



**HAL**  
open science

## Coevolution of the olfactory organ and its receptor repertoire in ray-finned fishes

Maxime Policarpo, Katherine Bemis, Patrick Laurenti, Laurent Legendre, Jean-Christophe Sandoz, Sylvie Rétaux, Didier Casane

► **To cite this version:**

Maxime Policarpo, Katherine Bemis, Patrick Laurenti, Laurent Legendre, Jean-Christophe Sandoz, et al.. Coevolution of the olfactory organ and its receptor repertoire in ray-finned fishes. BMC Biology, 2022, 20 (1), pp.195. 10.1186/s12915-022-01397-x . hal-03773266

**HAL Id: hal-03773266**

**<https://cnrs.hal.science/hal-03773266>**

Submitted on 9 Nov 2022

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

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

RESEARCH ARTICLE

Open Access



# Coevolution of the olfactory organ and its receptor repertoire in ray-finned fishes

Maxime Policarpo<sup>1†</sup>, Katherine E. Bemis<sup>2†</sup>, Patrick Laurenti<sup>3</sup>, Laurent Legendre<sup>1</sup>, Jean-Christophe Sandoz<sup>1</sup> , Sylvie Rétaux<sup>4\*</sup>  and Didier Casane<sup>1,5\*</sup> 

## Abstract

**Background:** Ray-finned fishes (Actinopterygii) perceive their environment through a range of sensory modalities, including olfaction. Anatomical diversity of the olfactory organ suggests that olfaction is differentially important among species. To explore this topic, we studied the evolutionary dynamics of the four main gene families (OR, TAAR, ORA/VR1 and OlfC/VR2) coding for olfactory receptors in 185 species of ray-finned fishes.

**Results:** The large variation in the number of functional genes, between 28 in the ocean sunfish *Mola mola* and 1317 in the reedfish *Erpetoichthys calabaricus*, is the result of parallel expansions and contractions of the four main gene families. Several ancient and independent simplifications of the olfactory organ are associated with massive gene losses. In contrast, Polypteriformes, which have a unique and complex olfactory organ, have almost twice as many olfactory receptor genes as any other ray-finned fish.

**Conclusions:** We document a functional link between morphology of the olfactory organ and richness of the olfactory receptor repertoire. Further, our results demonstrate that the genomic underpinning of olfaction in ray-finned fishes is heterogeneous and presents a dynamic pattern of evolutionary expansions, simplifications, and reacquisitions.

**Keywords:** Olfactory epithelium, Olfactory lamellae, Olfactory receptor genes, Actinopterygii, Gene family dynamics

## Background

With more than 34,000 valid species, Actinopterygii (ray-finned fishes) is the largest group of aquatic vertebrates [1]. Most species of ray-finned fishes belong to Teleostei (teleosts), but a few extant species belong to relictual clades: Polypteriformes, Acipenseriformes, Lepisosteiformes, and Amiiformes (Fig. 1). With a last common ancestor that lived 368–379 million years ago (Ma) [2, 3], the remarkable taxonomic diversity of

actinopterygians comes with striking anatomical, physiological, behavioral, and ecological adaptations [4]. Actinopterygians thrive in aquatic habitats from the tropics to the polar regions, in small temporary ponds to large oceans.

Ray-finned fishes have several sensory systems to process physical and chemical cues. Among them, the olfactory system serves in feeding, reproduction, predator avoidance and migration [5]. A seminal work described the great anatomical diversity in the olfactory organs of ray-finned fishes [6]. Since then, it has been assumed that fishes with a multilamellar olfactory epithelium have a better sense of smell than those with a flat olfactory epithelium, respectively classified as macrosmatic and microsomatic [7].

In most ray-finned fishes, the olfactory epithelium forms a rosette in which lamellae attach to a central

<sup>†</sup>Maxime Policarpo and Katherine E. Bemis participated equally to this work.

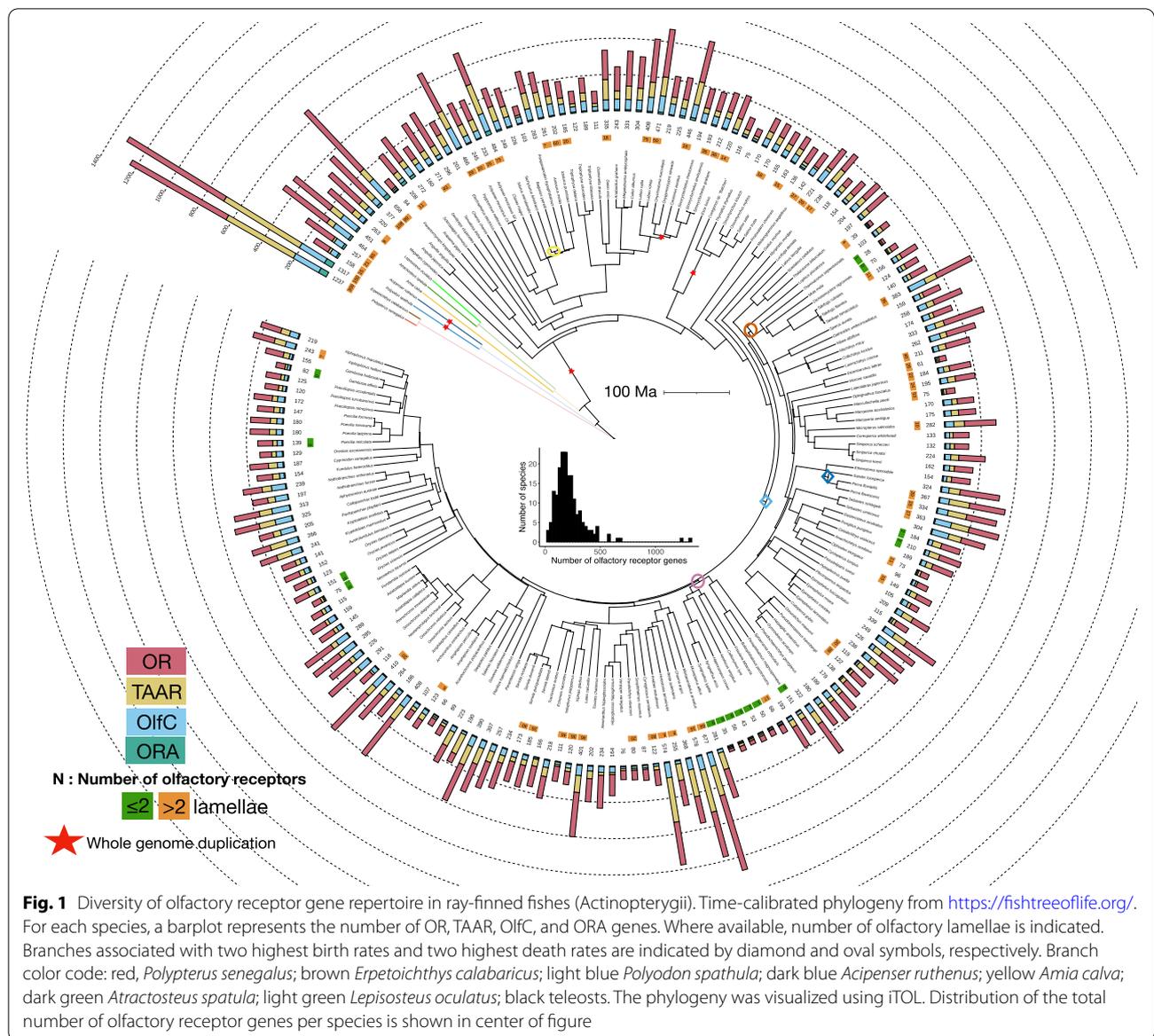
\*Correspondence: sylvie.retaux@cnrns.fr; didier.casane@egce.cnrns-gif.fr

