Evolutionary Transition From Water to Land in Vertebrates Illuminated by Basal Ray-Finned Fish Vomeronasal Type 2 Receptor (OlfC) Genes

Zicong Zhang

Tokyo Institute of Technology

Atsuhiko Sakuma

Tokyo Institute of Technology

Shigehiro Kuraku

RIKEN Center for Biosystems Dynamics Research (BDR)

Masato Nikaido (✉ mnikaido@bio.titech.ac.jp)

Tokyo Institute of Technology

Research Article

Keywords: Evolutionary, transition, water, land, finned fish, vomeronasal, genes, multigene family, peptide pheromones in mammals, tetrapods, environment

Posted Date: September 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-504504/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Evolutionary transition from water to land in vertebrates illuminated by basal ray-finned fish vomeronasal type 2 receptor (*OlfC*) genes

Zicong Zhang¹, Atsuhiko Sakuma¹, Shigehiro Kuraku^{2,3} and Masato Nikaido^{*,1}

¹School of Life Science and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo, 152-8550, Japan

²Molecular Life History Laboratory, National Institute of Genetics, Mishima, Shizuoka, 411-8540, Japan

³Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Hyogo, 650-0047, Japan

* Corresponding author:

E-mail: mnikaido@bio.titech.ac.jp

[Tel: +81-3-5734-2659](tel:+81-3-5734-2659), [Fax: +81-3-5734-2946](tel:+81-3-5734-2946),

Abstract

The vomeronasal type 2 receptor (*V2R*, also called *OlfC*) multigene family is found in a broad range of jawed vertebrates from cartilaginous fish to tetrapods. *V2Rs* encode receptors for food-related amino acids in teleost fish, whereas for peptide pheromones in mammals. In addition, *V2Rs* of teleost fish are phylogenetically distinct from those of tetrapods, implying a drastic change in the *V2R* repertoire during terrestrial adaptation. To understand the process of diversification of *V2Rs* in vertebrates from “fish-type” to “tetrapod-type”, we conducted an exhaustive search for *V2Rs* in cartilaginous fish (chimeras, sharks, and skates) and basal ray-finned fish (reedfish, sterlet, and spotted gar), and compared them with those of teleost, coelacanth, and tetrapods. Phylogenetic and synteny analyses on 1897 *V2Rs* revealed that basal ray-finned fish possess unexpectedly higher number of *V2Rs* compared with cartilaginous fish, implying that *V2R* gene repertoires expanded in the common ancestor of Osteichthyes. Furthermore, reedfish and sterlet possessed various *V2Rs* that belonged to both “fish-type” and “tetrapod-type”, suggesting that the common ancestor of Osteichthyes possess “tetrapod-type” *V2Rs* although they inhabited underwater environments. Thus, the unexpected diversity of *V2Rs* in basal ray-finned fish illuminates the process of how the osteichthyan ancestors adapt from water to land.

Introduction

Olfaction is a critical chemosensory system for eliciting social behaviors in vertebrates, including reproduction, kin recognition, aggression, and feeding. The vertebrate olfaction system has experienced drastic changes in anatomy and neurophysiology during adaptation from water to land, one of the most important events is vertebrate evolution. Specifically, aquatic vertebrates detect water-soluble chemicals using their olfactory epithelium (OE), whereas terrestrial vertebrates detect both volatile and nonvolatile chemicals by differentiating their OE into main olfactory epithelium (MOE) and vomeronasal organ (VNO). Accompanied by the terrestrial adaptation of olfaction, the chemosensory receptors are proposed to have undergone major innovative diversification, although the detailed evolutionary process of diversification remains to be poorly understood at the DNA level.

Olfaction of vertebrates is mainly composed of four types of G protein-coupled receptors (GPCRs), namely, the olfactory receptor (*OR*), vomeronasal receptor type I (*V1R*), vomeronasal receptor type II (*V2R*), and trace amine-associated receptor (*TAAR*), all forming multigene families¹. In teleost, *V2Rs*, also referred to as *OlfCs* (olfactory receptors classified as type C GPCRs)^{2,3}, are expressed in OE of the nasal cavity. Several independent studies have shown that teleost *V2Rs* detect amino acids that elicit certain feeding behaviors. For example, *V2Rs* are expressed in microvillous sensory neurons of zebrafish and respond to amino acids, but not bile acids or sex pheromones⁴. In addition, the genetic blockage of neural transmission in the *V2R*-expressing neurons abolishes the attractive response to a mixture of amino acids⁵. However, given that Yang *et al.*⁶ have proposed a possible contribution of some *V2Rs* to elicit fright reactions, it is premature to rule out the possibility that *V2Rs* detect some chemicals other than amino acids. In tetrapods, two anatomically distinct organs, that is, MOE and VNO, each mainly detecting odorants and pheromones. *V2Rs* are specifically expressed in VNO of mice⁷, frogs⁸, and reptiles⁹. Hence, they are believed to encode pheromone receptors. Indeed, it has been shown that *V2Rs* detect peptide pheromones^{10,11} and peptides for the major histocompatibility complex¹².

Until now, phylogenetic analyses of *V2Rs* have revealed that teleost possesses 20–60 *V2Rs*^{2,13,14}. These *V2Rs* are divided into 16 subfamilies, whereas mammalian *V2Rs* are genetically closely related to each other. *V2Rs* of teleost are revealed to be distinct from those of tetrapods in the phylogenetic tree^{13,15,16}, implying a drastic change in the *V2R* repertoire during terrestrial adaptation. Therefore, for simplicity, we used the term “fish-type” *V2Rs* to refer to putative amino acid receptors of teleosts, while “tetrapod-type” to refer to putative peptide pheromone receptors of tetrapods. The nomenclature with “fish-type” and “tetrapod-type” was also adopted for *V1Rs*, based on the initial discoveries from genomes of model organisms (mouse and zebrafish)¹⁷. The “fish-type” and “tetrapod-type” *V2Rs* are also found to be distinct in their synteny relationships in that the former *V2Rs* form single large cluster in particular chromosomes^{13,18} and the latter *V2Rs* are scattered on several chromosomes¹⁹. It is notable that the coelacanths, which are the close relatives of tetrapods, have both “fish-type” and “tetrapod-

type” *V2Rs*¹⁶, showing that the coelacanth is an important organism as it serves as a missing link to fill an evolutionary gap between vertebrates under water and land²⁰. Recent studies on several shark genomes have shown that olfaction in cartilaginous fish is dominated by *V2Rs* rather than conventional *ORs*^{21,22,23}, which is consistent with the ultrastructural observation that the presence of only microvillous sensory neurons in OE²⁴ and with the immunohistochemical observations that most neurons in OE are positive for Go antibodies²⁵. In addition, a previous study has shown that *V2Rs* were absent in lamprey genomes²⁶. Based on the above findings, *V2Rs* are considered to have originated in the jawed vertebrate ancestor before the split between the two extant descendant lineages, that is, cartilaginous fish and Osteichthyes (ray-finned and lobed-finned fish including tetrapods).

