actinopterygii AANAT2 form, and the third one gathered all Euteleostomi isoforms 1 (bony vertebrates). Some subgroups are also well supported within the last two clades. AANAT2 grouped by the taxonomic position of the species they belong to: Ostariophysi (*A. mexicanus*; *C. auratus*; *C. carpio*; *D. rerio*; *O. sifiontesi*); Protacanthopterygii (*E. lucius* and the Salmonidae species); Paracanthopterygii (*G. morhua* and *A. glacialis*); and Pleuronectiforms (*S. maximus*; *P. olivaceus*; *S. senegalensis*). The AANAT1 from Euteleostomi were subdivided into two robust clades including either tetrapod AANATs or teleost AANAT1b, but teleost AANAT1a forms were not robustly grouped.

Table 1. Sequences accession number for the different AANAT isoforms.

Species	AANAT	Accession number		Remarks/source
		Amino acid sequence	Nucleotide sequence	
<i>Acipenser sturio</i>	1	AAAY78956.1	DQ077733.1	NCBI
	2	AAAY78957.1	DQ077734.1	Partial/NCBI
<i>Anguilla anguilla</i>	1	scaffold 6643	Anguilla_anguilla_genome_v1_09_no-v_10	Partial/ www.zfgenomics.com
	2	scaffold639	Anguilla_anguilla_genome_v1_09_no-v_10	
<i>Anolis carolinensis</i>	1	ENSACAP00000008194	ENSACAT00000008370	Ensembl
<i>Arctogadus glacialis</i>	2	AFQ98400.1	JQ585748	NCBI
<i>Astyanax mexicanus</i>	1	ENSAMX-T00000016190	ENSAMXG00000015728	Partial/Ensemble
	2	ENSAMX-T00000016988	ENSAMXG00000016495	Ensemble
<i>Bos taurus</i>	1	NP_803475.1	NM_177509.2	NCBI
<i>Branchiostoma floridae</i>	NV beta	ACB14299.1	EU380676.1	NCBI
<i>Canis lupus familiaris</i>	1	NP_001070905.1	NM_001077437.1	NCBI
<i>Callorhynchus milii</i>	NV	AGX26123.1	KF246476.1	NCBI
	VT	AGX26122.1	KF246475.1	NCBI
<i>Carassius auratus</i>	2	ACZ65313.1	AB167078.1	NCBI
<i>Cyprinus carpio</i>	1	scaffold8139.1	carp_sept_scaffolds_v3.NEWALL.scaffolds	www.zfgenomics.com
	2	scaffold6738.1	carp_sept_scaffolds_v3.NEWALL.scaffolds	
<i>Chrysiptera cyanea</i>	1a	BAN63805.1	AB839763.1	NCBI
	1b	BAN63806.1	AB839764.1	NCBI
<i>Danio rerio</i>	1	AAQ54582.1	AY349158.1	NCBI
	2	NP_571486.1	NM_131411.1	NCBI
<i>Dicentrarchus labrax</i>	2			Besseau et al., Unpublished data
	1a	ACB13284.1	EU378922	NCBI
	1b	ACB13285.1	EU378923	NCBI
<i>Esox lucius</i>	1	AAD21316.1	AF034081.1	NCBI
	2	AAD21317.1	AF034082	NCBI
<i>Gadus morhua</i>	2	ENSGMO-P00000004524	ENSGMOT00000004660	Ensembl
	1	ENSGMO-P00000011553	ENSGMOT00000011867	Ensembl
<i>Gallus gallus</i>	1	NP_990489.1	NM_205158.1	NCBI
<i>Gasterosteus aculeatus</i>	2	ENSGAC-P00000018428	ENSGACT00000018464	Ensembl
	1a	ENSGAC-P00000025480	ENSGACT00000025530	Ensembl
	1b	ENSGAC-P00000009361	ENSGACT00000009381	Ensembl
<i>Homo sapiens</i>	1	NP_001160051.1	NM_001166579.1	NCBI
<i>Lampetra fluviatilis</i>	VT	AHA51209.1	KF290562.1	NCBI/partial
<i>Latimeria chalumnae</i>	1	ENSLACP00000017693	ENSLACT00000017823	Ensembl

Table 1. Cont.

Species	AANAT	Accession number		Remarks/source
		Amino acid sequence	Nucleotide sequence	
<i>Lepisosteus oculatus</i>	1	ENSGAC-P00000025480_1	ENSGACP00000025480_1	Partial
	2	ENSGAC-P00000018428_1	ENSGACP00000018428_1	Ensembl
<i>Lethenteron camtschaticum</i>	VT		unplaced genomic scaffold scaffold00451	NCBI
<i>Leucoraja erinacea</i>	NV	Contig1750401		skatebase.org
	VT	Contig244146, Contig28819 and Contig3043531+ LittleSkate_Transcriptome Contig50898		Partial/skatebase.org
<i>Macaca mulatta</i>	1	NP_001040592.1	NM_001047127.1	NCBI
<i>Mus musculus</i>	1	NP_033721.1	NM_009591.2	NCBI
<i>Oncorhynchus keta</i>	2	AAR07969.1	AY356363.1	Partial/NCBI
	1 brain	AAR07968.1	AY356362.1	NCBI
	1 retina	AAR07967.1	AY356361.1	NCBI
<i>Oncorhynchus mykiss</i>	2	AAD25333.1	AF106006.1	NCBI
	1 retina	BAA34809.1	AB007294.1	NCBI
<i>Oreochromis niloticus</i>	2	ENSONIP00000001811	ENSONIT00000001810	Ensembl
	1a	ENSONIP00000022987	ENSONIT00000023007	Ensembl
	1b	ENSONIP00000010440	ENSONIT00000010449	Ensembl
<i>Oryzias latipes</i>	2	NP_001098316.1	NM_001104846.1	NCBI
	1a	NP_001098302.1	NM_001104832.1	NCBI
	1b	NP_001098330.1	NM_001104860.1	NCBI
<i>Ovis aries</i>	1	NP_001009461.1	NM_001009461.1	NCBI
<i>Oxydoras sifontesi</i>	2	AFR32910.1	JQ478412.1	NCBI
<i>Pan troglodytes</i>	1	NP_001012442.1	NM_001012440.1	NCBI
<i>Paralichthys olivaceus</i>	2	AEK70413.1	HQ883478.1	NCBI
<i>Petromyzon marinus</i>	VT	AGX26120.1	KF246473.1	NCBI
<i>Psammomys obesus</i>	1	AAAY79348.1	DQ018371.1	NCBI
<i>Rana perezi</i>	1	/	/	Isorna et al., 2006
<i>Rattus norvegicus</i>	1	NP_036950.1	NM_012818.2	NCBI
<i>Salmo salar</i>	2	contig 354367743	AGKD01	http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AGKD01
	2	contig 354437952	AGKD01	
	1a	contig 354449422	AGKD01	
	1b	contig 354436144	AGKD01	
	1b	contig 354451409	AGKD01	
<i>Salvelinus alpinus</i>	2 VV	/	/	Cazamea-Cataman et al 2013
	2 IA	AFQ98280.1	JQ582623.1	NCBI
<i>Scophthalmus maximus</i>	1	ABM21526.1	EF033249.1	NCBI
	2	ABM21527.1	EF033250.1	NCBI
<i>Scyliorhinus canicula</i>	VT	ACB13283.1	EU378921	NCBI
<i>Siganus guttatus</i>	1	BAM83582.1	AB781600.1	NCBI

Table 1. Cont.