<sup>4</sup> Université Paris-Saclay, CNRS, Institut des Neurosciences Paris-Saclay, 91400, Saclay, France

<sup>5</sup> Université Paris Cité, UFR Sciences du Vivant, 75013 Paris, France

Full list of author information is available at the end of the article

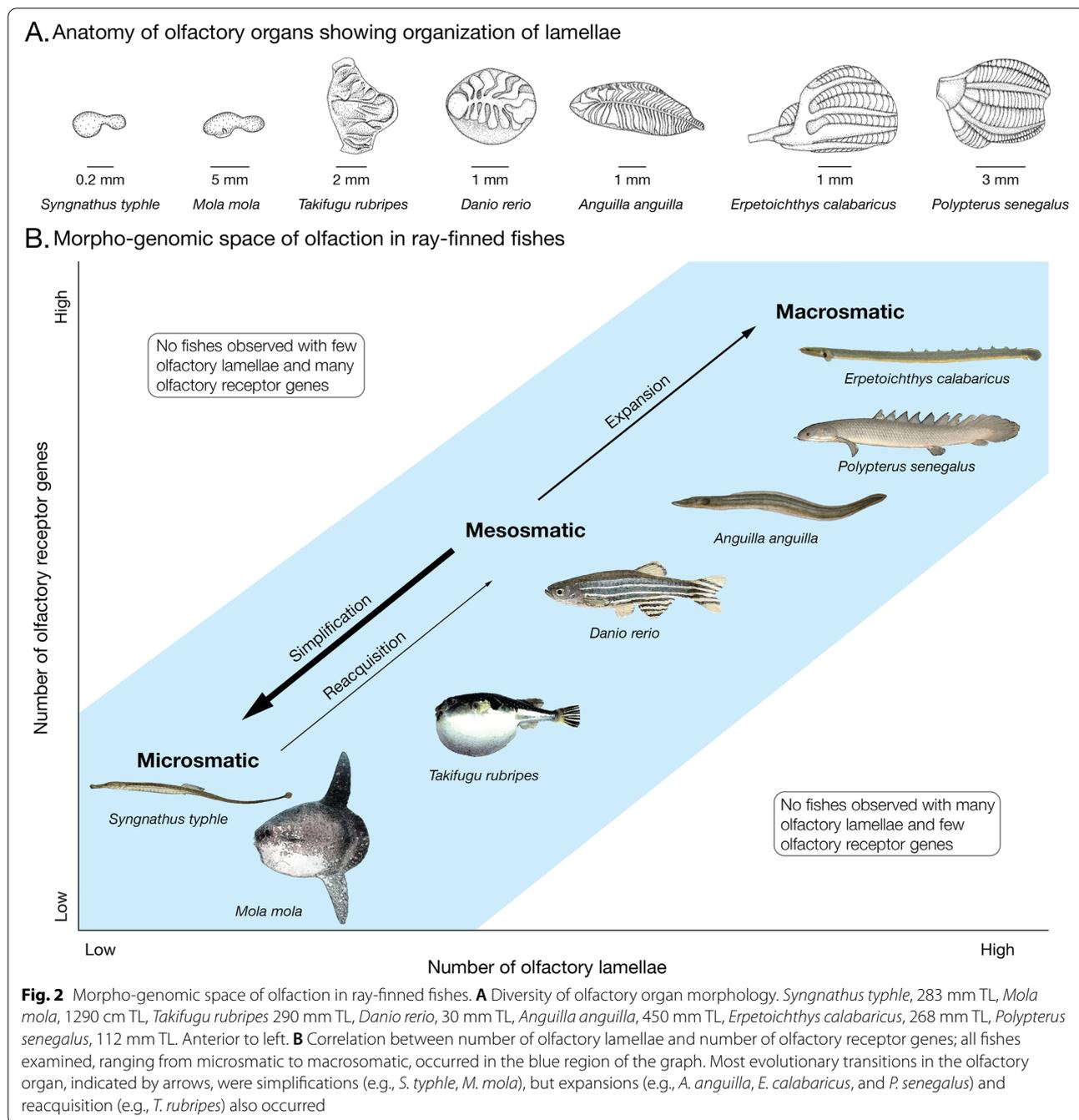




raphe (e.g., *Danio rerio* in Fig. 2A). Chondrichthyans (sharks, rays, chimeras) also have olfactory rosettes [8]; thus, it is likely that olfactory rosettes were present in the common ancestor of jawed vertebrates and conserved in the common ancestor of ray-finned fishes. However, the olfactory rosette has been simplified several times during the evolution of ray-finned fishes, leading in the most extreme cases to a small, flat olfactory epithelium with no lamellae (e.g., *Syngnathus typhle* in Fig. 2A) [9, 10]. In contrast, other groups have multilamellar organizations of the olfactory epithelium. The most extreme example of a multilamellar olfactory epithelium occurs in the Polypteriformes, which have a large and complex structure: a nasal capsule is

divided into six sectors, five in a main sac and one in a diverticulum, each with a rosette-like organization with a septum and lamellae attached to both sides (e.g., *Polypterus senegalus* and *Erpetoichthys calabaricus* in Fig. 2A) [11, 12].

