Biochimica et Biophysica Acta 1861 (2017) 644–651

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Genomic organization and evolution of olfactory receptors and trace amine-associated receptors in channel catfish, *Ictalurus punctatus*

Sen Gao [a,1], Shikai Liu [a,1], Jun Yao [a], Ning Li [a], Zihao Yuan [a], Tao Zhou [a], Qi Li [b], Zhanjiang Liu [a,*]

[a] The Fish Molecular Genetics and Biotechnology Laboratory, Aquatic Genomics Unit, School of Fisheries, Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL 36849, USA
[b] Key Laboratory of Mariculture Ministry of Education, Ocean University of China, Qingdao, China

ARTICLE INFO

Article history:
Received 8 July 2016
Received in revised form 5 October 2016
Accepted 19 October 2016
Available online 21 October 2016

Keywords:
Fish
Olfaction
Olfactory receptor
Trace amine-associated receptor

ABSTRACT

Background: Channel catfish (*Ictalurus punctatus*) live in turbid waters with limited visibility to chase prey within a certain distance. This can be compensated through detecting specific water-soluble substances by the olfactory receptors (ORs) and trace amine associated receptors (TAARs) expressed on the olfactory epithelium.
Methods: We identified the OR and TAAR repertoires in channel catfish, and characterized the genomic organizations of these two gene families by data mining available genomic resources.
Results: A total of 47 putative OR genes and 36 putative TAAR genes were identified in the channel catfish genome, including 27 functional OR genes and 28 functional TAAR genes. Phylogenetic and orthogroup analyses were conducted to illustrate the evolutionary dynamics of the vertebrate ORs and TAARs. Collinear analysis revealed the presence of two conserved orthologous blocks that contain OR genes between the catfish genome and zebrafish genome. The complete loss of a conserved motif in fish OR family H may contribute to the divergence of family H from other families. The dN/dS analysis indicated that the highest degree of selection pressure was imposed on TAAR subfamily 14 among all fish ORs and TAARs.
Conclusions: The present study provides understanding of the evolutionary dynamics of the two gene families (OR and TAAR) associated with olfaction in channel catfish.
General significance: This is the first systematic study of ORs and TAARs in catfish, which could provide valuable genomic resources for further investigation of olfactory mechanisms in teleost fish.

1. Introduction

Carnivorous fish use one of their chemosensory systems, the olfaction system, to detect and discriminate a broad spectrum of water-soluble substances. This system is mainly mediated by the olfactory receptors (ORs), a group of seven-transmembrane G protein-coupled receptors (GPCRs) in the class of rhodopsin. These receptors are expressed on the olfactory neurons, and can induce the signal transduction pathways that trigger behaviors [1–3]. When fish chase the preys in turbid waters, ORs are used as a compensation for limited visibility. Accordingly, preys have co-evolved to induce the OR-related aversion activities of predators. For instance, sea hares can release inks, which are composed of amino acids, to stimulate sea catfish to avoid predatory attacks [4]. In teleosts, ORs are used to mediate reproduction activities via sensation of odorants such as nucleotides, polyamines and bile salts [5,6]. ORs also have other roles, including self-expression regulation and auxiliary connection of sensory neurons [7].

In addition to ORs, trace amine-associated receptors (TAARs) are expressed on the olfactory epithelium. Although TAARs were initially considered as neurotransmitter receptors [8], several studies indicated that they have similar functions to ORs [9–16]. There are three clades in vertebrate TAARs: Clade I TAARs are found in both mammals and fish species, and they are expressed to detect primary amines through an aspartic acid on the third transmembrane domain ($Asp^{3.32}$; Ballesteros-Weinstein indexing); Clade II TAARs are found only in mammals, and they are expressed to detect tertiary amines; Clade III TAARs are found only in fish species, and they use $Asp^{5.42}$ instead of $Asp^{3.32}$ to detect amines [17,18]. However, some members of the fish TAAR subfamilies 13 and 14, which are not classified into any clades mentioned above, possess both $Asp^{3.32}$ and $Asp^{5.42}$ for detecting diamines [18].

The numbers of ORs and TAARs vary between mammals and fish species. Larger OR repertoires exist in mammalian species than in fish species. For instance, approximately 700 ORs have been identified in the human genome, and approximately 1200 ORs have been identified in the genomes of rodents [19,20]. The numbers of ORs in fish species are much smaller with zebrafish possessing the most ORs of 140 [21]. However, fish species possess more TAARs than mammals [22,23]. Moreover, the number of TAARs is far smaller than that of ORs in

* Corresponding author.
E-mail address: liuzhan@auburn.edu (Z. Liu).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.bbagen.2016.10.017
0304-4165/

mammals, while the numbers of these two gene families are roughly equal in fish species. It is speculated that the number variation of OR/TAAR genes between mammals and fish species might mirror the evolutionary dynamics of this gene family.

In catfish, previous efforts were mainly focused on the physiological and neural studies of the ORs [24–28], while few studies have been conducted on TAARs. Here, upon the completion of the catfish reference genome assembly [29], this study first report the complete repertoires of ORs and TAARs and their organizations in the channel catfish genome, and provide insights into the evolutionary dynamics of these two gene families in vertebrates.

2. Materials and methods

2.1. Retrieval of ORs and TAARs of other vertebrates

The OR and TAAR repertoires from 17 vertebrates, including amazon molly (*Poecilia formosa*), cave fish (*Astyanax mexicanus*), cod (*Gadus morhua*), fugu (*Takifugu rubripes*), medaka (*Oryzias Latipes*), platyfish (*Xiphophorus maculatus*), spotted gar (*Lepisosteus oculatus*), stickleback (*Gasterosteus aculeatus*), tetraodon (*Tetraodon nigroviridis*), tilapia (*Oreochromis niloticus*), zebrafish (*Danio rerio*), elephant shark (*Callorhinchus milii*), anole lizard (*Anolis carolinensis*), chicken (*Gallus gallus*), mouse (*Mus musculus*), cow (*Bos taurus*) and human (*Homo sapiens*), were retrieved from NCBI and ENSEMBL. In order to identify all the well-annotated OR and TAAR genes in the aforementioned species without mistakes, we took careful strategies as revealed below: sequences labeled with “pseudogene” and sequences labeled without specific subfamily name were removed; only full length OR and TAAR protein sequences were included in our analysis; only the longest sequences were selected when genes have multiple isoforms. For the ENSEMBL datasets, sequence descriptions were also downloaded using BioMart, and were combined with protein sequences using custom script. For the ORs, inconsistencies exist in the literature; e.g., the term olfactory receptor was used in mammals, birds, and cichlids, while the term odorant receptor was used in teleost fish species such as zebrafish, tetraodon, and stickleback. Therefore, “olfactory” and “odorant” were both used as keywords for sequence searching. We excluded genes from two olfactory families, including ORA (olfactory receptor class A-related) and OlfC (olfactory receptor C family), because these two families are not considered as canonical ORs [30,31]. All the sequences of ORs and TAARs used in the current project were provided in File S1 and File S2.