Species	AANAT	Accession number		Remarks/source
		Amino acid sequence	Nucleotide sequence	
<i>Solea senegalensis</i>	2	ACY66881.1	GQ340973.1	NCBI
	1a	ACY66879.1	GQ340971.1	NCBI
	1b	ACY66880.1	GQ340972.1	NCBI
<i>Sparus aurata</i>	1	AAT02159.1	AY533402.1	NCBI
	2	AAT02160.2	AY533403.2	NCBI
<i>Spheroides nephelus</i>	2	AAL73048.1	AY016023	NCBI
<i>Takifugu rubripes</i>	2	ENSTRUP00000029446	ENSTRUT00000029563	Ensembl
	1a	ENSTRUP00000044871	ENSTRUT00000045022	Ensembl
	1b	ENSTRUP00000024942	ENSTRUT00000025045	Ensembl
<i>Tetraodon nigroviridis</i>	2	ENSTNIP00000021755	ENSTNIT00000021990	Ensembl
	1a	ENSTNIP00000020660	ENSTNIT00000020893	Ensembl
	1b	ENSTNIP00000006942	ENSTNIT00000007096	Ensembl
<i>Thynnus thynnus</i>	2	AFQ98399.1	JQ585747.1	NCBI
<i>Xenopus laevis</i>	1a1	AAP57668.1	AY316296.1	NCBI
	1a2	AAP57669.1	AY316297.1	NCBI
<i>Xenopus tropicalis</i>	1	ENSXETG00000019534	ENSXETT00000042270	Ensembl
<i>Xiphophorus maculatus</i>	2	ENSXMAT00000000129	ENSXMAP0000000012	Ensembl
	1a	ENSXMA-P00000006030	ENSXMAT00000006036	Ensembl
	1b	ENSXMA-P00000008312	ENSXMAT00000008323	Ensembl

Only sequences used in the phylogenetic and synteny analysis were displayed. Empty cells correspond to sequences obtained in the lab which are under submission to databases. AANATs from genomic databases were manually assembled. The search of Actinopterygii AANATs meant to be exhaustive while we did not include all the publically available tetrapod AANATs. NV and VT stand for the non-vertebrate for and vertebrate AANAT subfamilies of Chondrichthyes and Cyclostomata [8]. Sarcopterygii AANATs were proven to be homologous of the Actinopterygii isoform 1 [4, 11] and were so labeled. Besides these exceptions, isoform description displayed here faithfully reflects the one provided in databases.

doi:10.1371/journal.pone.0112380.t001

Analysis of synteny

All *aanat* genes were located in a region presenting a noticeable degree of conservation among all of the investigated Actinopterygii. *aanat1* and *aanat2* contigs shared two paralogs: *rhb2* and *rhb1*, besides *cygb*, coding for a *cytoglobin*, and the globin cluster that framed *aanat1* and 2 (Fig. 3A-B, and S1 Data for precise gene locations). Four genes (*sphk1*, *ube2o*, *rhb2* and *cygb*) were close neighbors of *aanat1a* in Actinopterygii and Sarcopterygii. The Actinopterygii *aanat1b* region found no ortholog in Sarcopterygii.

No *aanat2* gene was found in Sarcopterygii, but eight orthologs of genes neighboring Actinopterygii *aanat2* were encountered (*snrnp25*, *polr3k*, *cdip1*, *mgnr1*, *rhb1*, *mpg*, *nprl3* and the globin cluster). The conservation of synteny is particularly strong in the spotted gar (*L. oculatus*) and the coelacanth (*L. chalumnae*), that lie at the base of the Actinopterygii and the Sarcopterygii lineages, respectively. In this region, and apart from the missing *aanat2* in the

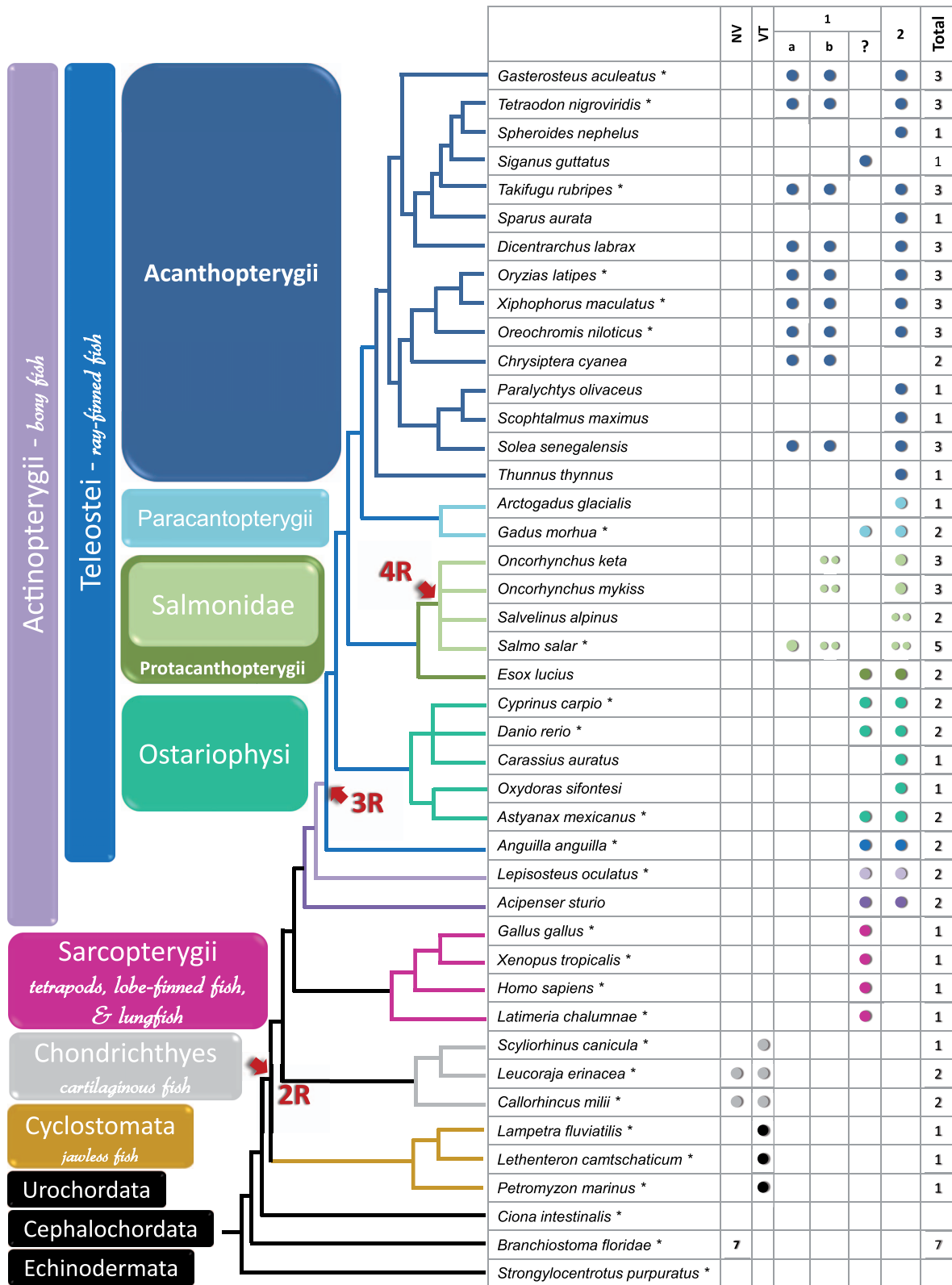


Figure 1. AANAT isoform distribution in vertebrates. AANATs isoforms were sorted according to the nomenclature displayed in [table 1](#). Taxa were classified according to their phylogenetic relationships defined in [\[65\]](#) and available genomes are indicated by an asterisk.

doi:10.1371/journal.pone.0112380.g001

coelacanth, six genes are conserved between the two species without any insertion/deletion (*cdip1*, *f100a*, *mgrn1*, *foxj1*, *rhbdf1*, *mpg*).