In this study, we have performed a comprehensive exploration and phylogenetic analyses for *V2Rs* in nine cartilaginous fish species (one elephant shark, one rabbit fish, four sharks, two skates, and sawfish), three basal ray-finned fish (reedfish, sterlet, and spotted gar), two teleost fish (eel, zebrafish), a coelacanth, two amphibians (western clawed frog and caecilian), an anole lizard, and a mouse to elucidate the process of diversification of “fish-type” and “tetrapod-type” *V2Rs* in vertebrates. As a result, we have characterized 9 to 42 *V2Rs* in cartilaginous fish, and a large amount (47 to 189) of *V2Rs* in basal ray-finned fish. Phylogenetic analyses of *V2Rs* of 19 various vertebrate species have revealed the existence of “tetrapod-type” *V2Rs* in the genomes of reedfish and sterlet, implying the presence of an evolutionary seed for mammalian peptide pheromone receptors in the basal ray-finned fish. The result of this fine-scale phylogenetic study would provide important insights into understanding the process of olfactory adaptations in the *V2Rs* of vertebrates from cartilaginous fish to tetrapods.

Results

Characterization of *V2Rs* from the genomes of a broad range of vertebrates

We explored *V2R* gene repertoires from the genomes of nine cartilaginous fish (one elephant shark, one rabbit fish, four sharks, and three rays), three basal ray-finned fish (reedfish, sterlet, and spotted gar), two teleost fish (zebrafish, Japanese eel), a coelacanth, two amphibians (western clawed frog, caecilian), an anole lizard, and a mouse using an original software “fate”²⁷. The copy numbers of *V2Rs* in individual species are summarized and provided in Table 1. The number of *V2Rs* in teleost fish was mostly comparable to previous studies with some update. For example, zebrafish that possessed 72 *V2Rs*, which was larger than in previous studies^{13,20}, comprising 19 genes from unplaced scaffolds and 53 genes from chromosome 18. The numbers of *V2Rs* in cartilaginous fish ranged from 8 (rays) to 41 (chimera), which were found to be smaller than those of the teleost fish. In contrast, basal ray-finned fish possessed an unexpectedly large number of *V2Rs*. In reedfish, we identified 188 intact *V2Rs*, including 68 genes from unplaced scaffolds, which was the largest in all ray-finned fish studied until now. The numbers of *V2Rs* in sterlet and spotted gars were also high (46 and 49, respectively). The copy number of *V2Rs* in the western clawed frog (691), which far exceeded that previous studies¹⁵, was the largest among vertebrates studied so far; it is much higher than those of other tetrapods, such as caecilian (275), anole lizard (64), and mouse (154). It is noteworthy that we found no *V2Rs* in the genomes of agnathans and amphioxus. According to a recent study by Bi *et al.*²⁸, reports on more than 50 *V2R*-like sequences in hagfish and a few in amphioxus were observed. However, our phylogenetic analyses revealed that they did not form a cluster with known *V2Rs*. Therefore, we need to be still cautious to designate these sequences as *V2Rs*, reaching a conclusion that no typical *V2Rs* exist in agnathans and amphioxus.

Phylogeny and classification of the identified *V2Rs*

To classify and infer the evolutionary history of *V2Rs* in vertebrates from cartilaginous fish to mammals, we constructed a phylogenetic tree of 1897 *V2Rs* identified in this study. The whole vertebrate *V2R* tree (Supplementary File 1) shows the dichotomy between two clades, including “fish-type” and “tetrapod-type” *V2Rs* supported by high BPs (89, 92). Therefore, the phylogenetic trees of “fish-type” (Fig. 1A) and “tetrapod-type” *V2Rs* (Fig. 1B) were shown separately for better visibility in detail classifications. A simplified and overall phylogenetic tree, including all *V2Rs* of “fish-type” and “tetrapod-type” *V2Rs*, and other two additional clades are shown in Fig. 1C.

Previous studies have shown that “fish-type” *V2Rs* were subdivided into 16 subfamilies¹³. In this study, the subfamily memberships of cartilaginous and basal ray-finned fish were consistent with those of previous studies; however, we identified additional subfamilies. Notably, a large number of *V2Rs* from the basal ray-finned fish were expanded in a species-specific manner at the stem of subfamily 4–9 (Fig. 1A). Although these *V2Rs* do not form an exclusive cluster in the tree, we named it subfamily

“a1” for convenience in classification. Over half of the *V2Rs* of the basal ray-finned fish belonged to this subfamily “a1” (Table 1, reedfish 101; sterlet 21; spotted gar 29). Also, we have found a cartilaginous fish-specific subfamily “a2” and “a3” at the stem of subfamily 4–10 and at the stem of subfamily 15, 16 (Table 1, Fig. 1A, Fig. S1). In particular, the large number of *V2Rs* of the cartilaginous fish belonged to subfamily “a2,” in which the genes were expanded in a species-specific manner (Table 1, Supplementary Fig. S1). Taken together, a large fraction of *V2Rs* in cartilaginous and basal ray-finned fish were classified into these new subfamilies “a1”–“a3,” which were the ancestral clades for some known 16 subfamilies. In addition to the subfamily “a1”, species-specific expansion of *V2Rs* in basal ray-finned fish was observed in subfamily two of sterlet (11) and in subfamily 16 of reedfish (67) and spotted gar (9). In contrast to ray-finned fish in which *V2R* subfamilies were highly diverse, cartilaginous fishes were relatively less diverse in that they lacked some subfamilies (Table 1). One exceptional finding was that only one sequence of caecilian *V2R* was classified as being a member of the “fish-type” subfamily 14 (Table 1, Fig. 1A, marked with an asterisk).