2.2. Identification, location determination, and nomenclature of ORs and TAARs in channel catfish

The catfish draft genome sequences [29] were first masked using RepeatMasker [32], and these sequences were used to predict putative protein sequences using FGENESH embedded in MolQuest [33]. The predicted catfish protein sequences were annotated through BLAST against the NCBI non-redundant database. The catfish protein sequences with hits to OR and TAAR genes were then extracted based on the annotation, and used for further analysis. In order to identify the OR and TAAR repertoires in the channel catfish genome as complete as possible, we determined the catfish ORs and TAARs based on: 1) clustering with other annotated genes in the phylogenetic tree; 2) clustering with other annotated genes into the same orthogroup; 3) possessing the seven trans-membrane topology structure. Then, the candidate OR and TAAR sequences of channel catfish were analyzed using PSF program embedded in MolQuest to identify pseudogenes. Three filters were used to identify pseudogenes, including in-frame premature stop codons, disruptive frameshifts, and poly-A tail found at the 3′ terminus in each sequence [34,35]. The genomic locations of ORs and TAARs were determined using BLAT [36].

We used a similar nomenclature method as described in previous studies for ORs and TAARs [21,37]. Genes were named with the same family/subfamily name if they were clustered with annotated genes from other species. Within a family/subfamily, genes were named with a number for ORs such as OR1, OR2, etc., and with a letter for TAARs such as TAARa, TAARb, etc., sequentially based on their genomic locations.

2.3. Phylogenetic and orthogroup analyses of ORs and TAARs in vertebrates

The ORs were aligned using MUSCLE [38], and the phylogenetic tree was constructed using FastTree with default settings [39,40]. The same were done for TAARs. The approximate-maximum-likelihood phylogenetic tree was constructed based on the JTT model of amino acid evolution and the approach of Bayesian. As a result, the local

Table 1
A summary description of OR and TAAR genes in the catfish genome.

Gene name	Chromosome location	Starting site	Ending site
OR134-3	11	13,099,351	13,100,296
OR107-1	17	1,328,968	1,332,011
OR111-2	17	1,358,658	1,360,632
OR111-3	17	1,365,698	1,367,760
OR111-4	17	1,377,637	1,379,669
OR111-5	17	1,382,204	1,386,070
OR106-1	17	1,389,660	1,390,602
OR106-2	17	1,395,184	1,396,126
OR103-1	17	1,422,180	1,423,164
OR102-2	17	1,467,906	1,504,626
OR102-3	17	1,508,517	1,515,209
OR129-1	17	22,333,164	22,334,549
OR118-1	17	4,146,465	4,163,108
OR118-2	17	4,164,004	4,178,618
OR118-3	17	4,189,595	4,196,693
OR119-2	17	6,363,016	6,374,383
OR117-1	17	6,378,556	6,381,065
OR113-2	17	9,849,401	9,850,343
OR113-3	17	9,854,555	9,855,497
OR113-5	17	9,864,054	9,875,885
OR132-3	18	14,614,357	14,618,677
OR132-6	18	14,648,565	14,649,528
OR115-2	18	23,034,072	23,035,008
OR109-3	19	2,961,239	2,963,779
OR109-4	19	2,972,007	2,986,472
OR109-5	19	3,010,029	3,023,962
OR128-1	24	20,517,994	20,523,574
TAAR1a	2	36,667,500	36,668,776
TAAR1c	2	36,675,995	36,683,695
TAAR11	2	36,685,914	36,689,038
TAAR1d	2	36,691,414	36,692,449
TAAR2a	2	36,693,494	36,701,745
TAAR13a	2	36,709,183	36,710,230
TAAR13b	2	36,712,698	36,717,655
TAAR13c	2	36,720,191	36,721,238
TAAR13d	2	36,724,908	36,725,937
TAAR13e	2	36,728,680	36,729,712
TAAR13f	2	36,733,256	36,736,935
TAAR2b	2	36,743,881	36,747,126
TAAR14a	2	36,763,888	36,764,869
TAAR14b	2	36,767,060	36,775,480
TAAR14c	2	36,778,121	36,781,493
TAAR14d	2	36,783,246	36,784,236
TAAR14f	2	36,794,780	36,795,770
TAAR14g	2	36,797,281	36,798,271
TAAR14i	2	36,803,798	36,812,629
TAAR14k	2	36,905,504	36,906,494
TAAR14l	2	36,911,863	36,918,794
TAAR14m	2	36,921,829	36,922,819
TAAR14n	2	36,932,944	36,933,934
TAAR14q	2	36,959,760	36,960,750
TAAR14r	2	36,965,881	36,966,871
TAAR14s	2	36,972,859	36,973,849
TAAR14u	2	36,984,592	36,988,579
TAAR15	16	22,550,131	22,553,100