The vast majority of teleosts presented two *aanat1* genes, 1a and 1b, which were part of a paralogon constituted of three genes: *ddx5*, *myh10* and *evpla/evplb*. Only *G. morhua* lacked *aanat1a* in the region where all other paralogs of *aanat1b* were present. Conversely, *D. rerio* lacked the *aanat1b* on chromosome 3 where all six paralog genes were present. The same pattern was found in the recently released genome of the blind cave fish *Astyanax mexicanus*: *aanat1b* was absent in its putative location otherwise containing all six paralog genes (data not shown). To investigate whether teleosts also presented two AANAT2 forms, we looked for paralogs of *aanat2* surrounding genes. All teleost genomes (apart from salmonids which we discuss separately below) were characterized by a paralogon of the *aanat2* region composed of *shisa9*, *mkl2*, *f100a*, *rhbdf1* and the globin cluster where the paralog of *aanat2* was missing, and henceforth named *aanat2'*.

At the time we wrote this article, the *S. salar* available genome was only a primary assembly version with low coverage. We were thus unable to obtain *aanat* orthologs of the same length as the one displayed for other species. Nevertheless, we found contigs containing the three *aanat* paralogs (1a, 1b, and 2). Two isoforms were identified for the *aanat1b* and *aanat2* genes, whereas only one *aanat1a* gene was encountered. The two *aanat2* nucleotide sequences are almost identical displaying only two synonymous changes. AANAT1b peptide sequences differed by substitution of nine amino-acids a deletion of two amino-acids. Both *aanat1b* were located next to a *sphk* gene whose paralogs were found close to Sarcopterygii and Actinopterygii *aanat1a*.

Discussion

This study sheds new light on the evolution of AANAT, overturning preconceptions and providing a totally new vision about the diversification of AANATs in the different classes of vertebrates. We have shown previously that both NV- and VT-AANAT enzymes were present in the common ancestor of vertebrates, prior to the split of Cyclostomata and Gnathostomata (jawless and jawed vertebrates) [\[8\]](#). Rampant gene loss (which rapidly erases signals of genome duplication) and the sparsity of Cyclostomata and Chondrichthyes genomic data precluded investigations on genes neighboring NV- and VT-AANAT. Without a synteny approach, the question of the orthology of Cyclostomata and Chondrichthyes VT-AANAT cannot be solved. Our phylogenetic analysis alone did not reveal whether the VT-AANAT was more closely related to the Euteleostomi. We discuss below new evidence about the time of the AANAT1/AANAT2 and AANAT1a/1b duplication.

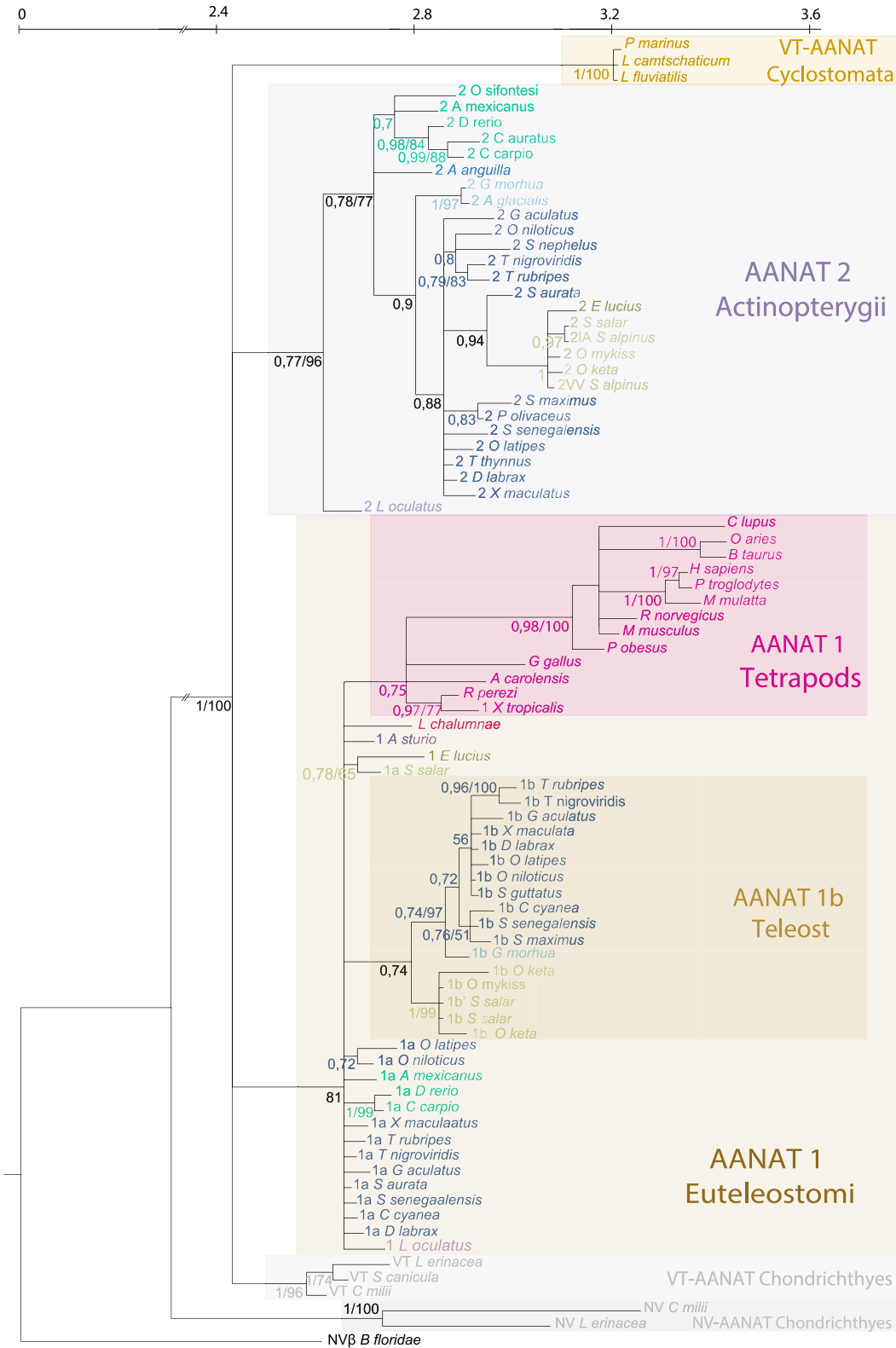


Figure 2. Phylogenetic tree of AANAT isoforms based on amino acid sequences. Bayesian and likelihood methods gave the same tree topology (only the Bayesian tree is shown). The amphioxus *Branchiostoma floridae* NV β -AANAT was used as outgroup. Numbers on branches indicate from left to right; (a) posterior probabilities in MrBayes reconstructions using a codon model and 4 million generations; (b) bootstrap support obtained in the maximum likelihood (ML) reconstruction (JTT+G) and 1,000 bootstrap replications. Note that posterior probabilities under 0.7 and Bootstrap values <50% were not displayed. Taxa color coding is similar to that used in [Figure 1](#).