The “tetrapod-type” *V2Rs* were shown to be dominated by closely related *V2Rs* that have duplicated more recently than those of cartilaginous and ray-finned fish^{9,13,15}. Notably, *V2Rs* of amphibians (western clawed frog and caecilian) are highly diverse in that they are subdivided into more than 10 and 3 species-specific clusters, while all of the 147 *V2Rs* of mouse belonged to a single cluster. The 63 *V2Rs* of anole lizard were grouped into 2 clusters, which were close to those of mammals (Fig. 1B, Table 1). The 75 *V2Rs* of coelacanth were classified into the “tetrapod-type,” while 13 *V2Rs* were classified into the “fish-type” (Fig. 1A, Table 1). The existence of both the “fish-type” and “tetrapod-type” *V2Rs* in coelacanth was consistent with a previous study¹⁶. One of the most striking results of this study was that some *V2Rs* of ray-finned fishes (three genes from each of reedfish and sterlet) were classified into “tetrapod-type”. Indeed, in the phylogenetic tree (as shown in Fig. 1B and Supplementary File S1), these *V2Rs* are nested within the “tetrapod-type” *V2Rs*. These “tetrapod-type” *V2Rs* of basal ray-finned fish were close to those of coelacanth.

Newly identified *V2R* orthologs conserved from cartilaginous fish to amphibians

Previous studies have shown that *V2Rs* were divided into three well-supported clades, namely, the *V2R2s*, “fish-type” *V2Rs*, and “tetrapod-type” *V2Rs*¹⁶. The orthologous *V2R2s* were shared in all vertebrates from cartilaginous fish to mammals with a few species-specific duplications in mice^{26,29}. In this study, only one *V2R2* ortholog of all species investigated was located at the basal position of the *V2R* tree (Table 1, Fig. 1C, Supplementary Fig. 1). In addition to the three clades, we identified a novel clade, in which *V2Rs* of cartilaginous fish, basal ray-finned fish, and amphibians were included (Fig. 1C). Since this clade contained one *V2R* from each species, and the tree topology was identical to the species tree, these *V2Rs* were considered orthologous. Considering that the origin of these *V2R* orthologs was comparably as old as *V2R2*, they would be evolutionarily distinct from the “fish-type”

and “tetrapod-type” *V2Rs*, which amplify in a species-specific manner. Therefore, we designated them as *anc* (ancient) *V2R*. In this analysis, highly conserved *ancV2R* sequences were found in the genomes of cartilaginous fish to amphibians, but not in teleost fish, coelacanth, and mammals (Fig. 1C).

Conserved gene clusters of the “fish-type” *V2Rs*

In addition to phylogenetic analysis, the synteny relationships have provided important insight into the classification of *V2Rs*. Previous studies have shown that the *V2Rs* in teleost were clustered in one particular chromosomal region, which was flanked by two landmark genes, *phospholipase C (PLC) etal* and *neprilysin*. In contrast, no *V2Rs* were found between these two genes in tetrapods, and “tetrapod-type” *V2Rs* were scattered into several chromosomes^{13,18,19}. Therefore, to ascertain if the *V2Rs* are “fish-type” or “tetrapod-type”, we examined the synteny relationships in vertebrates. Figure 2 shows the gene arrangements in the genomic region of two landmark genes in various vertebrates from cartilaginous fish to mammals. Notably, *V2R2* and *ancV2R* was found in tandem of the neighboring regions of the *PLC etal*, implying the evolutionary conservation of these two genes and evolutionary distinction from known “fish-type” *V2Rs*. Consistent with the previous study, no *V2R* was found in tetrapods in this region, except for only one caecilian *V2R*, which was classified as “fish-type” in the phylogenetic tree (Fig. 1A). In basal ray-finned fish (reedfish, sterlet spotted gar), it was obvious that the *V2Rs* were clustered between *PLC etal* and *neprilysin*. Importantly, we revealed that all “fish-type” *V2Rs* were located in this cluster region, while three “tetrapod-type” *V2Rs* identified in reedfish and sterlet were located on different chromosomes (chr.3 in reedfish, chr.52, 53 and VTUV01000346.1 in sterlet, Supplementary Table S1). In sterlet, we identified two distinct chromosomal regions of the “fish-type” *V2R* clusters, which were due to polyploidization specific to this group¹³. The synteny of coelacanth also showed a conservation of the “fish-type” *V2R* cluster. In the elephant shark and bamboo shark, some “fish-type” *V2Rs* were located outside the cluster, but considering that they were all found in short scaffolds, it is likely that the cluster region was not properly assembled. Overall, the phylogenetic and synteny analyses both supported the conservation of the cluster for “fish-type” *V2Rs* as well as the existence of the “tetrapod-type” *V2Rs* in basal ray-finned fish.

Expression of *V2Rs* in the olfactory epithelium of basal ray-finned fish

To evaluate the functional role of *V2Rs* found in the genomes of basal ray-finned fish, we examined the cellular expression patterns for these receptors in the OE. Figure 3 shows the location of the transcripts detected by *in situ* hybridization on frozen sections of the OE of *Polypterus senegalus* (bichir), which is a basal ray-finned fish closely related to reedfish. The probes of the four *V2Rs* – “fish-type” *V2R*, “tetrapod-type” *V2R*, *V2R2*, and *ancV2R* – each of which has been identified as a distinct clade in the phylogenetic tree (Fig. 1C), were used in the experiments. The expression of a

member of the “fish-type” *V2R* showed a sparse pattern in the sensory cells of the OE, typical of canonical *V2Rs* (Fig. 3A). The expression of a member of “tetrapod-type” *V2R*, which was newly identified in basal ray-finned fish, also showed similar sparse pattern in the OE of *P. senegalus* (Fig. 3B). The expression of *V2R2*, of which the ortholog was highly conserved among jawed vertebrates, has showed widespread patterns in their OE (Fig. 3C). The expression pattern of *V2R2* in *P. senegalus* was consistent with the ubiquitous expression in zebrafish³¹ and mouse²⁹. The expression of *ancV2R*, of which the orthologs were also highly conserved from cartilaginous fish to amphibians, showed a sparse pattern in their OE (Fig. 3D) distinct from that of *V2R2*. Overall, the *V2Rs* belonging to four clades were all expressed in the OE, suggesting their functions as olfactory receptors. However, the patterns of expressions were ubiquitous in *V2R2*, while they were sparse in *V2Rs* of other clades.

Discussion

In this study, we have conducted a comprehensive exploration of *V2R* sequences from the genomes of 19 vertebrate species. Phylogenetic analyses of a large number of *V2Rs* allowed us to gain a panoramic view of the diversity in terms of copy number and repertoire of *V2Rs* across vertebrates. Here, we discuss the tempo and mode of evolution of *V2Rs* and how these factors drive the adaptive evolution of the olfactory system in vertebrates.