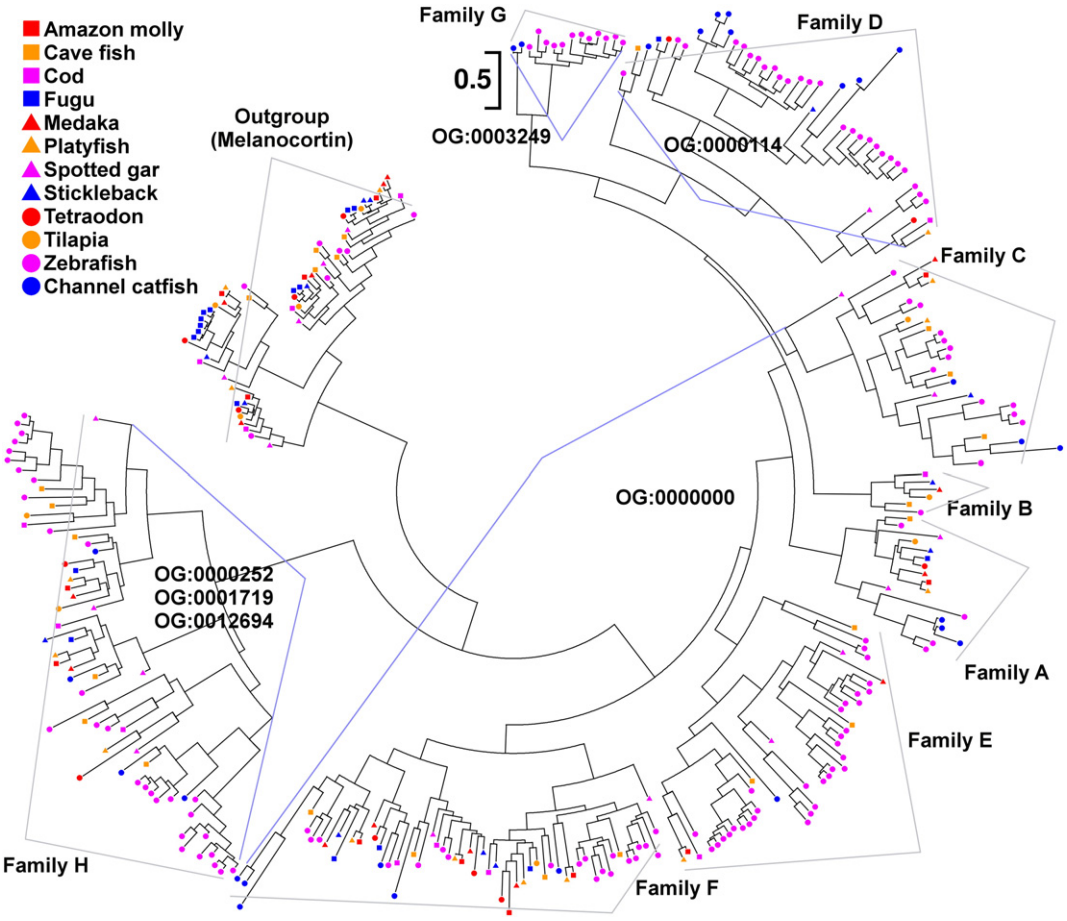

Fig. 1. A phylogenetic tree of ORs from 12 fish species (legends are indicated on the upper left side of the figure). Members from the same family are coverd with curves with their family names indicated outside of the curves in grey, while names of the orthogroups are indicated in the inside of the figure in blue. For example, orthogroup OG:0000000 included families A, B, C, E, and F.

support values were computed. Also, we used melanocortin and histamine receptor H2 as outgroups to root the phylogenetic trees for ORs and TAARs, respectively. MEGA6 was used for visualization [41]. We only used fish ORs to construct the phylogenetic tree, because the number of mammalian ORs is too large. Also, there are different nomenclature systems used among mammals, amphibians, birds and fish species.

Additionally, we conducted comparative genomic analysis using OrthoFinder [42], and orthogroups that contain ORs and TAARs were identified. Only clusters that consist of at least two orthologous or two paralogous were retained in our study [43].

2.4. Collinear analysis of ORs and TAARs between zebrafish and catfish

The genomic locations of the ORs and TAARs were compared between the catfish genome and zebrafish genome. Reciprocal blast (using catfish as query blast against zebrafish, and using zebrafish as query blast against catfish) and self-blast (self-blast of catfish, and self-blast of zebrafish) were conducted before running MCScanX for the detection of gene collinearity [44]. The results were visualized using Circos [45]. Tandem duplicated genes, which are defined as neighbors to each other while the distance between them is <10 kb, were also identified [46].

2.5. Identification of conserved motifs

Conserved motifs of ORs and TAARs were identified using MEME [47]. All amino acid sequences of ORs and TAARs were aligned using MUSCLE [38], and the gaps were removed using trimAl [48]. Only the top five conserved motifs were identified, with the motif length ranging from five to fifty.

2.6. dN/dS analysis

The dN/dS analysis was conducted for each subfamily of ORs and TAARs using Datamonkey [49]. Only genes that were found in at least two species were included for the analysis.

3. Results and discussion

3.1. Identification of ORs and TAARs in channel catfish

A total of 27 functional OR genes and 28 functional TAAR genes were identified in channel catfish. Furthermore, there were 20 OR pseudogenes and eight TAAR pseudogenes. The genomic locations of the ORs and TAARs were summarized in Table 1, and that of pseudogenes were summarized in Table S1. The identification of OR and TAAR pseudogenes in channel catfish was summarized in Table S2.

Mammals possess more ORs than TAARs, while fish species possess roughly equal ORs and TAARs. For instance, in humans, 339 functional ORs were identified while only 6 functional TAARs were identified [20,37]. In our study, the number of ORs was much closer to the number of TAARs. Similar results were also found in zebrafish after searching in ENSEMBL (159 ORs and 94 TAARs). One hypothesis for this phenomenon is that ligands or odorants detected by fish TAARs might be recognized by mammalian ORs, or vice versa [13,15,16,50]. Another hypothesis is that TAARs may play more important roles than ORs in the olfaction of fish species compared with mammals, owing to