doi:10.1371/journal.pone.0112380.g002

The emergence of two VT-AANAT isoforms

The occurrence of the two AANAT1 and AANAT2 forms is specific to all investigated Actinopterygii but is not reported in any Sarcopterygii, namely the Actinistia (coelacanth) and tetrapods ([Fig. 1](#)). Until now, the duplication of the ancestral AANAT to form the paralogous products *aanat1* and *aanat2* was assumed to have occurred at the base of the Actinopterygii lineage [[11](#)]. Our results contradict this view. Indeed, the present synteny analysis highlights the presence of an orthologon shared by all Euteleostomi (Actinopterygii and Sarcopterygii) in the region where only Sarcopterygii possesses the AANAT2. In other words many genes were conserved both upstream and downstream the region where the *aanat2* gene is now missing, suggesting that the most recent time at which the *aanat* duplication could have occurred would be in the common ancestor of Actinopterygii and Sarcopterygii, and that the absence of *aanat2* in the coelacanth and tetrapods results from a secondary loss. The loss of this gene probably did not involve deletion of a large genomic portion. The most parsimonious explanation is that the *aanat2* ortholog was lost from the genome by a unique small-scale event in the basal lineage of Sarcopterygii ([Fig. 4](#)). Similar patterns of secondary losses in Sarcopterygii, such as those in the opsin gene family, are considered to have played a crucial role in adaptive evolution of vertebrates to their light environment [[24](#)].

The emergence of *aanat1* and *aanat2* paralogs is difficult to time based on the current public data. This event is not the result of single gene duplication since Sarcopterygii regions surrounding *aanat1* and the 'absent' *aanat2* are characterized by another pair of paralogous genes (*rhbdf2* and *rhbdf1* respectively). It thus at least involved a regional duplication. Our analysis is restricted to the close neighborhood of *aanat* genes and does not allow us to conclude whether this duplication was part of the whole genome duplication (2R) reported at the basis of the vertebrates [[25](#), [26](#)] or if it rather occurred at the basis of Euteleostomi. We favor the first hypothesis as it involves the smallest number of evolutionary events (duplications and losses), while an independent duplication would lead to a less parsimonious scenario.

AANAT1 and AANAT2 diversification and losses in teleosts

The AANAT1a and AANAT1b forms have only been reported in teleosts. These two paralogs are framed by paralogs in the eight teleost genomes investigated, suggesting that the duplication event leading to these two forms is at least regional. The teleost specific WGD (the 3R or TGD, Teleost Genome Duplication

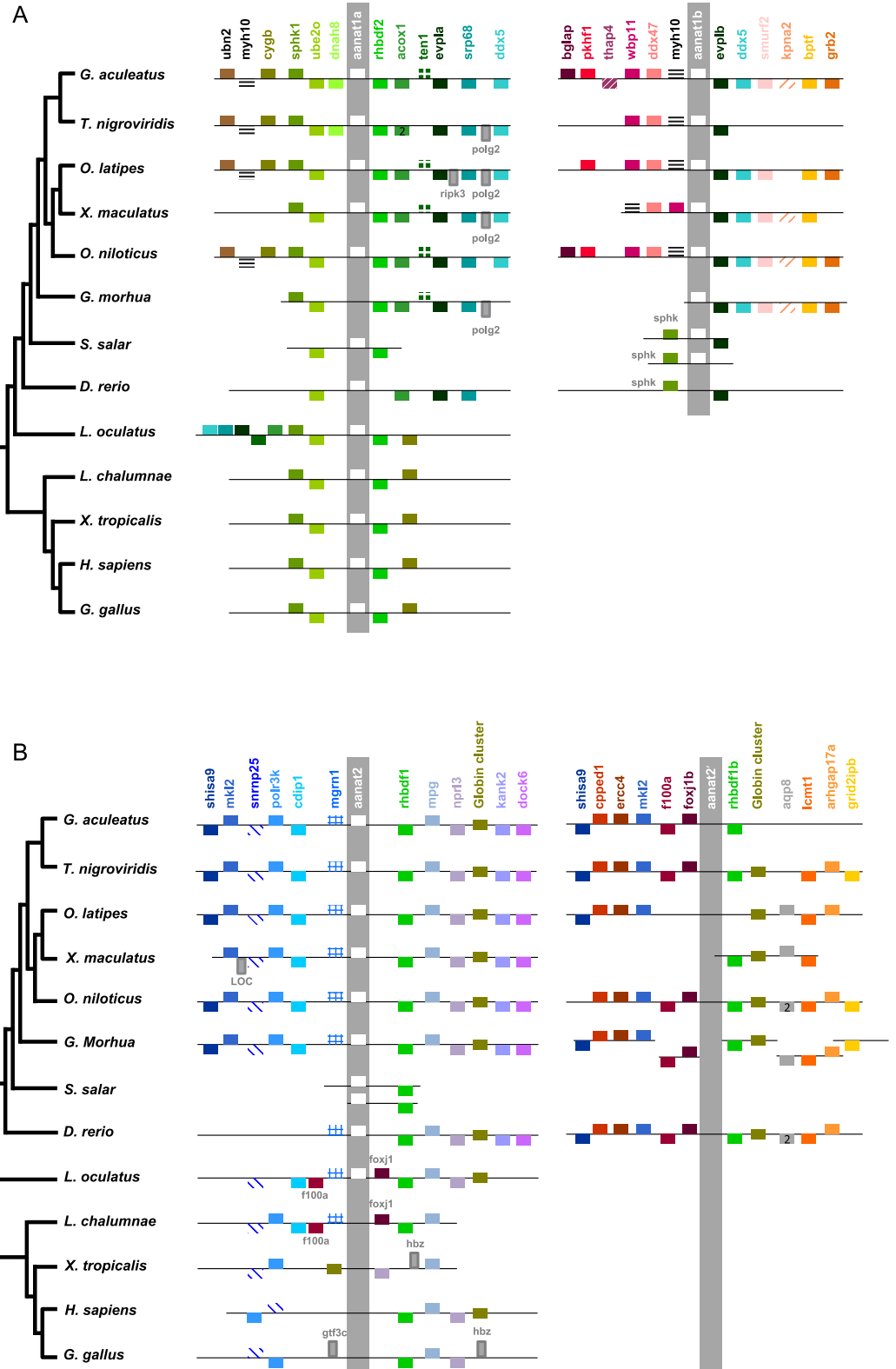


Figure 3. Patterns of synteny in genes neighboring *aanat1* and *aanat2*. Unscaled representations of conserved synteny of *aanat* genes: *aanat1* (A) and *aanat2* (B) paralogs are represented in parallel. The stickleback genome (*Gasterosteus aculeatus*) was used as a reference, and all *aanat* genes are presented in this same orientation. Six genes found upstream and downstream of each *aanat* paralog were selected in the stickleback and sought in all species of interest. Each box represents a gene placed on its chromosome portion. A number two in the box represents rare cases where two consecutive identical genes are found. Genes in the forward orientation are shown on top of the chromosome, whereas genes in the reverse orientation are shown below. The globin cluster encompasses several genes in various orientations, for details see [66]. Precise gene location and name is provided in the [S1 Data](#) spreadsheet.