It is obvious that *V2Rs* were abundant in ray-finned fish compared to cartilaginous fish, both in terms of copy number and repertoire, which is achieved by a species-specific expansion of “fish-type” *V2Rs*. The synteny analyses revealed that “fish-type” *V2Rs* constituted a large gene cluster between *PLC ϵ* and *neprilysin* in ray-finned fish (Fig. 2). Although the *V2R* clusters were ambiguous in the elephant shark, bamboo shark and saw fish, which may be due to the complex chromosomal rearrangements in cartilaginous fish²¹, the “fish-type” *V2Rs* exist near the two marker genes. Notably, species-specific expansions of *V2Rs* did not occur uniformly in all subfamilies, but were rather concentrated in certain ones. For example, the expansion of *V2Rs* was mainly observed in subfamilies 4–9, 16, ‘a1’, and ‘a2’, while the copy numbers of other subfamilies remained one or two. Therefore, distinction in copy numbers between subfamilies reflects the difference in ligand recognition and biological functions of *V2Rs* among each subfamily. At present, *V2Rs* are expected to detect amino acids and their derivatives, eliciting feeding behaviors in teleost fish^{4,5,32}. It is reasonable to assume that a limited number of amino acids in diets were received by evolutionarily conserved *V2R* subfamilies. However, it is plausible to assume that the *V2R* subfamily with frequent lineage-specific gene duplications is responsible for receiving some species-specific variable chemicals for social communication. For example, Yambe *et al.*³³ showed that an amino acid derivative, L-kynurenine, secreted in the female urine, acts as the male-attracting pheromone in masu salmon. In addition, a previous study showed a possible correlation between expansions of *V2Rs* in subfamily 9 and the evolution of fright reactions in teleost fish⁶. Thus, to elucidate the function of *V2Rs* in ray-finned fish in addition to amino acid reception, it is necessary to further examine the *V2Rs* from a multidisciplinary framework, including the ligand binding, and behavioral experiments using candidate chemicals.

We showed that orthologous sequences of *V2R2* and newly identified *ancV2R* have long been conserved during the evolution of vertebrates. The conservation pattern of orthologs in *V2R2* and *ancV2R* is distinct from the canonical *V2Rs*, such as those in the “fish-type” and “tetrapod-type” *V2Rs*, which were diversified via species-specific gene duplications. The existence of *V2R2* and *ancV2R* in close genomic proximity was also conserved in many vertebrates (Fig. 2). Studies by Silvotti *et al.* in the mouse²⁹ and DeMaria *et al.* in zebrafish³¹ have revealed that highly conserved *V2R2* was expressed in a broad area of the OE and was co-expressed with one of the many canonical *V2Rs*. Consistent with previous studies, *V2R2* of *Polypterus senegalus* showed widespread expression patterns in the OE (Fig.

3C), while the “fish-type” and “tetrapod-type” *V2Rs* showed sparse patterns (Figs. 3A, B). In contrast, although *ancV2R* was similar to *V2R2* in terms of evolutionary conservation and genomic proximity, the pattern of expression in the OE was sparse rather than widespread. Thus, it is implicative to note that *ancV2R* has characteristics between *V2R2* and canonical “fish-type” and/or “tetrapod-type” *V2Rs*. Thus, taking the evolutionary conservation as well as the sparse pattern of expression into account, *ancV2R* may have retained ancestral nature inherited from a protogene before the split between “fish-type” and “tetrapod-type” *V2Rs*, which are now highly diversified in jawed vertebrate genomes. This study showed that all four clades of *V2Rs*, including the “tetrapod-type” *V2Rs*, were expressed in the olfactory organs of basal ray-finned fish, which propose that they have functional roles in olfaction. At the same time, the degree of conservation and expression patterns was distinct among those clades. In particular, it was of interest that the expression patterns of *V2R2*, *ancV2R*, and “fish-type” *V2Rs* were distinct despite their location on the same genomic cluster. A detailed investigation into this genomic region would lead to the elucidation of a *cis*-regulatory mechanism that controls the expression of canonical *V2Rs*, as to say “one neuron one receptor” rule³⁴.

Figure 4 illustrates the evolutionary scenario of the four clades of *V2R* in vertebrates from agnathans (lamprey and hagfish) to mammals by showing the distribution of these four *V2R* clades in extant species and two ancestral nodes, as estimated by the parsimonious principle. No *V2R* was found in the agnathans, implying the acquisition of a *V2R*-mediated olfaction as a common ancestor of jawed vertebrates (cartilaginous fish and Osteichthyes). Given that agnathans have three other types of chemoreceptor genes, namely, *ORs*, *TAARs*, and *VRs*, it is likely that *V2Rs* were originated later than these three receptors during vertebrate evolution^{26,35}. *V2R2* was conserved in all extant vertebrates with no exceptions, which suggests the highly important and fundamental role of *V2R2* in the detection and subsequent signaling pathway for olfactory substances in both underwater and terrestrial environments. Importantly, except for evolutionarily conserved *V2R2*, mammal and teleost fish possessed only specific clades of *V2Rs*, namely, “tetrapod-type” and “fish-type” *V2Rs*, respectively. In contrast, basal ray-finned fish and amphibians possessed both “tetrapod-type” *V2Rs* and “fish-type” *V2Rs*. This finding suggests that all four clades of *V2Rs* were present at least in the common ancestor of Osteichthyes, or even earlier in the common ancestor of jawed vertebrates. However, subsequently, the repertoires of *V2Rs* were lost in teleost fish and mammals during adaptation to their specific environments. Because the previous studies of *V2Rs* have been limited to teleost fish and mammalians, the diversity of *V2Rs* that should exist in the other variety of species has not been recognized. Thus, the genome sequences of non-model animals such as cartilaginous fish, basal ray-finned fish, coelacanth, which become available due to the recent advancement of sequence technique, allowed us to uncover the hidden and unexpectedly large diversity in the genomes of vertebrates.