The diversity of odorants that can be detected depends on the size of the olfactory receptor gene repertoire [13]. In vertebrates, olfactory receptor genes belong to four main gene families with independent origins: odorant receptors (OR), trace amine-associated receptors (TAAR), and vomeronasal receptors 1 and 2 (named VR1 and VR2 in tetrapods). Actinopterygian fishes do not have a vomeronasal organ, and thus VR1 and VR2 gene families are referred to as ORA and



OlfC, respectively. Only a few olfactory receptor genes have been identified that do not belong to these four gene families [14].

Analyses of genomes of 13 teleosts and one non-teleost suggested that ORA is a small and stable gene family, with eight genes in the last common ancestor of ray-finned fishes and six genes in most teleosts [15]. More genes have been identified in OR, TAAR, and OlfC gene families [16–19]. The evolution of two families (OR and

TAAR) were analyzed separately using broad samples of ray-finned fishes [10, 20]. Both studies showed that olfactory receptor gene families are dynamic, for example, there is a ~30-fold variation in the number of OR genes (15 in ocean sunfish *Mola mola* and broad-nose pipefish *Syngnathus typhle* to 429 in Zig-zag Eel *Mastacembelus armatus*). Moreover, the number of olfactory lamellae was correlated with the richness of the OR gene repertoire [10].

A recent burst of highly complete (in terms of coding sequences), publicly available genomes for ray-finned fishes, in particular for non-teleost actinopterygians such as Polypteriformes, prompted us to analyze the evolution of the four olfactory gene families and the anatomy of the olfactory organ across the phylogeny of ray-finned fishes.

## Results and discussion

### Coevolution of olfactory receptor family sizes

We characterized the olfactory receptor gene repertoire, including OR, TAAR, OlfC, and ORA genes, for 185 species of ray-finned fishes selected based on high genome completeness (Fig. 1, Additional file 1: Fig. S1, Additional file 2: Supplementary Data 1).

The mean size of the total olfactory gene repertoire for actinopterygians was 224 genes. The largest (1317 genes) was found in the polypteriform *Erpetoichthys calabaricus* (reedfish) and the smallest (28 genes) in the tetraodontiform *Mola mola* (ocean sunfish) (Fig. 1). ORA is a small and stable family typically comprising six genes (ORA1 to ORA6) in teleosts [15]. Nevertheless, we found up to three ORA genes have been lost in several lineages, and, surprisingly, this gene family is much larger in some lineages, particularly Polypteriformes, which have nearly 50 functional ORA genes (Fig. 1, Additional file 1: Fig. S1A). Two genes, ORA7 and ORA8, were present in the last common ancestor of ray-finned fishes; ORA7 was lost in the common ancestor of teleosts, while ORA8 was lost in clupecocephalans [15] (Additional file 1: Fig. S2).

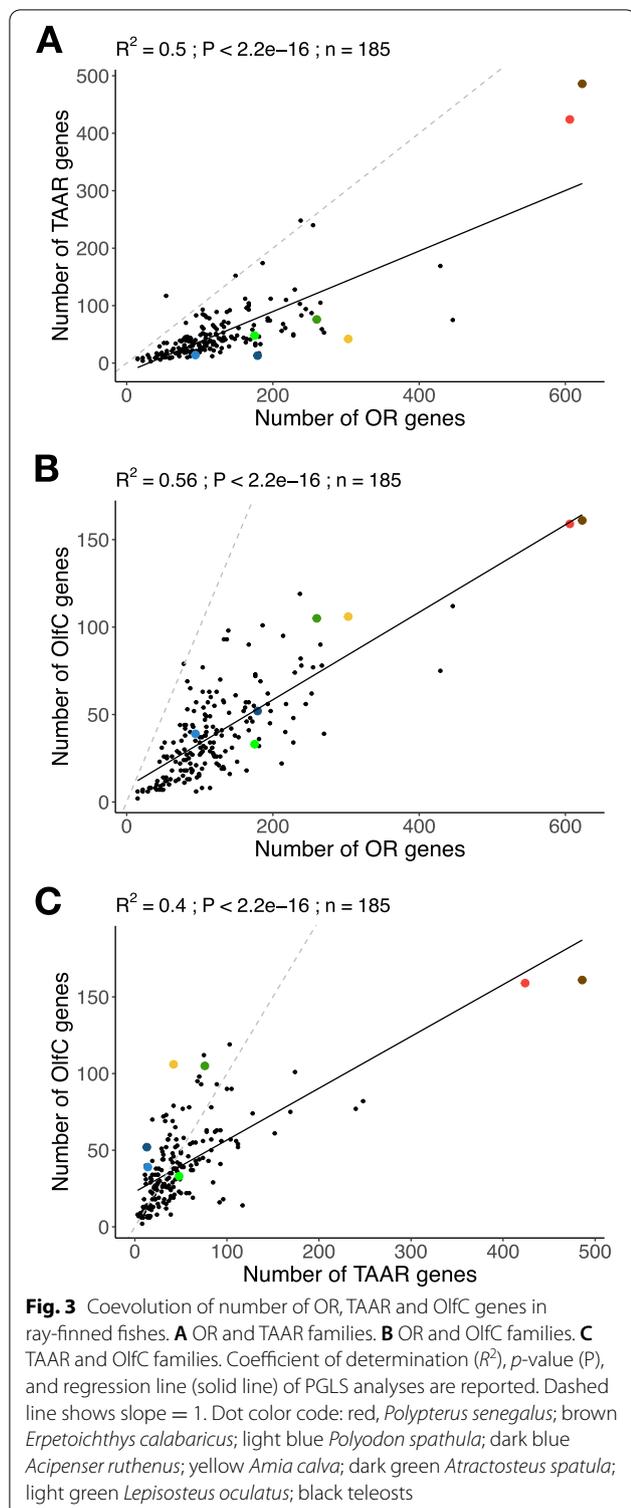
The evolution of the other three gene families (OR, TAAR, OlfC) has been more dynamic. For example, we identified an average of 126 functional OR genes in ray-finned fishes, but the variance is large, with 623 and 606 OR genes in the Polypteriformes *Erpetoichthys calabaricus* and *Polypterus senegalus*, respectively, and only 15 OR genes in the ocean sunfish *Mola mola* and broadnose pipefish *Syngnathus typhle* (Fig. 1, Additional file 1: Fig. S1B). The OR family is split into seven monophyletic subfamilies,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\xi$ , and  $\eta$  [21]. In tetrapods,  $\alpha$  and  $\gamma$  families expanded and other subfamilies are relictual or absent. In contrast, in teleosts, the  $\alpha$  family is absent and only one copy of a  $\gamma$  family gene occurs in Zebrafish *Danio rerio* [21]. Our analysis shows that  $\alpha$  family genes occur in all non-teleost actinopterygians but that the  $\alpha$  family was lost in the common ancestor of teleosts (Additional file 1: Fig. S1B). The  $\gamma$  family is well represented in non-teleost actinopterygians whereas only a few copies are scattered in the teleost phylogeny (Additional file 1: Fig. S1B). This suggests that the  $\gamma$  family was present in the common ancestor of teleosts but lost in most teleost lineages. The number of genes in the TAAR and OlfC repertoires is smaller than in the OR repertoire, with an average of 51 and 40 genes per species, respectively. For