doi:10.1371/journal.pone.0112380.g003

[27–32]) may be at the origin of the *aanat1* duplication into the *1a* and *1b* paralogs, and remains the most parsimonious scenario. This hypothesis tacitly implies that *aanat2* was duplicated in the same event. Yet no *aanat2* paralogs were found when screening the available teleost genomes. However paralogs were found for four of the genes framing *aanat2*, constituting a paralogon of the *aanat2* region (Fig. 3B). This conservation of synteny observed in all investigated teleosts suggests a secondary loss of *aanat2'*, the *aanat2* paralog in their common ancestor. These eight species belong to a wide diversity of taxa (Acantopterygii, Paracantopterygii, Protacanthopterygii, and Ostariophysii), suggesting that the secondary loss of *aanat2'* occurred shortly after the TGD, in the common ancestor of those taxa, *i.e.*, before the diversification of teleosts. Our findings change the prevailing scenario [11] that was built using fewer *aanat1a* and *1b* available

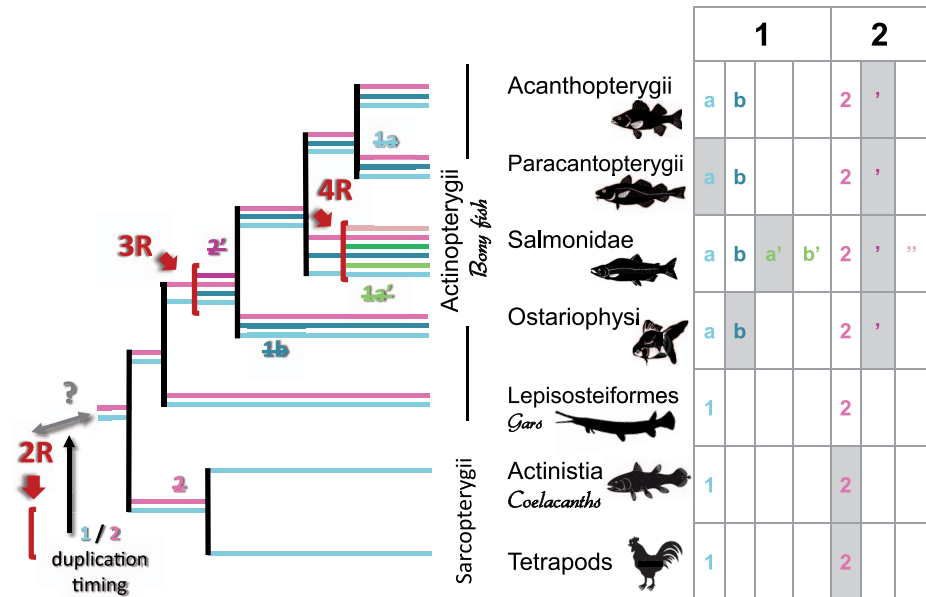


Figure 4. Evolutionary scenario proposing the potential role of whole genome duplications in AANAT diversification. Whole genome duplications are indicated by a red arrow: they concern vertebrates (2R), teleosts (3R) and salmonids (4R). Some *aanat* duplications were followed by secondary losses and are represented by shaded cells in the table or dotted branches on the tree. Other *aanat* paralogs diverged to form various isoforms as shown in different colors when present in a given phylum. The exact timing of the *aanat1/2* duplication cannot be inferred from current genomic data and ranges between the 2R and the Actinopterygii/Sarcopterygii split.

doi:10.1371/journal.pone.0112380.g004

sequences. Coon and Klein (2006) assumed that the duplication occurred in the main trunk of the teleost lineage after the zebrafish branch had split off, and did not consider the secondary loss of a putative second AANAT2 form.

Up to now, the evolutionary history of teleosts AANAT1a and AANAT1b has not led to any straightforward conclusion given that their pattern of presence/absence does not seem to be clade specific. Moreover, the investigation effort exerted to characterize AANAT enzymes has not been the same in all species and for all isoforms, and the well-known aphorism prevails here: “absence of evidence is not evidence of absence”. Our study brings some clarification.

First, the phylogeny analysis allows the AANAT1 nomenclature to be revised, especially for species where only one single AANAT1 form is known. Based on its position in the AANAT phylogeny, the orange-spotted spinefoot *Siganus guttatus* AANAT1 can be identified as a 1a form. All known Protacanthopterygii AANAT1 may all be the 1b form: the pike *E. lucius*, the rainbow trout *O. mykiss*, and the two undefined AANAT1 forms of chum salmon *O. keta* belong to the well supported 1b clade. But the phylogeny alone does not solve the question of AANAT1 nomenclature in the blind cave fish (*Astyanax mexicanus*), the common carp (*Cyprinus carpio*) and the zebrafish (*Danio rerio*) for which sequences fall in an intermediate position between 1a and 1b (Fig. 2). The synteny analysis allows this question to be circumvented, since this *aanat1* is located in the ortholog region of all other teleosts *aanat1a*. In addition, *aanat1b* is clearly absent in these three species whose genomes are available. Moreover, all genes surrounding *aanat1b* in other teleosts present a conservation of synteny, demonstrating the secondary loss of *aanat1b* in the blind cave fish *A. mexicanus*, the common carp *C. carpio* and the zebrafish *D. rerio*. Further investigations would be needed to generalize this statement for all Ostariophysi, but we hypothesize here that this event may have happened in the common ancestor of this clade (Fig. 4). Other genomic approaches on HOX genes have already shown that *D. rerio* presents a unique pattern of gene duplication/loss compared to other teleosts [33].

Only one *aanat1* was detected in the Protacanthopterygii *G. morhua* genome. This Atlantic cod form is closely related to other teleost *aanat1b* (Fig. 2). Its neighboring genes correspond to the 1b paralogon, while the 1a is missing in the region where its ortholog is found in other teleosts (Fig. 3A). This information allows us to conclude that there was a subsequent loss of *aanat1a* in the Atlantic cod *G. morhua* (Fig. 4). The pattern of presence/absence of AANAT1a/AANAT1b in Ostariophysi and Paracanthopterygii results from the reciprocal gene losses (RGL) previously defined [34]. The two paralogs were created by WGD (in this case the TGD) and were retained until speciation, after which each taxon lost a different copy.

A round of duplications specific to salmonids

The Atlantic salmon *S. salar* presents a unique diversity of AANATs with more than the three typical *aanat1a*, *aanat1b* and *aanat2* forms classically encountered in other teleosts. Its genome contains one form of *aanat1a*, but two forms for

both, *aanat1b* and *aanat2*. Although the salmon genome is still poor and does not allow synteny analysis on long fragments, we are confident each duplicate of the *aanat1b* and *aanat2* are located on two distinct genomic regions and truly correspond to two different gene copies ([S1 Data](#)). First, each belongs to a distinct contig, also containing respectively the *sphk* or *rhbdf1* genes. Second, these *aanat1b* or *aanat2* paralogs present more genetic divergence than sequencing errors alone could explain. But they are much less divergent than two paralogs generated by the TGD: *rhbdf1/1b* located on a given teleost *aanat2/2'* paralogon show 75 to 80% identity, while the two salmon *rhbdf1* neighboring these two *aanat2* copies show 94% identity.