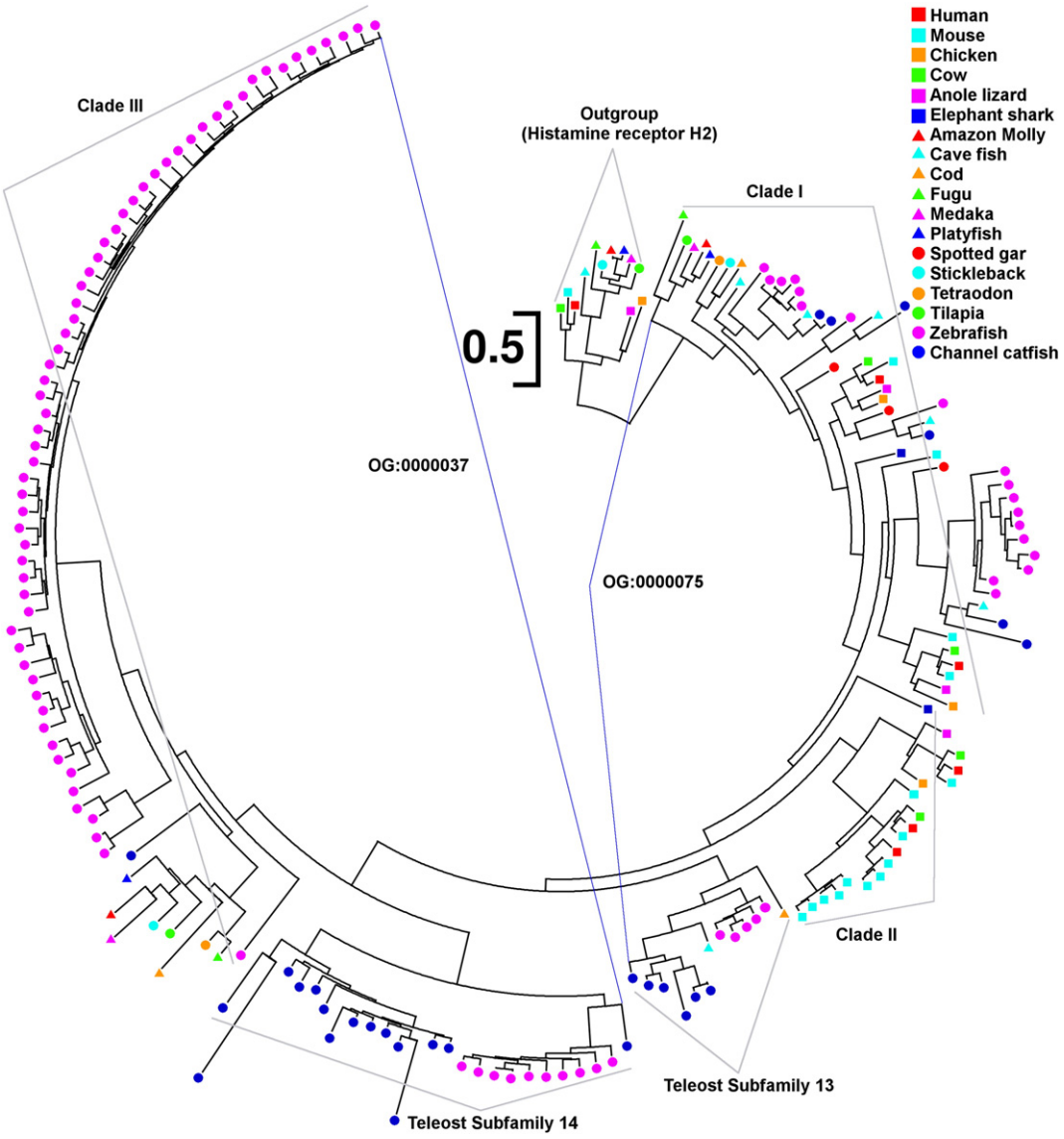

Fig. 2. A phylogenetic tree of TAARs from 18 vertebrate species (legends are indicated on the upper left side of the figure). Members from the same family are coverd with curves with their family names indicated outside of the curves in grey, while names of the orthogroups are indicated in the inside of the figure in blue. Orthogroup 0000075 included class I and class II TAAR genes while orthogroup 0000037 included class III TAAR genes.

different living environments. Generally, fish capture chemical compounds from water flux, while mammals capture odorant molecules from aspiratory flux [3].

3.2. Genomic organization of ORs and TAARs in channel catfish

The ORs (both functional genes and pseudogenes) were located on five chromosomes, including chromosomes 11, 17, 18, 19, and 24. As in zebrafish, ORs in catfish were organized in clusters, and each cluster contained at least five members [21]. Detailed gene coordinates were listed in Table 1. For instance, two clusters were found on chromosome 17, one on chromosome 18 and one on chromosome 19. In addition, several smaller clusters and single genes were scattered on chromosomes 11, 17, 18 and 24. As a matter of fact, members of each subfamily resided together, while a broad genomic distance existed between the subfamilies, suggesting that members within each subfamily could be derived from lineage-specific duplications.

The catfish TAARs (including functional genes and pseudogenes) were all located on chromosome 2 except for TAAR15, which was located on chromosome 16. On chromosome 2, TAARs spanned in a region of ~313 kb, and most of them were organized in the form of head-to-tail tandems (Table 1, and Fig. S1). Additionally, catfish TAARs were arranged in three clusters. Each cluster contained members from a single clade, and these clusters were distributed sequentially on chromosome 2 (Fig. S1). This result is quite distinct from the genomic organization of TAARs in zebrafish and Atlantic salmon, in which most members mainly clustered on two chromosomes [11,16].

3.3. Phylogenetic and orthogroup analyses of ORs and TAARs

A phylogenetic tree of ORs was displayed in Fig. 1. Following the nomenclature of ORs, seven families of ORs were identified in the catfish genome, including family A, family C, family D, family E, family F, family G and family H. Family B was not found from channel catfish (Fig. 1 and Fig. S2). Notably, zebrafish is the only fish species that have members across all families. For instance, families B and G were not found in pufferfish [21].

Six orthogroups were identified for fish ORs (abbreviated as OG in Fig. 1). The largest orthogroup, OG:0000000, comprised family A, family B, family C, family E, and family F. Family D belonged to orthogroup

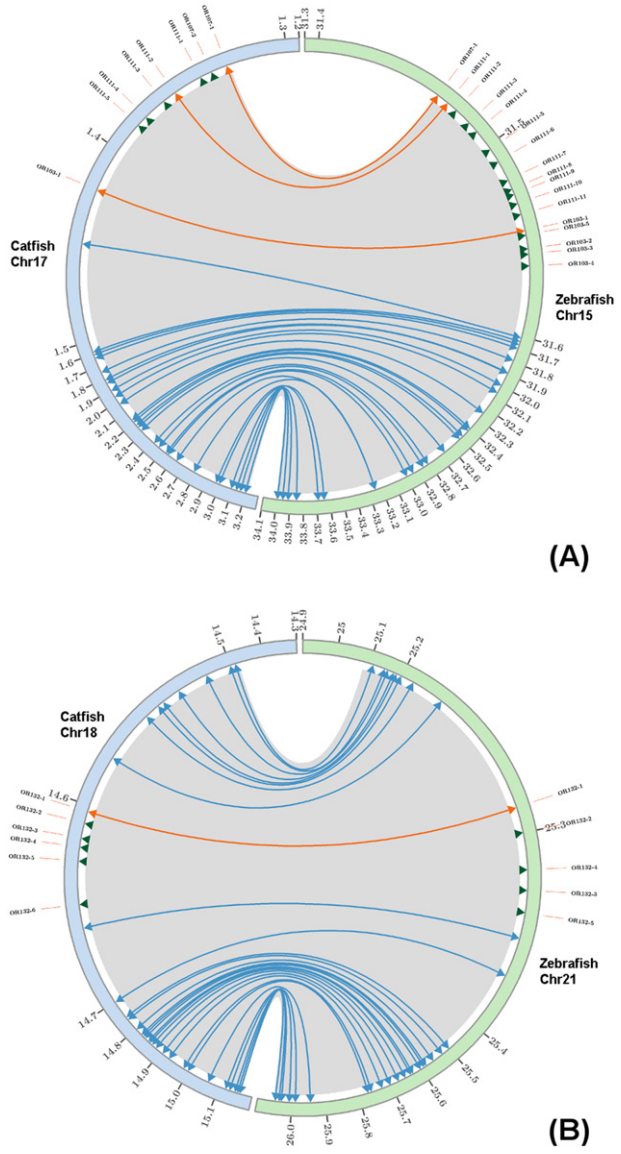

Fig. 3. Identification of a pair of orthologous chromosomal regions between channel catfish and zebrafish. The first region is between zebrafish chromosme 15 and catfish chromosome 17, and the second region is between zebrafish chromosome 21 and catfish chromosome 18. Anchor genes are linked with arrowed lines. ORs are linked with orange arrowed lines while other genes are linked with blue arrowed lines. Genomic positions for both fishes are indicated with numbers (in Mb) along its own chromosome. Tandem duplicated genes are indicated with green triangles.

OG:0000114, and family G belonged to orthogroup OG:0003249 (Fig. 1, and Table S3). For Family H, subfamilies 132, 133, 134, 135, 136, and 137 fell into orthogroup OG:0000252, subfamilies 130 and 131 fell into orthogroup OG:0001719, and subfamily 129 fell into orthogroup OG:0012694.

In this study, we used both the phylogenetic and orthogroup analyses to elucidate the evolutionary dynamics of ORs in fish species. The phylogenetic analysis was mainly based on pairwise comparison of ORs, while the orthogroup analysis was based on not only the self-comparison within each species genome but also the pairwise comparison among all species genomes used in the present study. Thus, the combination of these two analyses enables us to identify the catfish ORs properly.

Previous studies revealed that tandem duplication was a major type of duplication in teleosts [46], and might contribute to their evolution [44,51]. Furthermore, several tandem duplicated genes involved in sensory response pathways were enriched in zebrafish, and one olfaction-related GO term (olfactory receptor activity) was identified among these genes [46]. As in our results, most catfish ORs were originated from lineage-specific tandem duplication, while the remaining ORs were originated from the most recent common ancestor of all fish species.