these two gene families, the variance is also large. For example, *Erpetoichthys calabaricus* (Polypteriformes) has 486 TAAR and 161 OlfC genes. At the opposite extreme, only three TAAR genes were found in *Callionymus lyra* (Syngnathiformes) and two OlfC genes in *Mola mola* (Tetraodontiformes) (Fig. 1 and Additional file 1: Fig. S1C, D).

To analyze the evolutionary dynamics of the olfactory receptor gene families, we computed birth and death rates along branches of the phylogeny for the four families using the gene tree—species tree reconciliation method [22]. The mean birth and death rates were similar in OR, TAAR, and OlfC families, 0.0071/0.0071, 0.0101/0.0079, and 0.0059/0.0069 per gene per million years, respectively, but lower in the ORA family, 0.0018/0.0047 (Additional file 1: Fig. S3). Whereas birth and death rates are similar along most branches, we observed concomitant high death rates of OR, TAAR and OlfC genes in the common ancestor of two sampled species of Siluriformes (*Bagarius yarrelli* and *Tachysurus fulvidraco*), in the common ancestor of Lophiiformes and Tetraodontiformes, and in the common ancestor of Kurtiformes and Syngnathiformes. We also observed concomitant high birth rates of OR, TAAR, and OlfC genes in the common ancestor of Labriformes and Cyprinodontiformes and in the common ancestor of *Perca* + *Sander* (Fig. 1, Additional file 1: Fig. S3).

Despite variation in the number of genes in a family, we did not find evidence that contraction of one gene family is compensated by expansion of others. On the contrary, there is a correlation between the number of functional genes in each family (phylogenetic generalized least squares (PGLS);  $R^2 = 0.50$  between OR and TAAR,  $R^2 = 0.56$  between OR and OlfC,  $R^2 = 0.40$  between TAAR and OlfC, all  $p$ -values  $< 2e-16$ , Fig. 3). Moreover, in most species, the number of OR genes is greater than the number of TAAR or OlfC genes, and often, the number of TAAR genes is greater than the number of OlfC genes (Fig. 3). Although the number of ORA genes is less dynamic, particularly in teleosts, species with a high number of OR, TAAR, and OlfC genes, such as Polypteriformes or Anguilliformes, tend to have more ORA genes, whereas species with few genes in these three families, such as *Mola mola*, tend to have fewer ORA genes (Fig. 1).

The coevolution of the three dynamic receptor gene families (OR, TAAR, OlfC) is also supported by the correlation between the number of gene losses along the branches of the phylogenetic tree (Pearson's  $r = 0.8$  between OR and TAAR, 0.52 between OR and OlfC and 0.62 between TAAR and OlfC, all  $p$ -values  $< 2e-16$ ) and gene gains ( $r = 0.79$  between OR and TAAR, 0.78 between OR and OlfC and 0.69 between TAAR and OlfC,



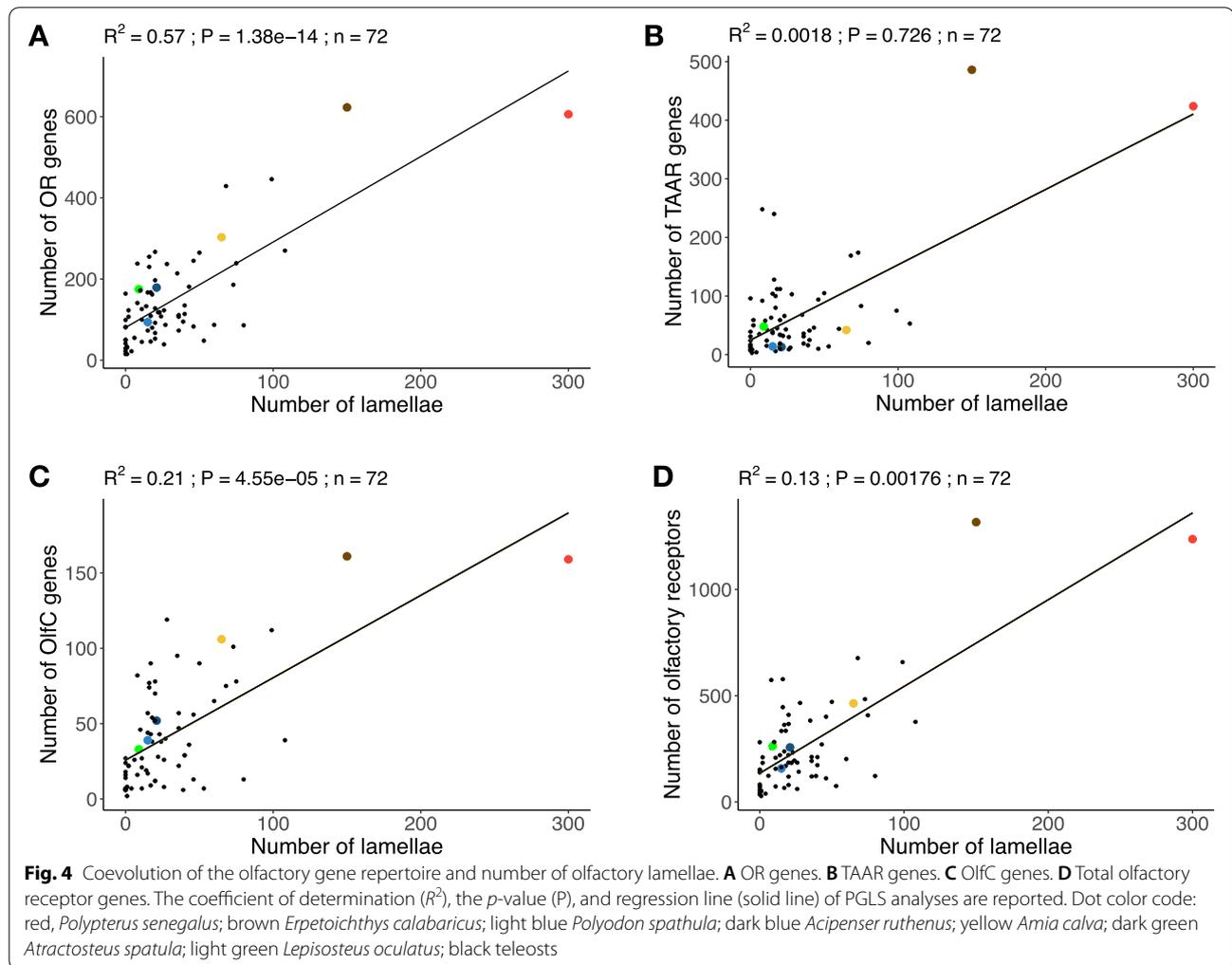
all  $p$ -values  $< 2e-16$ ) (Additional file 1: Fig. S4). The coevolution of the OR, TAAR, and OlfC receptor gene families is further supported by a correlation of the number and proportion of pseudogenes, which agrees with