One of the worthwhile discussions is the possible great diversity of *V2Rs* in the Osteichthyes ancestor. The repertoires of the subfamily of the “fish-type” *V2Rs* are abundant in ray-finned fish and

coelacanth compared with those in cartilaginous fish (Table 1). In addition, all four clades of *V2Rs* were shown to be present in basal ray-finned fish and amphibians. The abundance of copy numbers and repertoires in these groups show that *V2Rs* were highly diversified in the common ancestor of Osteichthyes. Given that some of the olfaction-related genes also emerge in the common ancestor of the Osteichthyes (e.g., *ancVIR*³⁶; *OMP*³⁷), it might be possible that an innovative evolution of the olfactory system occurred in this timing. It is worth mentioning here that the polypterids (bichir and reedfish) possesses large paired openings (spiracles) on top of their head, in which they use for breathing air³⁸. Similar spiracle-like structures were observed in the fossil records of stem tetrapods³⁹. Thus, breathing air using spiracles may have been an important respiratory strategy in the stem Osteichthyes, which inhabit shallow freshwater environments and use lungs in addition to gills for respiration⁴⁰. Specifically, the evolution of air-breathing by spiracles may increase the opportunity to raise their head above water, which led to the acquisition of the primitive capabilities of detecting airborne chemicals before terrestrial adaptation. Thus, such dual functional roles of the olfactory system in stem Osteichthyes were related to the diversification of *V2Rs*, including “tetrapod-type” *V2Rs*. Deorphanization of *V2Rs* of various vertebrates in the near future is necessary to evaluate the above possibility.

Materials and methods

Sequence retrieval

To estimate the evolutionary history of *V2Rs*, we conducted a comprehensive exploration of *V2R* sequences in the genome assemblies of a broad range of vertebrates, including nine cartilaginous fishes and three basal ray-finned fishes. In addition to two teleost fishes (Japanese eel and zebrafish), coelacanth, two amphibians (caecilian and western clawed frog), anole lizard, and mouse were explored (Table 1). To identify the *V2R* sequences from the genome assemblies, we performed tBLASTn searches using the transmembrane (TM) domain of *V2Rs* as queries against all of the genomic sequences. The query sequences were generated by aligning *V2R* sequences of mouse, anole lizard, tropical clawed frog, coelacanth, spotted gar, and zebrafish deposited in the Ensembl database and of elephant shark and cloudy catshark²² using the MA-T --dash option⁴¹. Then, the TM domains, from 50-aa sequences upstream of the 1st TM region to 50-aa downstream of the 7th TM region, were extracted using a protein structure of the glutamate receptor (PDB: 4OR2)⁴² as reference.

Next, the *V2R*-coding sequences were predicted for tBLASTn hit regions using GeneWise v.2.4.1⁴³. Sequences shorter than 600 nucleotides were discarded. The homology of intact outputs was then determined by the next phylogenetic analysis. The queries used in the BLAST search were critical to the estimated number of *V2Rs*. Indeed, the number of *V2Rs* identified in this study was consistent with previous studies, whereas a few differences were observed in some cases. For example, the numbers of *V2Rs* in elephant shark and catshark identified in this study (29 and 24 copies) were slightly smaller than those reported by Sharma *et al.*²². The difference in query sequences used for the BLAST search reflects results. Specifically, we used the seven-transmembrane regions of the *V2R*, whereas Sharma *et al.*²² used the entire region of *V2R*, including the extracellular Venus flytrap module region (VFTM)⁴⁴. The VFTM region of the *V2R* protein has been determined to be highly diverse and difficult in predicting the exon-intron structure of the gene. This result affects the judgment of intact or pseudogenes. Therefore, we compared the number of *V2Rs* estimated under a unified condition using the seven-transmembrane region as a query.

Phylogenetic and synteny analyses

Sequences were translated into amino acid sequences and aligned using MAFFT with the -ginsi option⁴⁵ and default parameters. The sites with <50 % coverage among all sequences were removed. Maximum likelihood trees for *V2R* genes were then inferred using RAXML-NG v.1.0.1⁴⁶ with the JTT+G+F amino acid substitution model. This result was the best fitting model selected by the ModelTest-NG^{47,48} based on AIC scores. Rapid bootstrap analyses were performed using 100 replicates to assess the reliability of nodes. The *V2R* sequences for coelacanth, western clawed frog, and mouse were then clustered based on the criteria of 80 % similarity at the amino acid level to save computational costs. To avoid false positives in identification of the *V2R* genes and prevent long

branch attractions, we obtained other GPCR family C sequences from GenBank or Ensembl for use as outgroups, including CaSR (NM_013803.3), Tas1r1 to Tas1r3 (ENSMUSG00000028950, ENSMUSG00000028738, ENSMUSG00000029072), GPCR6 (NM_153071.1), GRM1a to GRM8 (NM_016976.1, NM_001160353.1, NM_181850.2, NM_001081414.2, NM_001081414.2, NM_173372.2, NM_177328.3, NM_001361125.1), and GABA B1 to GABA B2 (NM_019439.3 and NM_001081141.2), for constructing an initial gene tree. All genes in the sister clade to *CaSR* were named homologs of *V2R* (including *V2R2*). Using these genes, we constructed the *V2R* gene tree again (*CaSR* was used as an outgroup). We also included sequences of all 16 teleost fish subfamilies classified in previous studies^{13,14}, as markers to indicate *V2R* subfamilies.

The synteny relationships of “fish-type” *V2Rs* of vertebrates were then illustrated based on the genomic location of the identified *V2Rs* (Supplementary Table S1). The numeric data for the genomic position of *V2Rs*, which were identified as “fish-type” *V2Rs*, *V2R2*, and *ancV2R* in the phylogenetic tree, were then compiled, with those of *PLC ϵ 1* and *neprilysin*.

***In situ* hybridization**

Bichir (*Polypterus senegalus*) individuals of 11–25 cm long, which were used for the preparation of frozen sections and the extraction of total RNA from the olfactory organs, were purchased from a commercial supplier; they were kept under standard conditions suitable for tropical fish breeding until experimental manipulations. All experimental studies using the animals were approved by the Institutional Animal Experiment Committee of the Tokyo Institute of Technology were performed in accordance with the institutional, governmental ARRIVE guidelines. TRIzol (Invitrogen) was then used for total RNA extraction from the olfactory organs of the bichir. Using the total RNA extracted from the olfactory organs of the bichir, cDNA was synthesized by reverse transcription reaction using SuperScript III RTase (Invitrogen). Coding regions of *V2R* were amplified by PCR using the primer sets, which were designed on the basis of *V2R* sequences of reedfish, as has been summarized in Supplementary Table S2. The PCR products were cloned using the pGEM-T vector (Promega) and the DH5 α strain of *E. coli*. Digoxigenin-labeled RNA probes were synthesized using the plasmid vector as a template using T7 or SP6 RNA Polymerase (Roche) and DIG RNA labeling mix (Roche) as well. The olfactory organs of the bichir were then fixed with 4 % PFA, replaced with sucrose, and embedded in an O.C.T compound (Sakura Finetek). *In situ* hybridization was performed according to the method as previously described^{36,37}. Briefly, hybridization was performed using DIG-labeled RNA probes. The antibody reaction was conducted using anti-digoxigenin-POD, Fab fragments (Sigma-Aldrich), or anti-fluorescein-POD, Fab fragments (PerkinElmer). The signal was then amplified with Tyramide Signal Amplification Plus Biotin kit (Kiko Tech), and detected by streptavidin, Alexa Fluor 488 conjugate (Thermo Fisher). Finally, the sections were sealed using a VECTASHIELD Mounting Medium with DAPI (VECTOR). The sealed sections were observed using a fluorescence microscope