A phylogenetic tree of TAARs was displayed in Fig. 2. Twenty-eight catfish TAARs were identified in the catfish genome, of which six belonged to Clade I, one belonged to Clade III, six belonged to subfamily 13, and 15 belonged to subfamily 14 (Fig. 2 and Fig. S3). It is apparent that fish TAARs possess a characteristic of species-specific gene expansion. For instance, in subfamily TAAR14, gene expansion led to the presence of 10 TAARs in zebrafish, and 15 TAARs in catfish (Fig. 2). Similar results were also found in Atlantic salmon [16]. TAAR gene expansion was also found in subfamily TAAR13 for both catfish and zebrafish (Fig. 2).

Two orthologous groups were identified for vertebrate TAARs. Orthogroup OG:0000075 covered all genes in Clade I, Clade II and subfamily 13 (Fig. 2, and Table S3). Orthogroup OG:0000037 comprised all genes from Clade III and subfamily 14 (Fig. 2, and Table S3). In the phylogenetic analysis, Clade I and Clade II were also clustered together in a single clade, indicating that these two clades of TAARs might have evolved slower and were not divergent from each other over the course of evolution. This is consistent with the hypothesis that Clade III was derived later than Clades I and II, but evolved faster than Clades I and II after the separation of teleosts and tetrapods [11].

3.4. Collinear analysis of ORs and TAARs

Two putative orthologous regions that contain ORs as anchor genes were identified between the catfish genome and zebrafish genome (expectation value = 0) (Fig. 3, and Table S4). One of them comprised 30 pairs of anchor genes, while the other comprised 31 pairs. The orders of anchor genes were well conserved, indicating reliable results as revealed by collinear analysis (Fig. 3, and Table S4). Actually, only one pair was not annotated with the same gene name. Zebrafish OR131-1 was listed as the collinear gene of catfish OR131-2 based on two facts: 1) catfish OR131-1 was identified as pseudogene; and 2) catfish OR131-1 and OR131-2 were identified as tandem duplicated genes (Table S5). Putative orthologous regions that contain TAARs as anchor genes were not identified.

Considering that gene expansion was originated from tandem duplication, the relative birth time of catfish ORs can be inferred within subfamilies 111 and 132 (Fig. 3). The catfish OR111-1 and zebrafish OR111-1 were listed as a pair of collinear genes, while the others were identified as tandem duplicated genes (Fig. 3A). Previous studies reported that collinearity could be introduced after the divergence of most teleosts [52], and species-specific tandem duplications were found in several teleost fishes [46]. Therefore, it is reasonable to infer that both catfish OR111-1 and zebrafish OR111-1 are the most ancient genes in their respective subfamilies. These ORs were originated before the divergence of catfish and zebrafish, while the other members were derived from more recent tandem duplication events. The same conclusion can be drawn for catfish subfamily OR132 (Fig. 3B).

3.5. Identification of conserved motifs for ORs and TAARs

To characterize the two-dimensional structure of fish ORs, five conserved motifs were identified. The motifs' logos, as well as their corresponding locations in the topology structure of OR (here, we used catfish OR115-2 as an example), were displayed in Fig. 4. In general, most fish ORs possessed all these five conserved motifs, with the exception of family H. We succeeded to recap the sequence pattern of "MAYDRYVAIC" within motif 1 (Fig. 4A), which was highly conserved in teleost fish ORs [21]. Four of these conserved motifs were spanned on the junction of trans-membrane regions and intracellular loops,

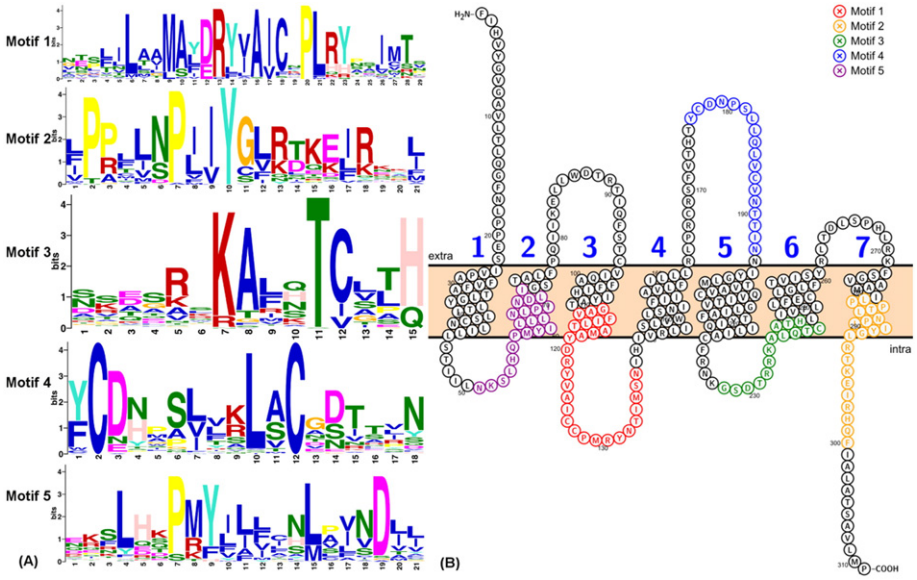

Fig. 4. Logo representation and distribution of the five best conserved motifs identified for teleost ORs. (A) Sequence logos of the conserved motifs, as the degree of conservation is indicated by the height of amino acid code. (B) The distribution of these motifs as displayed in the two-dimensional topology structure of ORs. The blue numbers represent the number of each trans-membrane domain.

while only one motif was spanned on an extracellular loop (Fig. 4B). Strikingly, the positions of these conserved motifs were overlapped with potential binding sites in the mammalian ORs [53], indicating that the binding sites of fish ORs might be similar to that of mammals.

Five best-conserved motifs were identified for fish TAARs. The logo presentations of these motifs, as well as their corresponding locations in the topology structure of TAARs (here, we used catfish taar14n as an example), were displayed in Fig. 5. Of these five motifs, motif 2 contained "NSXXNPXXYXXXYXWF" (where "X" represents any amino acid residue) (Fig. 5A), which is considered as the TAAR fingerprint motif [8]. Motif 3 possessed the sequence pattern of "DRY" (Fig. 5A), which can coordinate the conformational status of TAARs and then affect the binding affinity of TAAR [54]. The rest of the five motifs all contained conserved amino acid(s) identified in a previous study [8], most of which were located on trans-membrane domains (Fig. 5B). The distribution of all the conserved motifs was well consistent with the distribution of predicted ligand-binding sites [16], which are essential for the formation of ligand pocket vector [37]. We speculate that even though the sequence divergences are large among fish TAARs, the components and positions of ligand binding residues of fish TAARs remain highly conserved.