similar gene death rates in the three dynamic gene families (Additional file 1: Fig. S5).

Together, these results suggest that dramatic changes in evolutionary constraints on the size of the olfactory repertoire occurred several times, with periods of expansion or contraction affecting OR, TAAR, and OlfC olfactory receptor families the same way. Hence, they do not constitute independent evolutionary units in ray-finned fishes.

### Coevolution of olfactory organ and olfactory gene repertoire

Using data for 72 species, 66 teleosts and 6 non-teleost ray-finned fishes (Additional file 2: Supplementary Data 1), we confirmed the correlation between the number of OR genes and the number of lamellae in the olfactory organ (PGLS;  $R^2 = 0.57$ ,  $p = 1.38e-14$ , Fig. 4A) reported recently for a smaller sample of 35 teleosts and two non-teleost ray-finned fishes [10]. While no significant correlation was found between the number of lamellae and the number of TAAR genes (PGLS;  $R^2 = 0.00177$ ,  $p = 0.726$ , Fig. 4B), a correlation was found with the number of OlfC genes (PGLS;  $R^2 = 0.21$ ,  $p = 4.55e-05$ , Fig. 4C) and the total number of olfactory receptor genes (PGLS;  $R^2 = 0.13$ ,  $p = 0.00176$ , Fig. 4D). The smallest olfactory repertoires occur in ocean sunfish *Mola mola* (28 genes) and broad-nosed pipefish *Syngnathus typhle* (35 genes). These extreme reductions of olfactory receptor diversity evolved independently and in parallel with the simplification of the olfactory organ, which is a small, flat olfactory epithelium in both species [10, 23] (Fig. 2A). Moreover, *M. mola* has greatly reduced olfactory nerves and reduced olfactory bulbs [24]. Limited data suggests that ocean sunfish are highly visual predators of gelatinous organisms [25–27]; however, more research on molid ecology is essential to determine if this is an example of a sensory tradeoff. At the other extreme is the unique organization of the olfactory organ of Polypteriformes. In both species studied, the olfactory organ consists of six sectors, each with a rosette-like structure [11, 12], resulting in many more olfactory lamellae than any other ray-finned fishes (Fig. 2A). Polypteriformes also have a much larger olfactory gene repertoire with many more genes in all four gene families than in any other ray-finned fishes (*Polypterus senegalus*: 1237 olfactory receptors, 300 olfactory lamellae; *Erpetoichthys calabaricus*: 1317 olfactory receptors, 150 olfactory lamellae; Fig. 1). The two other species studied that had the most olfactory receptor genes also had many olfactory lamellae (*Anguilla anguilla*: 658 olfactory receptors and 99 olfactory lamellae; *Mastacembelus armatus*: 677 olfactory receptors, 68 olfactory lamellae; Fig. 1). Interestingly, *P. senegalus*, *E. calabaricus*, *A. anguilla*, and *M. armatus* are nocturnal



[28, 29], perhaps making them more reliant on olfaction. They also have other specializations of the olfactory system, such as prominent, anteriorly directed incurrent narial tubes (Additional file 1: Fig. S6). Such tubes direct water flow into the olfactory organ, which allows the fish to sample water above its boundary layer and thus more rapidly detect odors [30].

After an extreme contraction of the olfactory gene repertoire and simplification of the olfactory epithelium, a secondary expansion in the gene repertoire occurred in parallel with the reacquisition of a multilamellar epithelium in the Tetraodontiformes genus *Takifugu*. The genomes of *Takifugu rubripes*, *T. flavidus*, and *T. bimaculatus* have more olfactory genes (156, 124, and 140 respectively) than other Tetraodontiformes with a flat olfactory epithelium, *Dichomyctere nigroviridis* (70 genes) and *Mola mola* (28 genes). The increased number of genes in the three species of *Takifugu* is due to duplications of OR, TAAR, and OlfC genes (Additional file 1:

Fig. S1B–D). We dissected a specimen of *T. rubripes* and found a non-rosette, but multilamellar, organization of parallel lamellae on the floor of the olfactory chamber that continues on the ventral surface of the nasal bridge between the incurrent and excurrent nares (Fig. 2A and Additional file 1: Fig. S7). This novel organization supports the hypothesis of a reacquisition of a multilamellar olfactory epithelium in association with secondary expansion of the olfactory receptor gene repertoire.

Together, our results indicate a functional link between the number of receptors and the number of lamellae in the olfactory organ of ray-finned fishes (Fig. 2B). In the most extreme cases, this leads to the loss of the rosette (e.g., *Mola mola* and *Syngnathus typhle*) or anatomical innovations with several rosettes (e.g., Polypteriformes) or a novel organization of olfactory lamellae (e.g., species of *Takifugu*). This link limits the morpho-genomic space occupied by ray-finned fishes (Fig. 2B). Accordingly, we did not observe ray-finned fishes with many olfactory

genes and few olfactory lamellae or fishes with few olfactory genes and many olfactory lamellae (Figs. 2 and 4). Because many olfactory neurons expressing each olfactory receptor are necessary for efficient olfaction, there is a functional limit to the number of olfactory receptor genes that can be expressed on a given area of olfactory epithelium. This would explain why there are no species with many olfactory receptor genes and few olfactory lamellae. We did not find any examples of macrosmatic fishes with a low number of olfactory receptor genes, which would favor high sensitivity for a small set of odors. There is probably no functional limit moderating the evolution of such a specialization, and perhaps cartilaginous fishes, which have few olfactory receptor genes and large multilamellar olfactory organs [31], may occupy this area of the morpho-genomic space.