376 Axioplan (Carl Zeiss). All fluorescence photographs were taken using an AxioCam 503 color (Carl
377 Zeiss) and optimized for brightness and contrast in Adobe Photoshop.
378
379

Acknowledgments

This work was funded by JSPS KAKENHI (20H03307) and Asahi Glass Foundation to M.N. We thank the researchers, who determined and assembled the genome sequences of cartilaginous fish and reedfish as part of the Vertebrate Genome Project (VGP), for releasing their unpublished data.

Author contribution

Project design, coordination: M.N.; bioinformatics and evolutionary analyses: Z.Z.; molecular experiments: A.S.; manuscript writing: Z.Z., S. K., M.N.

Competing interests

The authors declare no competing interests

Additional information

The online version contains supplementary materials available at XXXX.

Correspondence and request for the materials should be addressed to M.N.

Literature Cited

1. Nei, M., Niimura, Y. & Nozawa, M. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nat. Rev. Genet.* **9**, 951–963 (2008).
2. Alioto, T. S. & Ngai, J. The repertoire of olfactory C family G protein-coupled receptors in zebrafish: candidate chemosensory receptors for amino acids. *BMC Genomics* **7**, 309 (2006).
3. Johnstone, K. A. *et al.* Genomic organization and evolution of the vomeronasal type 2 receptor-like (OlfC) gene clusters in Atlantic salmon, *Salmo salar*. *Mol. Biol. Evol.* **26**, 1117–1125 (2009).
4. Specca, D. J. *et al.* Functional identification of a goldfish odorant receptor. *Neuron* **23**, 487–498 (1999).
5. Koide, T. *et al.* Olfactory neural circuitry for attraction to amino acids revealed by transposon-mediated gene trap approach in zebrafish. *Proc. Natl. Acad. Sci. USA* **106**, 9884–9889 (2009).
6. Yang, L. *et al.* Expansion of vomeronasal receptor genes (OlfC) in the evolution of fright reaction in Ostariophysan fishes. *Commun. Biol.* **2**, 235 (2019).
7. Matsunami, H. & Buck, L. B. A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell* **90**, 775–784 (1997).
8. Hagino-Yamagishi, K. *et al.* Expression of vomeronasal receptor genes in *Xenopus laevis*. *J. Comp. Neurol.* **472**, 246–256 (2004).
9. Brykczynska, U., Tzika, A. C., Rodriguez, I. & Milinkovitch, M. C. Contrasted evolution of the vomeronasal receptor repertoires in mammals and squamate reptiles. *Genome Biol. Evol.* **5**, 389–401 (2013).
10. Kimoto, H., Haga, S., Sato, K. & Touhara, K. Sex-specific peptides from exocrine glands stimulate mouse vomeronasal sensory neurons. *Nature* **437**, 898–901 (2005).
11. Haga, S. *et al.* The male mouse pheromone ESP1 enhances female sexual receptive behaviour through a specific vomeronasal receptor. *Nature* **466**, 118–122 (2010).
12. Leinders-Zufall, T. *et al.* MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science* **306**, 1033–1037 (2004).
13. Hashiguchi, Y. & Nishida, M. Evolution and origin of vomeronasal-type odorant receptor gene repertoire in fishes. *BMC Evol. Biol.* **6**, 76 (2006).
14. Nikaido, M. *et al.* Lineage-specific expansion of vomeronasal type 2 receptor-like (OlfC) genes in cichlids may contribute to diversification of amino acid detection systems. *Genome Biol. Evol.* **5**, 711–722 (2013).
15. Shi, P. & Zhang, J. Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. *Genome Res.* **17**, 166–174 (2007).
16. Picone, B. *et al.* Taste and odorant receptors of the coelacanth—a gene repertoire in transition. *J. Exp. Zool. B Mol. Dev. Evol.* **322**, 403–414 (2014).

17. Nikaido, M. Evolution of *VIR* pheromone receptor genes in vertebrates: diversity and commonality. *Genes Genet. Syst.* **94**, 141–149 (2019).
18. Silva, L. & Antunes, A. Vomeronasal Receptors in Vertebrates and the Evolution of Pheromone Detection. *Annu. Rev. Anim. Biosci.* **5**, 353–370 (2017).
19. Rodriguez, I., Del Punta, K., Rothman, A., Ishii, T. & Mombaerts, P. Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors. *Nat. Neurosci.* **5**, 134–140 (2002).
20. Nikaido, M. *et al.* Coelacanth genomes reveal signatures for evolutionary transition from water to land. *Genome Res.* **23**, 1740–1748 (2013).
21. Hara, Y. *et al.* Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat. Ecol. Evol.* **2**, 1761–1771 (2018).
22. Sharma, K., Syed, A. S., Ferrando, S., Mazan, S. & Korsching, S. I. The chemosensory receptor repertoire of a true shark is dominated by a single olfactory receptor family. *Genome Biol. Evol.* **11**, 398–405 (2019).
23. Marra, N. J. *et al.* White shark genome reveals ancient elasmobranch adaptations associated with wound healing and the maintenance of genome stability. *Proc. Natl. Acad. Sci. USA* **116**, 4446–4455 (2019).
24. Theisen, B., Zeiske, E. & Breucker, H. Functional Morphology of the Olfactory Organs in the Spiny Dogfish (*Squalus acanthias* L.) and the Small-spotted Catshark (*Scyliorhinus canicula* (L.)) *Acta Zoologica* **67**, 73–86 (1986).
25. Ferrando, S. *et al.* Immunolocalization of G-protein alpha subunits in the olfactory system of the cartilaginous fish *Scyliorhinus canicula*. *Anat. Rec.* **292**, 1771–1779 (2009).
26. Grus, W. E. & Zhang, J. Origin of the genetic components of the vomeronasal system in the common ancestor of all extant vertebrates. *Mol. Biol. Evol.* **26**, 407–419 (2009).
27. Moriya-Ito, K., Hayakawa, T., Suzuki, H., Hagino-Yamagishi, K. & Nikaido, M. Evolution of vomeronasal receptor 1 (V1R) genes in the common marmoset (*Callithrix jacchus*). *Gene* **642**, 343–353 (2018).
28. Bi, X. *et al.* Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. *Cell* **184**, 1377–1391 (2021).
29. Silvotti, L., Moiani, A., Gatti, R. & Tirindelli, R. Combinatorial co-expression of pheromone receptors. *J. Neurochem.* **103**, 1753–1763 (2007).
30. Du, K. *et al.* The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization. *Nat. Ecol. Evol.* **4**, 841–852 (2020).
31. DeMaria, S. *et al.* Role of a ubiquitously expressed receptor in the vertebrate olfactory system. *J. Neurosci.* **33**, 15235–15247 (2013).
32. Sato, Y., Miyasaka, N. & Yoshihara, Y. Mutually exclusive glomerular innervation by two distinct