Other multiple AANAT isoforms were previously reported in related salmonid taxa. Two AANAT1b were observed in the chum salmon *O. keta* [35] and two AANAT2 were characterized in the Arctic char *S. alpinus* [36, 37]. Unlike in *S. alpinus* where the two variants differ by two amino-acids, the *S. salar* AANAT2 variants only differ by two nucleotides and are identical at the peptide level.

A specific WGD [38, 39] which occurred in the Salmonidae family at least 88 Mya [40] could explain this particular AANAT diversity. We hypothesize the three *aanat* genes - resulting from two previous whole genome duplication rounds (2R and the TGD 3R) and a subsequent loss of the *aanat2'* - could have been duplicated during this event. We propose that the second copy of the *aanat1a* has been lost. Other mechanisms such as independent regional duplications could also explain this pattern. Given this last scenario is less parsimonious we favor the one involving the well documented salmonid WGD [38, 39].

Functional implications of AANAT diversification

We suggest that *aanat* genes present in vertebrates went through a series of three consecutive duplications. Based on available evidence we propose these duplications are the result of the Vertebrata, teleosts and salmonid-specific WGDs. Several *aanat* duplicates were lost, possibly after an accumulation of deleterious mutations or chromosome rearrangements. This is mainly the case for the most ancient duplicates such as the NV- and VT-*aanat* in early vertebrates. Secondary losses are suspected to be the most frequent scenario following gene duplication, because only a minority of gene pairs adopts a new function (neofunctionalization) or partitions old functions (subfunctionalization) quickly enough to persist in face of deleterious mutations [26, 41, 42].

Falcon and colleagues [8] demonstrated that after an ancestral NV-*aanat* gene duplication in early vertebrates, one copy acquired the new function of melatonin production while the other copy maintained its detoxification role and was eventually lost in most Vertebrate lineages. This scenario obviously corresponds to a neofunctionalization event. Based on tissue or organ localization, enzyme activity and substrate affinity, we see that that all VT-*aanat* duplications leading to the AANAT1/AANAT2 forms, and subsequently the AANAT1a/AANAT1b forms, did not generate new functions but rather led to subfunctionalization or functional refinement or shuffling. These subfunctions are more diverged in the

AANAT1 and AANAT2 forms that emerged earlier in the evolutionary history of the Timezyme. On the contrary, finer specializations are seen between the AANAT1a and AANAT1b isoforms, whose duplication and differentiation is more recent.

The acquisition of *aanat1/aanat2* paralogs occurred before the divergence of Sarcopterygii and Actinopterygii and we clearly deciphered the two distinct evolutionary histories followed in each lineage. Both genes were maintained in Actinopterygii, possibly corresponding to the complementary partitioning of ancestral structural or regulatory functions between *aanat1* and *aanat2*. Several clues favor this hypothesis. A first one is the selective expression of the paralogs; *aanat1* is expressed in the retina and brain, while *aanat2* is expressed in the pineal gland [43–45]. A second one is the partitioning and refinement of the kinetics properties of these enzymes. AANAT1 acetylates indoleethylamines and phenylethylamines equally well, whereas AANAT2 exhibits strong preference for indoleethylamines [43, 46–48]. A third one is the difference in regulation by temperature both at the cellular [49, 50] – and molecular levels [43, 46]. The subfunctionalization of AANAT1 and 2 may have contributed to the functional specialization of the organs where they are respectively expressed. In the retina, acetylation of both *N*-acetylserotonin (indoleethylamine) and dopamine (phenylethylamine) contributes to regulating the circadian rhythmicity displayed by the tissue, which involves the reciprocal melatonin/dopamine interaction [51–53]. The pineal gland became a melatonin factory that allows integrating the photoperiod and temperature information in a species-dependent manner. This pattern typically corresponds to the duplication, degeneration, complementation (or DDC) model where the addition of both paralogs' functions may provide at least that of the pre-duplicated gene [26].

On the contrary, *aanat2* was lost in Sarcopterygii and only the *aanat1* paralog remained. This is the most frequently observed scenario after duplication [26, 41]. Even if gene redundancy may be maintained for a certain period, nonfunctionalization (or pseudogenization) by accumulation of detrimental mutations is the most common fate of one of the duplicates. However, this loss of *aanat2* does not mean a loss of function. The remaining Sarcopterygii AANAT gathers functions observed for both Actinopterygii AANAT1 and AANAT2. Phylogeny and synteny demonstrate that this gene is the ortholog of Actinopterygii *aanat1*. Yet, Sarcopterygii expression sites resemble more Actinopterygii AANAT2. Except for frog AANAT, which is equally expressed in the retina, the pineal photoreceptors and in some discrete brain areas [54], chicken and mammalian AANAT display high levels of expression in the pineal gland and much lower levels in the retina [6, 55]. The kinetics of the avian and mammalian AANAT appear to be closer to teleost AANAT2 rather than teleost AANAT1 in terms of substrate selectivity, while displaying an even higher catalytic efficiency [56–61]. *Aanat* expression and AANAT enzyme activity are both higher at night than during day, as is always the case for teleost AANAT2 but not necessarily for teleost AANAT1. To summarize, Sarcopterygii AANAT is the ortholog of retinal AANAT1 from Actinopterygii but displays many characteristic of the Actinopterygii AANAT2 pineal enzyme, which

controls production of circulating melatonin. The two orthologs (*i.e.* Sarcopterygii AANAT and Actinopterygii AANAT2) either display a convergent evolution towards the endocrine production of melatonin or retained the ancestral features of the AANAT from which they diverged. Finer characterization of early vertebrate AANAT should help decipher the functional features of the ancestral VT-AANAT. In any case the present pattern highlights the crucial role of melatonin in vertebrate biological rhythms.

The physiological refinement produced by the split of AANAT1 into the 1a and 1b forms is more tenuous. While some species possess both genes, others possess either one or the other, as a result of gene secondary losses. When both genes are present *aanat1b* shows higher levels of expression. However this observation concerns only two species, the sole *Solea senegalensis* [62] and the sea bass *Dicentrarchus labrax* (Besseau et al, unpublished data). While we observe a marked specialization in expression sites for AANAT1s (retina/brain) vs. AANAT2 (pineal) the subfunctionalization of AANAT1a/1b is only observed at the cell type level. It concerns different cell layers of the retina (Besseau et al, unpublished data) or a different timing in expression across larval development [62]. Information concerning the kinetics and tissue distributions of the two isoforms is lacking. It is possible that expression of either form is related to some specific life history traits of the species considered.

We show that *aanat1b* has been lost by three Ostariophysi (*D. rerio*, *C. carpio*, and *A. mexicanus*). Based on our phylogeny, the primary structure of these AANAT1a is an intermediate between AANAT1a and 1b orthologs from other teleosts. This would provide a case of functional shuffling where surviving paralogs re-acquire subfunctions related to loss of ohnologs (OGMs, Ohnologs Gone Missing) [26]. Further functional assays would be suitable to test the hypothesis that these Ostariophysi AANAT1a enzymes have re-acquired subfunctions lost with the 1b paralog.