No significant correlations were found between other morphological characters (maximum length of the fish, relative eye size (eye diameter/standard length of the fish)), ecological parameters (trophic level, preferred temperature, maximum depth), or genome size and the number of functional OR [10] or TAAR or OlfC genes (present study, data not shown).

## Conclusions

Our analysis of 185 highly complete genomes of ray-finned fishes highlights the diversity of the olfactory receptor repertoire. The number of genes is highly dynamic for three (OR, TAAR, OlfC) of the four gene families, but the reasons for large gene gains or losses are still unknown. In marine tetrapods, including cetaceans and sea snakes, extreme reductions in the number of olfactory genes occurred likely because air-adapted olfactory systems were not useful in marine environments [32]. No such major ecological transition is associated with gene losses of similar magnitude in Syngnathiformes and Tetraodontiformes, and it remains unknown why their olfaction degenerated at both morphological and genomic levels. The complexity of the olfactory organ and large olfactory gene repertoire in Polypteriformes is also surprising. These fishes have a high olfactory sensitivity [33]. An olfactory organ with a large olfactory epithelium surface is probably involved in high sensitivity; however, a link between sensitivity and gene repertoire size is less obvious. For example, some *Astyanax mexicanus* cavefish have a higher sensitivity ( $10^5$ ) to some molecules than surface conspecifics [34], while their olfactory gene repertoires are very similar (present study). To date, few olfactory receptor genes have been de-orphanized, and such functional information, combined with behavioral studies, may shed light on the dynamics of losses and specializations. Together, our analyses of the olfactory gene repertoire and morphology

of the olfactory epithelium show that olfaction is a heterogeneous sensory modality in ray-finned fishes. Our identification of non-model species with particularly poorly developed olfaction (e.g., *Mola mola*) or exceptionally well-developed sense of smell (e.g., *Erpetoichthys calabaricus*) opens new possibilities for comparative and functional research on olfaction.

## Methods

### Olfactory epithelium data

We surveyed the literature on olfactory organs in fishes and found data on number of lamellae for 60 species for which a draft genome assembly was available. We also dissected olfactory organs and made lamellae counts for 12 species at the National Museum of Natural History, Washington, DC, USA. Literature and specimen data were collected from adults because the number of lamellae often increases with total length (TL); we did not consider sexual dimorphism or individual age, which can impact number of olfactory lamellae [35, 36]. We classified the olfactory epithelium as flat if it had  $\leq 2$  lamellae and multilamellar if it had  $> 2$  lamellae following Hansen et al. (2005). Olfactory lamellae data used in the analyses is summarized in Additional file 2: Supplementary Data 1.

### Genome selection and mining of olfactory receptor genes

Using BUSCO (v5.2.2) [37], highly complete genomes (in terms of coding sequences) of 185 ray-finned fishes were selected, including 178 teleost genomes with a BUSCO score  $> 90\%$  and seven non-teleost genomes with BUSCO score ranging from 81% to 93% (Additional file 2: Supplementary Data 1).

A time-calibrated phylogenetic tree of ray-finned fishes was downloaded from <https://fishtreeoflife.org> [2] and pruned using the R package ape (v5.0) [38] to the 185 species included in our study.

Single-exon genes that code for OR receptors were mined following methods described by Policarpo et al. (2021). TAAR, OlfC, and ORA genes, which consist of several exons, were identified following [39], with slight modifications. In brief, a TBLASTN [40] was performed using known TAAR, OlfC, or ORA sequences as queries with a threshold  $e$ -value  $< 1e-10$  to select regions containing putative TAAR, OlfC, or ORA genes. Non-overlapping hit regions were extracted and extended 5000 bp upstream and downstream using SAMtools [41]. For each extended non-overlapping hit region, the protein with the best TBLASTN match was aligned to the DNA sequence using EXONERATE (v2.2) [42], and the resulting protein-coding sequence was used as query for a BLASTX against a custom database of OR, TAAR, OlfC, ORA, and other G protein-coupled

receptors (GPCRs). Protein-coding sequences that best matched TAAR, OlfC, or ORA receptors were retained and manually curated. Each protein-coding sequence was translated and aligned to known olfactory receptors and other GPCR genes with MAFFT (v7.487) [43], and maximum likelihood trees were computed with IQ-TREE (v1.6.12) [44]. Only protein-coding sequences that clustered with known olfactory receptors by visual inspection using iTOL [45] were retained as olfactory receptor genes. When several identical sequences were retrieved in a genome, only one was kept using CD-HIT [46].

Retrieved coding sequences were classified as (1) ‘gene’ if complete and without loss-of-function mutation (premature stop codon or frameshift), (2) ‘pseudogene’ if with at least one loss-of-function mutation, (3) ‘truncated’ if incomplete and without loss-of-function mutation, and (4) ‘edge’ if incomplete and less than 30 bp from a contig border.

We assessed the quality of our mining pipeline by comparing the olfactory gene repertoires we identified with those published by other authors for four teleost species. We systematically found more genes than previous studies; in particular, in *P. senegalus* [47], we identified three times more OR genes (Additional file 3: Supplementary Data 2).

### Phylogenies and gene classification

For each species, we aligned protein sequences coded by putative OR genes with known OR genes [21] using MAFFT. A maximum likelihood tree was computed with IQ-TREE, and genes were classified according to their position in the tree. To assess the relative diversity of OR subfamilies, a phylogenetic tree with OR genes of 44 species, each species belonging to a different order based on fishreeoflife (<https://fishreeoflife.org/>), was computed. The root was placed between type I and type II genes (Additional file 4: Supplementary Data 3). Using MAFFT, putative TAAR genes were aligned with TAARs and non-TAAR GPCRs genes obtained from [20]. A maximum likelihood tree was computed with IQ-TREE and genes were classified according to their position in the phylogenetic tree (Additional file 4: Supplementary Data 3). The same method was used for putative OlfC and ORA genes. For putative OlfC genes, we used genes from [16] and CasR and V2R2 genes as outgroups (Additional file 4: Supplementary Data 3). For ORA sequences, we used genes from [15] and T2R genes as an outgroup (Additional file 4: Supplementary Data 3).