- types of olfactory sensory neurons revealed in transgenic zebrafish. *J. Neurosci.* **25**, 4889–4897 (2005).
33. Yambe, H. *et al.* L-Kynurenine, an amino acid identified as a sex pheromone in the urine of ovulated female masu salmon. *Proc. Natl. Acad. Sci. USA* **103**, 15370–15374 (2006).
34. Chess, A., Simon, I., Cedar, H. & Axel, R. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**, 823–834 (1994).
35. Poncelet, G. & Shimeld, S. M. The evolutionary origins of the vertebrate olfactory system. *Open Biol.* **10**, 200330 (2020).
36. Suzuki, H. *et al.* A single pheromone receptor gene conserved across 400 million years of vertebrate evolution. *Mol. Biol. Evol.* **35**, 2928–2939 (2018).
37. Suzuki, H., Nikaido, M., Hagino-Yamagishi, K. & Okada, N. Distinct functions of two olfactory marker protein genes derived from teleost-specific whole genome duplication. *BMC Evol. Biol.* **15**, 245 (2015).
38. Graham, J. B. *et al.* Spiracular air breathing in polypterid fishes and its implications for aerial respiration in stem tetrapods. *Nat. Commun.* **5**, 3022 (2014).
39. Clack, J. A. Devonian climate change, breathing, and the origin of the tetrapod stem group. *Integr. Comp. Biol.* **47**, 510–523 (2007).
40. Graham, J. B. Air-breathing fishes: evolution, diversity, and adaptation. (Academic Press, 1997).
41. Rozewicki, J., Li, S., Amada, K. M., Standley, D. M. & Katoh, K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* **47**, W5–W10 (2019).
42. Wu, H. *et al.* Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **344**, 58–64 (2014).
43. Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004).
44. O'Hara, P. J. *et al.* The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron* **11**, 41–52 (1993).
45. Katoh, K., Kuma, K. I., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
46. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
47. Flouri, T. *et al.* The Phylogenetic Likelihood Library. *Syst. Biol.* **64**, 356–362 (2014).
48. Darriba, D. *et al.* ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* **37**, 291–294 (2020).

Figure legends

Figure 1. Phylogenetic relationships of *V2Rs* of a broad range of vertebrates from cartilaginous fish to tetrapods. (A) Phylogenetic tree of “fish-type” *V2Rs*. Note that “fish-type” *V2Rs* were subdivided into 16 known and 3 novel subfamilies, as indicated by gray thick bars. Triangles in red, orange, blue, white, and brown indicate the expanded *V2R* clusters specific to reedfish, sterlet, spotted gar, cartilaginous fish, and teleost fish, respectively. Only one “fish-type” *V2R* found in the caecilian species was marked using an asterisk. (B) Phylogenetic tree of “tetrapod-type” *V2Rs*. Triangles in violet, pink, yellow, green, and gray indicate expanded *V2R* clusters specific to coelacanth, caecilian, western clawed frog, anole lizard, and mouse, respectively. Asterisks were used to mark the “tetrapod-type” *V2Rs* identified in reedfish and sterlet. Note that the “tetrapod-type” *V2Rs*, in contrast to the “fish-type” *V2Rs*, are composed of many clusters that are expanded in a species-specific manner. (C) Overview of the phylogenetic tree of all *V2Rs* showing novel orthologous clade *ancV2R*. The calcium-sensing receptor (*CaSR*) gene was used for outgrouping all *V2Rs*. The OTU names consist of the common name and locus as summarized in Supplementary Table S1. The “fish-type” *V2Rs*, “tetrapod-type” *V2Rs*, and *V2R2* clades were compressed into black triangles. The number on the branches indicates the bootstrap support values for particular nodes. Note that the grouping of the orthologous *ancV2Rs* of cartilaginous fish, basal ray-finned fish, and western clawed frog was suggested by maximum bootstrap support (100 %). The numbers next to triangles indicate the copy number of *V2Rs* included in the clusters. The filled circles on each node indicate bootstrap supports (black > 75, 75 ≥ gray ≥ 50, 50 > white). Scale bar indicates the number of amino acid substitutions per site.

Figure 2. Synteny relationships for the “fish-type” *V2R* clusters among basal ray-finned fish (reedfish, sterlet, spotted gar), teleost fish (zebrafish), lobe-finned fish (coelacanth), tetrapods (caecilian, western clawed frog, mouse), and cartilaginous fish (elephant shark, bamboo shark, and smalltooth sawfish), respectively. Triangles in yellow, blue, red, and green indicate two landmark genes (*PLC ϵ 1*, *neprilysin*), *V2R2*, and *ancV2R*, respectively. Black squares indicate “fish-type” *V2Rs*. Indicated at the upper-left of each line were several chromosomes or scaffolds and its directions. Indicated below the ends of the line are the start and end of the cluster regions. Unplaced scaffolds are not shown in this figure. Note that “fish-type” *V2Rs* were flanked by two landmark genes and that *V2R2*s and *ancV2Rs* were located in tandem close to the clusters. No “tetrapod-type” *V2Rs* were observed in these cluster regions. In the elephant shark, some “fish-type” *V2Rs* are located outside the cluster because the cluster regions were not properly assembled.