In summary vertebrates provide a panel of evolutionary strategies with regard to the synthesis of melatonin, wherein the Timezyme plays a pivotal role. As demonstrated earlier, VT-AANAT results from a duplication followed by neofunctionalization (acetylation of indolamines) of one paralog and disappearance of the other [8]. This new AANAT displayed a more specific tissue distribution than its paralog (as seen when both are present, *i.e.*, in Chondrichthyes) [8]. This specialization may be at the origin of the nocturnal melatonin secretion signal. From there, and after a new duplication event, two scenarios are observed. The lack of public information and the incomplete nature of released genomes of the Cyclostomata and Chondrichthyes preclude providing a definitive picture based on synteny. However, our data support the idea that the 2R generated AANAT1 and AANAT2 and that the two paralogs were conserved only in Actinopterygii, with clear subfunctionalization and tissue specific expression. These factors might have been crucial in the conservation of AANAT2 by Actinopterygii. From Cyclostomata to mammals AANAT expression became concentrated in the pineal gland and catalytic efficiency increased. Thus, whatever the strategy, they all converged to make pineal melatonin production the signal

for the night from Cyclostomata to mammals. In other words, the time-keeping nocturnal signal is preserved. This common outcome observed at the molecular level has been accompanied by spectacular modifications of the pineal anatomy and structure as well as of the mechanisms of regulation by light, as extensively reported elsewhere [10, 63, 64], emphasizing the importance of the melatonin message in vertebrates.

Supporting Information

S1 Data. Detailed information about genes involved in the synteny analysis. Each of the four tables displays precise name, accession number and Ensembl location of genes involved in the synteny analysis for *aanat1a*, *aanat1b*, *aanat2*, and *aanat2'*. Six genes found upstream and downstream of each *aanat* paralog were annotated, using the stickleback genome (*Gasterosteus aculeatus*) as a reference. Cells highlighted in yellow provide the contig number on which were found the targeted genes in the AGKD01 salmon genome, since *S. salar* is not referenced in Ensembl.

[doi:10.1371/journal.pone.0112380](https://doi.org/10.1371/journal.pone.0112380) (XLSX)

Acknowledgments

We thank Sylvie Dufour and Jeremy Pasquier (UMR BOREA, CNRS-MNHN): We drew on inspiration from the enthusiastic discussions we had together. We also thank Nigel Grimsley (UMR BIOM CNRS-UPMC) for his thorough editing.

Author Contributions

Conceived and designed the experiments: EM DCC. Performed the experiments: DCC EM. Analyzed the data: DCC EM. Contributed reagents/materials/analysis tools: JF. Wrote the paper: EM DCC JF LB.

References

1. Falcon J (1999) Cellular circadian clocks in the pineal. *Prog Neurobiol* 58: 121–162.
2. Reiter RJ (1991) Melatonin synthesis: multiplicity of regulations. *Adv Exp Med Biol* 294: 149–158.
3. Falcon J, Besseau L, Sauzet S, Boeuf G (2007) Melatonin effects on the hypothalamo-pituitary axis in fish. *Trends Endocrinol Metab* 18: 81–88.
4. Falcon J, Migaud H, Munoz-Cueto JA, Carrillo M (2010) Current knowledge on the melatonin system in teleost fish. *Gen Comp Endocrinol* 165: 469–482.
5. Iuvone PM, Tosini G, Pozdeyev N, Haque R, Klein DC, et al. (2005) Circadian clocks, clock networks, arylalkylamine N-acetyltransferase, and melatonin in the retina. *Prog Retin Eye Res* 24: 433–456.
6. Klein DC, Coon SL, Roseboom PH, Weller JL, Bernard M, et al. (1997) The melatonin rhythm-generating enzyme: molecular regulation of serotonin N-acetyltransferase in the pineal gland. *Recent Prog Horm Res* 52: 307–357; discussion 357–308.
7. Klein DC (2007) Arylalkylamine N-acetyltransferase: “the Timezyme”. *J Biol Chem* 282: 4233–4237.

8. **Falcon J, Coon SL, Besseau L, Cazamea-Catalan D, Fuentes M, et al.** (2014) Drastic neofunctionalization associated with evolution of the timezyme AANAT 500 Mya. *Proc Natl Acad Sci U S A* 111: 314–319.
9. **Klein DC** (2006) Evolution of the vertebrate pineal gland: the AANAT hypothesis. *Chronobiol Int* 23: 5–20.
10. **Collin JP, Arendt J, Gern WA** (1988) The third eye. *La Recherche* 19: 1154–1165.
11. **Coon SL, Klein DC** (2006) Evolution of arylalkylamine N-acetyltransferase: emergence and divergence. *Mol Cell Endocrinol* 252: 2–10.
12. **Falcon J, Besseau L, Sauzet S, Fuentes M, Boeuf G** (2007) Melatonin and neuroendocrine regulations in fish. *J Soc Biol* 201: 21–29.
13. **Louis A, Muffato M, Roest Crollius H** (2013) Genomicus: five genome browsers for comparative genomics in eukaryota. *Nucleic Acids Res* 41: D700–705.
14. **Burge C, Karlin S** (1997) Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268: 78–94.
15. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
16. **Davidson WS, Koop BF, Jones SJ, Iturra P, Vidal R, et al.** (2010) Sequencing the genome of the Atlantic salmon (*Salmo salar*). *Genome Biol* 11: 403.
17. **Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, et al.** (2010) Geneious v5.3. Available: <http://www.geneious.com>.
18. **Thompson JD, Higgins DG, Gibson TJ** (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673–4680.
19. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al.** (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Likelihood, Distance, and Parsimony methods. *Mol Biol Evol* 28: 2731–2739.
20. **Castresana J** (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17: 540–552.
21. **Abascal F, Zardoya R, Posada D** (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104–2105.
22. **Guindon S, Gascuel O** (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
23. **Ronquist F, Huelsenbeck JP** (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
24. **Davies WL** (2009) Adaptive gene loss in vertebrates: photosensitivity as a model case. *Encyclopedia of Life Sciences*: DOI: 10.1002/9780470015902.a9780470022890.
25. **Kuraku S, Meyer A, Kuratani S** (2009) Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol Biol Evol* 26: 47–59.
26. **Canestro C, Albalat R, Irimia M, Garcia-Fernandez J** (2013) Impact of gene gains, losses and duplication modes on the origin and diversification of vertebrates. *Semin Cell Dev Biol* 24: 83–94.
27. **Christoffels A, Koh EG, Chia JM, Brenner S, Aparicio S, et al.** (2004) Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Mol Biol Evol* 21: 1146–1151.
28. **Froschauer A, Braasch I, Voff JN** (2006) Fish genomes, comparative genomics and vertebrate evolution. *Current Genomics* 7: 43–57.
29. **Hoegg S, Brinkmann H, Taylor JS, Meyer A** (2004) Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J Mol Evol* 59: 190–203.
30. **Lundin LG** (1993) Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* 16: 1–19.
31. **Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, et al.** (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431: 946–957.