Pseudogenes, truncated genes, and edge gene classification were based on the best blastx match.

### Phylogenetic comparative analyses

We estimated phylogenetic signal (Pagel’s  $\lambda$ ) of each trait with the function `phylosig` in the R package `phytools` with the option `test = TRUE` [48]. The R package `caper` (v1.0.1) [49] was used to perform phylogenetic generalized least square analyses using the function “`pgls`” with `lambda = “ML”` (Additional file 2: Supplementary Data 1).

### Gene tree—species tree reconciliation

The number of gene gains and number of gene losses along each branch of the species phylogenetic tree were inferred using the gene tree—species tree reconciliation method. The OR family is large, as described previously [10], and thus OR genes belonging to different subfamilies were aligned separately. For the smaller TAAR, OlfC, and ORA gene families, one alignment was obtained for each gene family separately. All alignments were obtained using MAFFT. Maximum likelihood trees were computed with IQ-TREE. Nodes with low bootstrap values (< 90%) were collapsed into polytomies using the R package `ape`. We then used `Treerects` [22] to root and reconcile genes trees with the species tree.

For each olfactory receptor family, we computed birth and death rates using equations in [50] excluding branches with length < 2 Mya because differences in gene retrieval and genome qualities greatly impacted inferred birth and death rate [10].

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-022-01397-x>.

**Additional file 1: Figure S1.** Diversity of the olfactory receptor gene repertoire in ray-finned fishes. **Figure S2.** Distribution of ORA7 and ORA8 subfamilies in ray-finned fishes. **Figure S3.** Distribution of birth and death rates of OR, TAAR, OlfC and ORA genes in ray-finned fishes. **Figure S4.** Correlation of the number of gene losses (or gene gains) between gene families, estimated using the 368 branches of the phylogenetic tree. **Figure S5.** Correlation between the number of OR, TAAR and OlfC pseudogenes in 185 ray-finned fishes. **Figure S6.** Narial tubes of four species of ray-finned fishes with complex olfactory organs and large gene repertoires. **Figure S7.** *Takifugu rubripes*, USNM 57620, 290 mm TL.

**Additional file 2.** Supplementary Data 1. Sheet 1: NCBI Assembly accession and assembly level of the 185 genomes studied, their species name in NCBI and in Eschmeyer’s Catalog of Fishes. Sheet 2: results of BUSCO analyses on the 185 genomes studied. Sheet 3: species’ name and order based on the taxonomy of [fishreeoflife.org](https://fishreeoflife.org/). Sheet 4: summary of the number of genes in each olfactory receptor family. Sheet 5: Olfactory epithelium shape and number of lamellae in 72 ray-finned fishes for which a genome assembly is available. Sheet 6: phylogenetic signal of the number of genes in each family and of the number of lamellae in the epithelium computed with `phytools`. Values of phylogenetic regression described in this study are also given.

**Additional file 3.** Supplementary Data 2. Comparison of olfactory receptor gene repertoires from the present and previous studies. (A) Summary of the number of TAAR genes retrieved in our study and previous studies of four teleost species. (B) Summary of the number of OlfC genes retrieved in our study and previous studies of four teleost species. (C) Summary

of the number of ORA genes retrieved in our study and previous studies of four teleost species. (D) Phylogenetic tree of *Danio rerio* TAAR genes retrieved in Hashiguchi and Nishida 2007 and our study. (E) Phylogenetic tree of *Gasterosteus aculeatus* TAAR genes retrieved in Azzouzi et al. 2015 and our study. (F) Phylogenetic tree of *Oryzias latipes* TAAR genes retrieved in Azzouzi et al. 2015 and our study. (G) Phylogenetic tree of *Takifugu rubripes* TAAR genes retrieved in Hashiguchi and Nishida 2007 and our study. (H) Phylogenetic tree of *Danio rerio* OlFC genes retrieved in Yang et al. 2019 and our study. (I) Phylogenetic tree of *Gasterosteus aculeatus* OlFC genes retrieved in Yang et al. 2019 and our study. (J) Phylogenetic tree of *Oryzias latipes* OlFC genes retrieved in Yang et al. 2019 and our study. (K) Phylogenetic tree of *Takifugu rubripes* OlFC genes retrieved in Yang et al. 2019 and our study. (L) Phylogenetic tree of *Danio rerio* ORA genes retrieved in Zapilko and Korsching 2016 and our study. (M) Phylogenetic tree of *Gasterosteus aculeatus* ORA genes retrieved in Zapilko and Korsching 2016 and our study. (N) Phylogenetic tree of *Oryzias latipes* ORA genes retrieved in Zapilko and Korsching 2016 and our study. (O) Phylogenetic tree of *Takifugu rubripes* ORA genes retrieved in Zapilko and Korsching 2016 and our study. (P) Phylogenetic tree of *Polypterus senegalus* OR genes retrieved in Bi X et al. 2021 and our study.

**Additional file 4.** Supplementary Data 3. (A) Phylogeny of OR genes from 44 species representing 44 orders of ray-finned fishes sampled in this study. Branches are colored according to the gene subfamily classification. (B) Phylogeny of all TAAR genes retrieved from 185 ray-finned fishes. Branches are colored according to gene family classification. Outgroup sequences (nonTAAR GPCRs) are colored in black. (C) Phylogeny of all OlFC genes retrieved from 185 ray-finned fish. Branches are colored according to the gene subfamily classification. Outgroup sequences (CasR and V2R2 genes) are colored in black. (D) Phylogeny of all ORA genes retrieved from 185 ray-finned fishes. Branches are colored according to the gene subfamily classification. Outgroup sequences (T2R genes) are colored in black.

#### Acknowledgements

J. Galbraith, NOAA Northeast Fisheries Science Center, provided the examined specimen of *Mola mola*.

#### Authors' contributions

MP, KEB, SR, and DC conceived the project and designed the research. MP and KEB collected the data. MP, KEB, and DC analyzed the data. MP, KEB, and DC wrote the manuscript with contributions from PL, LL, JCS and SR. SR and DC acquired the funding. All authors read and approved the final manuscript.