Figure 3. The expression patterns of *V2Rs* in the olfactory epithelium of the basal ray-finned fish, *Polypterus senegalus*. Expression of four *V2R* genes were confirmed by FISH on horizontal sections of the olfactory organ of the bichir using DIG-labeled RNA antisense probes for “fish-type” *V2R* (A)

and “tetrapod-type” *V2R* (B), *V2R2* (C), and *ancV2R* (D). Green indicated the expression signals. The blue area indicates the cell nucleus stained with DAPI. *V2R2* was globally expressed in the deep layers of olfactory folds (C). In contrast, “fish-type” *V2R*, “tetrapod-type” *V2R*, and *ancV2R* were sparsely expressed in a small number of neurons in the deeper layers of the olfactory folds (A, B, and D). (A'–D') High magnification view of the dotted squares in A–D. Scale bars show 100 μ M (A–D) and 20 μ M (A'–D').

Figure 4. Evolutionary scenario of *V2Rs* during vertebrate evolution. The presence/absence of the four major *V2R* clades was plotted on the phylogenetic tree of vertebrates from agnathans to mammals. The red circle with “2,” yellow with “a,” gray with “t,” and blue with “f” indicate *V2R2*, *ancV2R*, “tetrapod-type” *V2R*, and “fish-type” *V2R*, respectively. In contrast to basal ray-finned fish with all four clades of *V2Rs*, teleost fish, mammals, and lizards were determined to possess only two of them. The reduction of specific *V2R* clades in these lineages would be due to adaptation to specific oceanic and terrestrial environments. Note that the origin of “tetrapod-type” *V2Rs* dates back to the era of the common ancestor of extant Osteichthyes, but its antiquity in the jawed vertebrate ancestor remains to be examined with complete genome sequences of more cartilaginous fish (dotted circle with “t” inside).

Table 1. Number of intact *V2R* genes identified in the genomes of a broad range of vertebrates.

Supplementary Table S1. Summary of the intact *V2Rs* in the genomes of vertebrates. Gene name, locus (position in the chromosomes, scaffolds), and directions were summarized in a bed file format.

Supplementary Table S2. PCR primers used to amplify four *V2R* genes of *Polypterus senegalus*.

Supplementary Figure S1. A phylogenetic tree of subfamily “a2” and the cartilaginous fish-specific *V2Rs*. Note that this subfamily consists of 132 *V2Rs*, most of which were expanded in species-specific birth and death processes. Only one orthologous *V2R* was shared among cartilaginous fish (indicated by asterisk).

Supplementary File 1. The Newick format phylogenetic tree of all 1897 intact *V2Rs* identified by a broad range of vertebrates. Note that grouping of “fish-type” *V2Rs*, “tetrapod-type” *V2Rs*, *V2R2*, and *ancV2R* was supported in sufficient bootstrap values.

575 **Supplementary File 2.** The aligned *V2R* sequences of all 1897 intact *V2Rs* identified from various
576 vertebrates.

Table 1. Number of intact *V2R* genes identified in the genomes of broad range of vertebrates.

			V2R2	ancV2R	tetrapod-type V2R	fish-type V2R																	
Scientific name	Common name	Assembly name	sum	sum	sum	sum	sub 1	sub 2	sub 3	sub 4~9	sub 10	sub 11	sub 12	sub 13	sub 14	sub 15	sub 16	sub a1	sub a2	sub a3	unplaced		
Callorhinchus milii	Elephant shark	Callorhinchus_milii-6.1.3	1	1	0	26	1	0	0	0	0	0	1	0	1	0	0	0	19	2	2		
Hydrolagus affinis	Small-eyed rabbitfish	UP_Haf	1	1	0	39	0	0	0	0	0	0	0	0	1	0	0	0	36	0	2		
Scyliorhinus torazame	Cloudy catshark	Storazame_v1.0	1	1	0	21	0	0	1	0	0	0	1	4	1	0	0	0	13	1	0		
Carcharodon carcharias	Great white shark	ASM360424v1	1	1	0	12	2	0	1	0	0	0	0	0	0	0	0	0	8	0	1		
Rhincodon typus	Whale shark	ASM164234v2	1	0	0	21	2	0	1	0	0	0	0	1	0	0	0	0	16	0	1		
Chiloscyllium punctatum	Brownbanded bambooshark	Cpunctatum_v2.1	1	1	0	24	1	0	1	0	0	0	0	2	1	0	0	0	18	1	0		
Leucoraja erinacea	Little skate	LER_WGS_1	1	1	0	6	1	0	1	0	0	0	0	1	0	0	0	0	2	1	0		
Amblyraja radiata	Thorny skate	sAmbRad1.pri	1	1	0	6	1	0	1	0	0	0	0	1	0	0	0	0	2	1	0		
Pristis pectinata	Smalltooth sawfish	sPriPec2.pri	1	1	0	22	2	0	1	0	0	0	0	0	0	0	0	0	18	1	0		
Erpetoichthys calabaricus	Reedfish	fErpCal1.1	1	1	3	183	1	2	1	0	4	1	1	2	1	1	67	101	0	0	1		
Acipenser ruthenus	Sterlet	ASM1064508v1	1	1	3	41	1	11	1	0	1	2	1	0	1	0	1	21	0	0	1		
Lepisosteus oculatus	Spotted gar	LepOcu1	1	1	0	47	0	1	2	0	1	0	1	1	1	1	9	29	0	0	1		
Anguilla japonica	Japanese eel	Ajp_01	1	0	0	63	1	4	2	39	1	4	1	2	1	3	5	0	0	0	0		
Danio rerio	Zebrafish	GRCz11	1	0	0	71	2	6	1	24	1	3	1	6	2	2	23	0	0	0	0		
Latimeria chalumnae	Coelacanth	LatCha1	1	0	75	13	1	1	1	0	0	1	4	1	1	0	0	0	0	0	3		
Xenopus tropicalis	Western clawed frog	UCB_Xtro_10.0	1	1	689	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Geotrypetes seraphini	Cacilian	aGeoSer1.1	1	0	273	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0		
Anolis carolinensis	Anole lizard	AnoCar2.0	1	0	63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Mus musculus	Mouse	GRCm38.p6	7	0	147	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		

Note: The gray highlighted the cells in which no genes were found in the corresponding subfamilies.

Fig. 1(A)

Fig. 1(B)

Fig. 1(C)

Fig. 2

Fig. 3

Fig. 4

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.pdf](#)