We observed that the conserved motifs' arrangement, in most cases (four pairs of conserved motifs), were generally identical between fish ORs and TAARs. For example, a conserved motif, which was spanned on the junction of the sixth trans-membrane region and the third intracellular loop, was identified in both ORs and TAARs. The sequence pattern of "DRY" was located on the same position of these two receptors, at the beginning of the second intracellular loop. Since both ORs and TAARs are involved in olfaction, they must initiate the same or similar intracellular signal cascades. Therefore, the similar distribution pattern of conserved motifs may indicate that ORs and TAARs have similar functions. We further explored this hypothesis through identifying conserved motifs for ORs and TAARs together. Three conserved motifs (data not shown) were overlapped with regions described above,

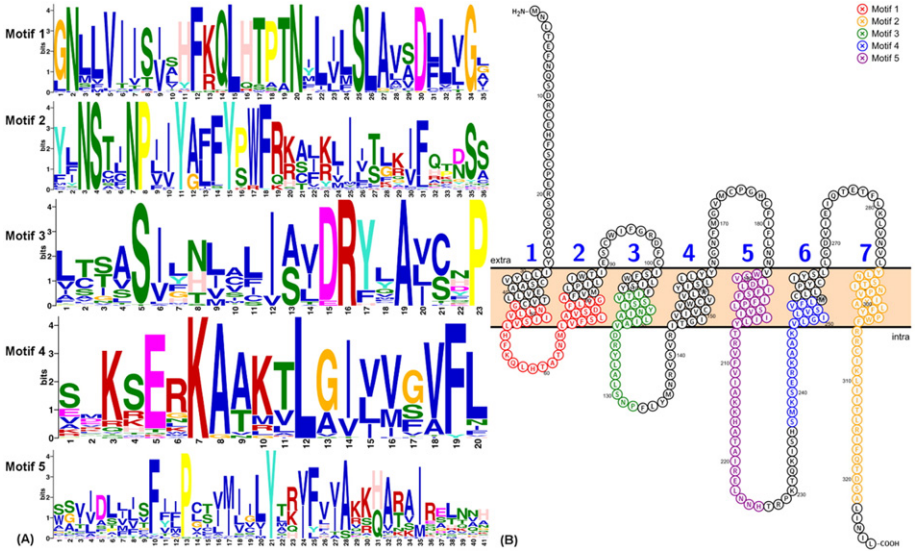

Fig. 5. Logo representation and distribution of the five best conserved motifs identified for teleost TAARs. (A) Sequence logos of the conserved motifs, as the degree of conservation is indicated by the height of amino acid code. (B) The distribution of these motifs as displayed in the two-dimensional topology structure of TAARs. The blue numbers represent the number of each trans-membrane domain.

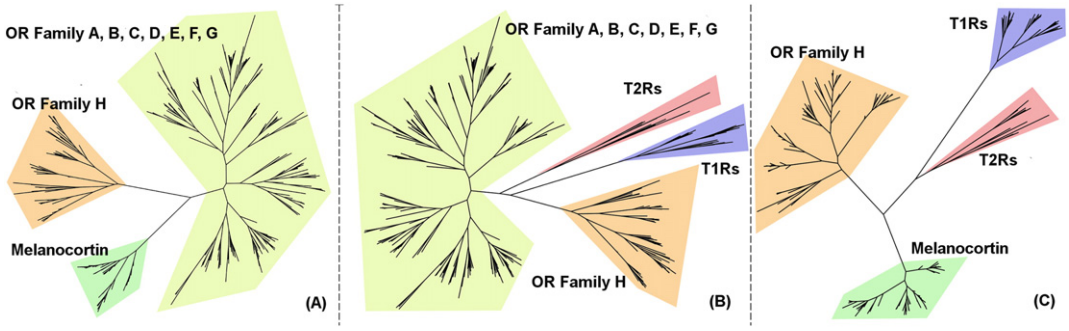

Fig. 6. Divergence of family H from other ORs. (A) A phylogenetic tree of all OR families with melanocortin. (B) A phylogenetic tree of all OR families with taste receptors. (C) A phylogenetic tree of family H with melanocortin and taste receptors.

indicating that fish ORs and TAARs share similar sequence patterns at certain locations. Taken together, we inferred that fish TAARs might share the same function with fish ORs based on the motif analysis, but functional studies for validation are still needed in the future.

3.6. Divergence of family H from other OR families

Family H is considered as a group of fish ORs that were originated from an ancient duplication event [21]. This was also found in our phylogenetic analysis, with the first event being the divergence of family H from other families (Fig. 5). Family H was not clustered with other families (Fig. 6A) even using more evolutionary distant gene families as outgroups (Fig. 6B). However, family H was clustered into a single clade that was clearly separated from these outgroups (Fig. 6C). Here, the conserved motif analysis for all the fish OR families allowed us to unveil a new mechanism underlying this phenomenon. Motif 4 was not found in family H, suggesting that the loss of motif 4 may contribute to the divergence of family H from other families (Fig. 4B). Interestingly, this motif, spanning entirely on the second extracellular loop, contains the cysteine residue that is essential for the formation of ligand binding pocket [53]. Thus, we inferred that the fish ORs from family H might lose their sensing ability completely, or that they may possess weaker ligand-receptor affinities compared with fish ORs in other OR families.

3.7. dN/dS analysis

To measure the natural selection pressure that was imposed on fish ORs and TAARs, the global synonymous (dS) and non-synonymous (dN) rates were calculated for selected subfamilies. The global ratios of dN/dS were well below 1.0 for all subfamilies, a theoretical boundary for positive and negative selection (Fig. 7A). The fish TAAR14s exhibited the highest dN/dS ratio, followed by the TAAR13s. We speculated that these two subfamilies were inclined to increase the frequency of some certain alleles under selective pressure. To explore this hypothesis, we conducted site-by-site (or codon-by-codon) analysis for each subfamily. As expected, positive selection sites ($p < 0.1$) were found in subfamily 14 (Fig. 7B). Previous study reported that two zebrafish TAAR14s contain both $Asp^{3.32}$ and $Asp^{5.42}$, indicating that a transformation occurred in fish TAARs [18]. In our results, nine catfish TAAR14s contained both $Asp^{3.32}$ and $Asp^{5.42}$, including TAAR14c, TAAR14f, TAAR14g, TAAR14k, TAAR14m, TAAR14n, TAAR14q, TAAR14r and TAAR 14s, and they could be candidate diamine receptors. Therefore, we conclude that fish TAAR subfamily 14 was imposed with the highest degree of natural selection pressure among all fish ORs and TAARs.

4. Conclusions

In the present study, we report the complete repertoires of ORs and TAARs in channel catfish. Two conserved orthologous blocks that contain ORs as anchor genes were identified between the catfish genome and zebrafish genome. The arrangements of conserved motifs were generally identical between fish ORs and TAARs. The complete loss of a conserved motif in OR family H might contribute to its divergence from other families. The highest level of selection pressure was imposed on fish TAAR subfamily 14 among all fish OR and TAAR subfamilies.

Supplementary data to this article can be found online at doi:10.1016/j.bbagen.2016.10.017.