32. **Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, et al.** (2005) The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res* 15: 1307–1314.
33. **Hoegg S, Meyer A** (2005) Hox clusters as models for vertebrate genome evolution. *Trends Genet* 21: 421–424.
34. **Semon M, Wolfe KH** (2007) Reciprocal gene loss between tetraodon and zebrafish after whole genome duplication in their ancestor. *Trends Genet* 23: 108–112.
35. **Shi Q, Ando H, Coon SL, Sato S, Ban M, et al.** (2004) Embryonic and post-embryonic expression of arylalkylamine N-acetyltransferase and melatonin receptor genes in the eye and brain of chum salmon (*Oncorhynchus keta*). *Gen Comp Endocrinol* 136: 311–321.
36. **Cazamea-Catalan D, Magnanou E, Helland R, Besseau L, Boeuf G, et al.** (2013) Unique arylalkylamine N-acetyltransferase-2 polymorphism in salmonids and profound variations in thermal stability and catalytic efficiency conferred by two residues. *J Exp Biol* 216: 1938–1948.
37. **Cazamea-Catalan D, Magnanou E, Helland R, Vanegas G, Besseau L, et al.** (2012) Functional diversity of Teleost arylalkylamine N-acetyltransferase-2: is the timezyme evolution driven by habitat temperature? *Mol Ecol* 21: 5027–5041.
38. **Allendorf FW, Thorgaard GH** (1984) Tetraploidy and the evolution of Salmonid fishes. In: (ed.) BJT, editor. *Evolutionary genetics of fishes*: Plenum Publishing Corporation, New York. pp.1–53.
39. **Koop BF, Davidson WS.** Genomics and the genome duplication in salmonids; 2008. TERRAPUB. pp. 77–86.
40. **Macqueen DJ, Johnston IA** (2014) A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc R Soc B* 281: 20132881.
41. **Dehal P, Boore JL** (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol* 3: e314.
42. **Kasahara M** (2007) The 2R hypothesis: an update. *Curr Opin Immunol* 19: 547–552.
43. **Coon SL, Begay V, Deurloo D, Falcon J, Klein DC** (1999) Two arylalkylamine N-acetyltransferase genes mediate melatonin synthesis in fish. *J Biol Chem* 274: 9076–9082.
44. **Gothilf Y, Coon SL, Toyama R, Chitnis A, Namboodiri MA, et al.** (1999) Zebrafish serotonin N-acetyltransferase-2: marker for development of pineal photoreceptors and circadian clock function. *Endocrinology* 140: 4895–4903.
45. **Begay V, Falcon J, Cahill GM, Klein DC, Coon SL** (1998) Transcripts encoding two melatonin synthesis enzymes in the teleost pineal organ: circadian regulation in pike and zebrafish, but not in trout. *Endocrinology* 139: 905–912.
46. **Zilberman-Peled B, Benhar I, Coon SL, Ron B, Gothilf Y** (2004) Duality of serotonin-N-acetyltransferase in the gilthead seabream (*Sparus aurata*): molecular cloning and characterization of recombinant enzymes. *Gen Comp Endocrinol* 138: 139–147.
47. **Zilberman-Peled B, Bransburg-Zabary S, Klein DC, Gothilf Y** (2011) Molecular evolution of multiple arylalkylamine N-acetyltransferase (AANAT) in fish. *Mar Drugs* 9: 906–921.
48. **Zilberman-Peled B, Ron B, Gross A, Finberg JP, Gothilf Y** (2006) A possible new role for fish retinal serotonin-N-acetyltransferase-1 (AANAT1): Dopamine metabolism. *Brain Res* 1073–1074: 220–228.
49. **Falcon J, Bolliet V, Collin JP** (1996) Partial characterization of serotonin N - acetyltransferases from northern pike (*Esox lucius*, L.) pineal organ and retina: effects of temperature. *Pflugers Arch* 432: 386–393.
50. **Benyassi A, Schwartz C, Coon SL, Klein DC, Falcon J** (2000) Melatonin synthesis: arylalkylamine N-acetyltransferases in trout retina and pineal organ are different. *Neuroreport* 11: 255–258.
51. **Bartell PA, Miranda-Anaya M, Mclvor W, Menaker M** (2007) Interactions between dopamine and melatonin organize circadian rhythmicity in the retina of the green iguana. *J Biol Rhythms* 22: 515–523.
52. **Ribelayga C, Wang Y, Mangel SC** (2004) A circadian clock in the fish retina regulates dopamine release via activation of melatonin receptors. *J Physiol* 554: 467–482.
53. **Tosini G, Baba K, Hwang CK, Iuvone PM** (2012) Melatonin: an underappreciated player in retinal physiology and pathophysiology. *Exp Eye Res* 103: 82–89.

54. **Isorna E, Besseau L, Boeuf G, Desdevises Y, Vuilleumier R, et al.** (2006) Retinal, pineal and diencephalic expression of frog arylalkylamine N-acetyltransferase-1. *Mol Cell Endocrinol* 252: 11–18.
55. **Iuvone PM, Brown AD, Haque R, Weller J, Zawilska JB, et al.** (2002) Retinal melatonin production: role of proteasomal proteolysis in circadian and photic control of arylalkylamine N-acetyltransferase. *Invest Ophthalmol Vis Sci* 43: 564–572.
56. **Coon SL, Roseboom PH, Baler R, Weller JL, Namboodiri MA, et al.** (1995) Pineal serotonin N-acetyltransferase: expression cloning and molecular analysis. *Science* 270: 1681–1683.
57. **Coon SL, Weller JL, Korf HW, Namboodiri MA, Rollag M, et al.** (2001) cAmp regulation of arylalkylamine N-acetyltransferase (AANAT, EC 2.3.1.87): a new cell line (1E7) provides evidence of intracellular AANAT activation. *J Biol Chem* 276: 24097–24107.
58. **De Angelis J, Gastel J, Klein DC, Cole PA** (1998) Kinetic analysis of the catalytic mechanism of serotonin N-acetyltransferase (EC 2.3.1.87). *J Biol Chem* 273: 3045–3050.
59. **Ferry G, Loynel A, Kucharczyk N, Bertin S, Rodriguez M, et al.** (2000) Substrate specificity and inhibition studies of human serotonin N-acetyltransferase. *J Biol Chem* 275: 8794–8805.
60. **Ferry G, Ubeaud C, Daully C, Mozo J, Guillard S, et al.** (2004) Purification of the recombinant human serotonin N-acetyltransferase (EC 2.3.1.87): further characterization of and comparison with AANAT from other species. *Protein Expr Purif* 38: 84–98.
61. **Voisin P, Namboodiri MA, Klein DC** (1984) Arylamine N-acetyltransferase and arylalkylamine N-acetyltransferase in the mammalian pineal gland. *J Biol Chem* 259: 10913–10918.
62. **Isorna E, Aliaga-Guerrero M, M'Rabet AE, Servili A, Falcon J, et al.** (2011) Identification of two arylalkylamine N-acetyltransferase 1 genes with different developmental expression profiles in the flatfish *Solea senegalensis*. *J Pineal Res* 51: 434–444.
63. **Ekstrom P, Meissl H** (2003) Evolution of photosensory pineal organs in new light: the fate of neuroendocrine photoreceptors. *Philos Trans R Soc Lond B Biol Sci* 358: 1679–1700.
64. **Falcon J, Besseau L, Fuentes M, Sauzet S, Magnanou E, et al.** (2009) Structural and functional evolution of the pineal melatonin system in vertebrates. *Ann N Y Acad Sci* 1163: 101–111.
65. **Betancur-R R, Broughton RE, Wiley EO, Carpenter K, López JA, et al.** (2013) The tree of life and a new classification of bony fishes PLOS Currents Tree of Life Edition 1. doi: 10.1371/currents.tol.53ba26640df0ccae75bb165c8c26288.
66. **Opazo JC, Butts GT, Nery MF, Storz JF, Hoffmann FG** (2013) Whole-genome duplication and the functional diversification of teleost fish hemoglobins. *Mol Biol Evol* 30: 140–153.