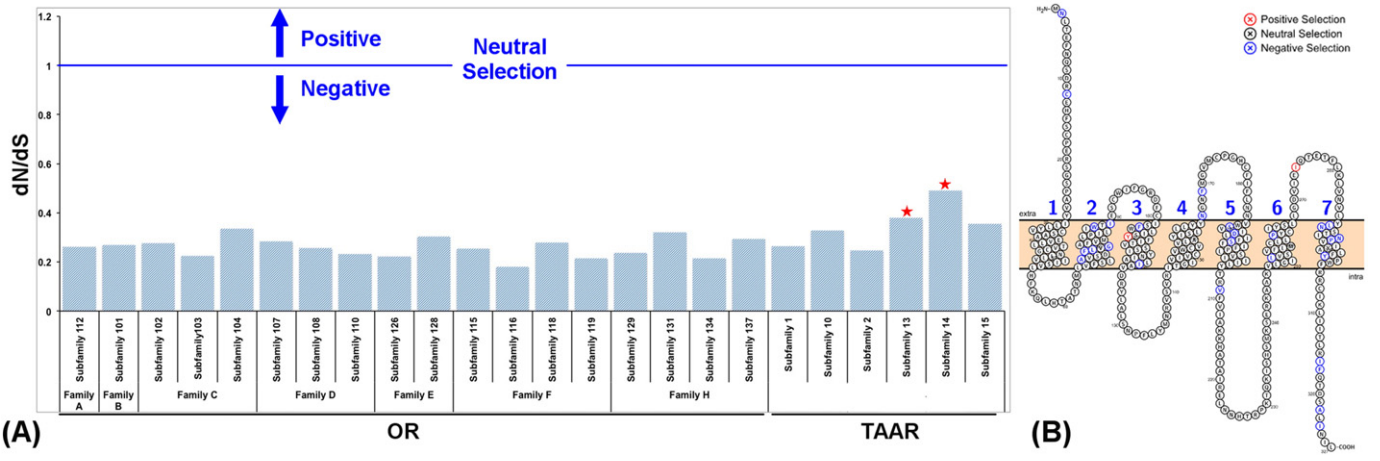

Fig. 7. Selection pressure imposed on both ORs and TAARs. (A) dN/dS ratios of each subfamily for both OR and TAAR. (B) The distribution of positive, neutral and negative sites for teleost Class III TAAR as displayed in two-dimensional topology structure.

Transparency Document

The Transparency document associated with this article can be found in the online version.

Acknowledgements

This project was supported by USDA National Institute of Food and Agriculture through grants from Agriculture and Food Research Initiative Animal Genomics, Genetics and Breeding Program (2015-67015-22907) and Animal Disease Program (2015-67015-22975). S. Gao was supported by a scholarship from the China Scholarship Council (CSC 201206330079).

References

[1] J. Caprio, M. Shimohara, T. Marui, J. Kohbara, S. Harada, S. Kiyohara, Amino acid specificity of fibers of the facial/trigeminal complex innervating the maxillary barbel in the Japanese sea catfish, *Plotosus japonicus*, Physiol. Behav. 152 (2015) 288–294.
[2] S. DeMaria, J. Ngai, The cell biology of smell, J. Cell Biol. 191 (2010) 443–452.
[3] J. Fleischer, H. Breer, J. Strotmann, Mammalian Olfactory Receptors, Design Principles of Sensory Receptors, 2013 45.
[4] A. Sheybani, M. Nusnbaum, J. Caprio, C.D. Derby, Responses of the sea catfish *Ariopsis felis* to chemical defenses from the sea hare *Aplysia californica*, J. Exp. Mar. Biol. Ecol. 368 (2009) 153–160.
[5] T.L. Meredith, J. Caprio, S.M. Kajiura, Sensitivity and specificity of the olfactory epithelia of two elasmobranch species to bile salts, J. Exp. Biol. 215 (2012) 2660–2667.
[6] K.A. Johnstone, K.P. Lubieniecki, B.F. Koop, W.S. Davidson, Identification of olfactory receptor genes in Atlantic salmon *Salmo salar*, J. Fish Biol. 81 (2012) 559–575.
[7] I. Rodriguez, Singular expression of olfactory receptor genes, Cell 155 (2013) 274–277.
[8] L. Lindemann, M.C. Hoener, A renaissance in trace amines inspired by a novel GPCR family, Trends Pharmacol. Sci. 26 (2005) 274–281.
[9] D.M. Ferrero, J.K. Lemon, D. Fluegge, S.L. Pashkovski, W.J. Korzan, S.R. Datta, M. Spehr, M. Fendt, S.D. Liberles, Detection and avoidance of a carnivore odor by prey, Proc. Natl. Acad. Sci. 108 (2011) 11235–11240.
[10] A. Hussain, L.R. Saraiva, D.M. Ferrero, G. Ahuja, V.S. Krishna, S.D. Liberles, S.I. Korsching, High-affinity olfactory receptor for the death-associated odor cadaverine, Proc. Natl. Acad. Sci. 110 (2013) 19579–19584.
[11] A. Hussain, L.R. Saraiva, S.I. Korsching, Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts, Proc. Natl. Acad. Sci. 106 (2009) 4313–4318.
[12] Q. Li, W.J. Korzan, D.M. Ferrero, R.B. Chang, D.S. Roy, M. Buchi, J.K. Lemon, A.W. Kaur, L. Stowers, M. Fendt, Synchronous evolution of an odor biosynthesis pathway and behavioral response, Curr. Biol. 23 (2013) 11–20.
[13] S.D. Liberles, Trace amine-associated receptors are olfactory receptors in vertebrates, Ann. N. Y. Acad. Sci. 1170 (2009) 168–172.
[14] S.D. Liberles, Trace amine-associated receptors: ligands, neural circuits, and behaviors, Curr. Opin. Neurobiol. 34 (2015) 1–7.
[15] S.D. Liberles, L.B. Buck, A second class of chemosensory receptors in the olfactory epithelium, Nature 442 (2006) 645–650.
[16] J.A. Tessarolo, M.J. Tabesh, M. Nesbitt, W.S. Davidson, Genomic organization and evolution of the trace amine-associated receptor (TAAR) repertoire in Atlantic salmon (*Salmo salar*), G3: Genes|Genomes|Genetics (2014) (g3. 114.010660).
[17] D.M. Ferrero, D. Wacker, M.A. Roque, M.W. Baldwin, R.C. Stevens, S.D. Liberles, Agonists for 13 trace amine-associated receptors provide insight into the molecular basis of odor selectivity, ACS Chem. Biol. 7 (2012) 1184–1189.
[18] Q. Li, Y. Tachie-Baffour, Z. Liu, M.W. Baldwin, A.C. Kruse, S.D. Liberles, Non-classical amine recognition evolved in a large clade of olfactory receptors, eLife 4 (2015), e10441.
[19] X. Zhang, S. Firestein, The olfactory receptor gene superfamily of the mouse, Nat. Neurosci. 5 (2002) 124–133.
[20] B. Malnic, P.A. Godfrey, L.B. Buck, The human olfactory receptor gene family, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 2584–2589.
[21] T.S. Alioto, J. Ngai, The odorant receptor repertoire of teleost fish, BMC Genomics 6 (2005) 173.
[22] D.E.I. Gloriam, T.K. Bjarnadóttir, Y.-L. Yan, J.H. Postlethwait, H.B. Schiöth, R. Fredriksson, The repertoire of trace amine G-protein-coupled receptors: large expansion in zebrafish, Mol. Phylogenet. Evol. 35 (2005) 470–482.
[23] Y. Hashiguchi, M. Nishida, Evolution of trace amine–associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class of vertebrate chemosensory receptors expressed in the olfactory epithelium, Mol. Biol. Evol. 24 (2007) 2099–2107.
[24] J.M. Parker, Q. Chang, J. Caprio, Citrate enhances olfactory receptor responses and triggers oscillatory receptor activity in the channel catfish, J. Neurophysiol. 83 (2000) 2676–2681.
[25] A. Hansen, S.H. Rolen, K. Anderson, Y. Morita, J. Caprio, T.E. Finger, Correlation between olfactory receptor cell type and function in the channel catfish, J. Neurosci. 23 (2003) 9328–9339.
[26] A.A. Nikonov, J. Caprio, Electrophysiological evidence for a chemotopy of biologically relevant odors in the olfactory bulb of the channel catfish, J. Neurophysiol. 86 (2001) 1869–1876.
[27] A.A. Nikonov, J. Caprio, Odorant specificity of single olfactory bulb neurons to amino acids in the channel catfish, J. Neurophysiol. 92 (2004) 123–134.
[28] A.A. Nikonov, J.M. Parker, J. Caprio, Odorant-induced olfactory receptor neural oscillations and their modulation of olfactory bulbar responses in the channel catfish, J. Neurosci. 22 (2002) 2352–2362.
[29] Z. Liu, S. Liu, J. Yao, L. Bao, J. Zhang, Y. Li, C. Jiang, L. Sun, R. Wang, Y. Zhang, T. Zhou, Q. Zeng, Q. Fu, S. Gao, N. Li, S. Koren, Y. Jiang, A. Zimin, P. Xu, A.M. Phillippy, X. Geng, L. Song, F. Sun, C. Li, X. Wang, A. Chen, Y. Jin, Z. Yuan, Y. Yang, S. Tan, E. Peatman, J. Lu, Z. Qin, R. Dunham, Z. Li, T. Sonstegard, J. Feng, R.G. Danzmann, S. Schroeder, B. Scheffler, M.V. Duke, L. Ballard, H. Kucuktas, L. Kaltenboeck, H. Liu, J. Armbruster, Y. Xie, M.L. Kirby, Y. Tian, M.E. Flanagan, W. Mu, G.C. Waldbieser, The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts, Nat. Commun. 7 (2016) 11757.
[30] T.S. Alioto, J. Ngai, The repertoire of olfactory C family G protein-coupled receptors in zebrafish: candidate chemosensory receptors for amino acids, BMC Genomics 7 (2006) 1.
[31] L.R. Saraiva, S.I. Korsching, A novel olfactory receptor gene family in teleost fish, Genome Res. 17 (2007) 1448–1457.
[32] A.F.A. Smit, R. Hubley, P. Green, RepeatMasker Open-4.0. 2013–2015, 2014.
[33] V. Solovyev, P. Kosarev, I. Seledsov, D. Vorobyev, Automatic annotation of eukaryotic genes, pseudogenes and promoters, Genome Biol. 7 (2006) S10.
[34] R.C. Pink, K. Wicks, D.P. Caley, E.K. Punch, L. Jacobs, D.R.F. Carter, Pseudogenes: pseudo-functional or key regulators in health and disease? RNA 17 (2011) 792–798.
[35] D. Zheng, A. Frankish, R. Baertsch, P. Kapranov, A. Reymond, S.W. Choo, Y. Lu, F. Denoeud, S.E. Antonarakis, M. Snyder, Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution, Genome Res. 17 (2007) 839–851.
[36] W.J. Kent, BLAT—the BLAST-like alignment tool, Genome Res. 12 (2002) 656–664.
[37] L. Lindemann, M. Ebeling, N.A. Kratochwil, J.R. Bunzow, D.K. Grandy, M.C. Hoener, Trace amine-associated receptors form structurally and functionally distinct subfamilies of novel G protein-coupled receptors, Genomics 85 (2005) 372–385.
[38] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (2004) 1792–1797.
[39] M.N. Price, P.S. Dehal, A.P. Arkin, FastTree: computing large minimum evolution trees with profiles instead of a distance matrix, Mol. Biol. Evol. 26 (2009) 1641–1650.
[40] M.N. Price, P.S. Dehal, A.P. Arkin, FastTree 2–approximately maximum-likelihood trees for large alignments, PLoS One 5 (2010), e9490.
[41] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729.
[42] D.M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy, Genome Biol. 16 (2015) 1–14.
[43] R.L. Tatusov, E.V. Koonin, D.J. Lipman, A genomic perspective on protein families, Science 278 (1997) 631–637.
[44] Y. Wang, H. Tang, J.D. DeBarry, X. Tan, J. Li, X. Wang, T.-h. Lee, H. Jin, B. Marler, H. Guo, MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity, Nucleic Acids Res. 40 (2012) e49.
[45] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S.J. Jones, M.A. Marra, Circos: an information aesthetic for comparative genomics, Genome Res. 19 (2009) 1639–1645.
[46] J. Lu, E. Peatman, H. Tang, J. Lewis, Z. Liu, Profiling of gene duplication patterns of sequenced teleost genomes: evidence for rapid lineage-specific genome expansion mediated by recent tandem duplications, BMC Genomics 13 (2012) 246.
[47] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME SUITE: tools for motif discovery and searching, Nucleic Acids Res. 37 (2009) W202–W208.
[48] S. Capella-Gutierrez, J.M. Silla-Martinez, T. Gabaldon, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses, Bioinformatics 25 (2009) 1972–1973.
[49] W. Delport, A.F. Poon, S.D. Frost, S.L. Kosakovsky Pond, Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology, Bioinformatics 26 (2010) 2455–2457.
[50] N. Azzouzi, F. Barloy-Hubler, F. Galibert, Identification and characterization of cichlid TAAR genes and comparison with other teleost TAAR repertoires, BMC Genomics 16 (2015) 1–14.
[51] M. Freeling, Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition, Annu. Rev. Plant Biol. 60 (2009) 433–453.
[52] A. Meyer, Y. Van de Peer, From 2R to 3R: evidence for a fish-specific genome duplication (FSGD), BioEssays 27 (2005) 937–945.
[53] O. Man, Y. Gilad, D. Lancet, Prediction of the odorant binding site of olfactory receptor proteins by human-mouse comparisons, Protein Sci. 13 (2004) 240–254.
[54] G.E. Rovati, V. Capra, R.R. Neubig, The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state, Mol. Pharmacol. 71 (2007) 959–964.