#### Funding

This work was supported by a collaborative grant from Institut Diversité Ecologie et Evolution du Vivant (to SR and DC). MP was supported by a PhD fellowship from the French Ministry of Research.

#### Availability of data and materials

All sequences analyzed in this study are available on [https://figshare.com/articles/dataset/Olfactory\\_receptor\\_sequences\\_for\\_185\\_ray-finned\\_fishes/17061632](https://figshare.com/articles/dataset/Olfactory_receptor_sequences_for_185_ray-finned_fishes/17061632) [51].

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, 91198 Gif-sur-Yvette, France. <sup>2</sup>NOAA National Systematics Laboratory, National Museum of Natural History, Smithsonian Institution, Washington, D.C. 20560, USA. <sup>3</sup>Université Paris Cité, Laboratoire Interdisciplinaire des

Energies de Demain, 75013 Paris, France. <sup>4</sup>Université Paris-Saclay, CNRS, Institut des Neurosciences Paris-Saclay, 91400, Saclay, France. <sup>5</sup>Université Paris Cité, UFR Sciences du Vivant, 75013 Paris, France.

Received: 23 February 2022 Accepted: 24 August 2022

Published online: 01 September 2022

#### References

- Fricke R, Eschmeyer WN, Van der Laan R. Eschmeyer's Catalog of Fishes: genera, species, references. (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Electronic version accessed 16 October 2021. 2021.
- Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, et al. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*. 2018;559(7714):392–5.
- Hughes LC, Ortí G, Huang Y, Sun Y, Baldwin CC, Thompson AW, et al. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc Natl Acad Sci USA*. 2018;115(24):6249–54.
- Helfman G, Collette BB, Facey DE, Bowen BW. The diversity of fishes: biology, evolution, and ecology: Wiley-Blackwell; 2009.
- Hara TJ. Olfaction in fish. *Prog Neurobiol*. 1975;5:271–335.
- Burne RH. The anatomy of the olfactory organ of teleostean fishes. *Proc Zoological Soc London*. 1909;2:610–63.
- Teichmann H. Vergleichende untersuchungen an der nase der fische. *Zeitschrift für Morphologie und Ökologie der Tiere*. 1954;43(2):171–212.
- Ferrando S, Gallus L, Ghigliotti L, Amaroli A, Abbas G, Vacchi M. Clarification of the terminology of the olfactory lamellae in Chondrichthyes. *Anatomical Record*. 2017;300(11):2039–45.
- Hansen A, Zielinski BS. Diversity in the olfactory epithelium of bony fishes: development, lamellar arrangement, sensory neuron cell types and transduction components. *J Neurocytol*. 2005;34(3):183–208.
- Policarpo M, Bemis KE, Tyler JC, Metcalfe CJ, Laurenti P, Sandoz J-C, et al. Evolutionary dynamics of the OR gene repertoire in teleost fishes: evidence of an association with changes in olfactory epithelium shape. *Mol Biol Evol*. 2021;38(9):3742–53.
- Pfeiffer W. Das geruchsorgan der Polypteridae (Pisces, Brachiopterygii). *Z Morph Tiere*. 1968;63(1):75–110.
- Theisen B. The morphology and vascularization of the olfactory organ in *Calamoichthys calabaricus* (Pisces, Polypteridae). *Vidensk Meddr dansk naturh Foren*. 1970;133:31–50.
- Buck L, Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*. 1991;65(1):175–87.
- Kowatschew D, Korsching SI. An ancient adenosine receptor gains olfactory function in bony vertebrates. *Genome Biol Evol*. 2021;13(9):evab211.
- Zapilko V, Korsching SI. Tetrapod V1R-like ora genes in an early-diverging ray-finned fish species: the canonical six ora gene repertoire of teleost fish resulted from gene loss in a larger ancestral repertoire. *BMC Genom*. 2016;17(1):83.
- Yang L, Jiang H, Wang Y, Lei Y, Chen J, Sun N, et al. Expansion of vomeronasal receptor genes (OlFC) in the evolution of fright reaction in Ostariophysan fishes. *Commun Biol*. 2019;2(1):235.
- Liu H, Chen C, Lv M, Liu N, Hu Y, Zhang H, et al. A chromosome-level assembly of Blunt Snout Bream (*Megalobrama amblycephala*) genome reveals an expansion of olfactory receptor genes in freshwater fish. *Mol Biol Evol*. 2021;38(10):4238–51.
- Hussain A, Saraiva LR, Korsching SI. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. *Proc Natl Acad Sci*. 2009;106(11):4313–8.
- Hashiguchi Y, Nishida M. 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*. 2007;24(9):2099–107.
- Dieris M, Kowatschew D, Korsching SI. Olfactory function in the trace amine-associated receptor family (TAARs) evolved twice independently. *Sci Rep*. 2021;11(1):7807.

21. Niimura Y. On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. *Genome Biol Evol.* 2009;1:34–44.
22. Comte N, Morel B, Hasić D, Guéguen L, Boussau B, Daubin V, et al. Treerecs: an integrated phylogenetic tool, from sequences to reconciliations. *Bioinformatics.* 2020;36(18):4822–4.
23. Dymek J, Rosenqwert G, Kuciel M, Lauriano ER, Capillo G, Zaccone G, et al. Micro- and macro-morphology of the olfactory organ of *Syngnathus typhle* (Syngnathidae, Actinopterygii). *Acta Zoologica.* 2021;102(2):206–19.
24. Burr HS. The central nervous system of *Orthogoriscus mola*. *J Comparative Neurol.* 1928;45(1):33–128.
25. Kino M, Miayzaki T, Iwami T, Kohbara J. Retinal topography of ganglion cells in immature ocean sunfish, *Mola mola*. *Environ Biol Fish.* 2009;85(1):33–8.
26. Nakamura I, Goto Y, Sato K. Ocean sunfish rearm at the surface after deep excursions to forage for siphonophores. *J Anim Ecol.* 2015;84(3):590–603.
27. Phillips ND, Harrod C, Gates AR, Thys TM, Houghton JDR. Seeking the sun in deep, dark places: mesopelagic sightings of ocean sunfishes (Molidae). *J Fish Biol.* 2015;87(4):1118–26.
28. Znotinas KR, Standen EM. Aerial and aquatic visual acuity of the grey bichir *Polypterus senegalus*, as estimated by optokinetic response. *J Fish Biol.* 2019;95(1):263–73.
29. LaBar GW, Hernando Casal JA, Delgado CF. Local movements and population size of European eels, *Anguilla anguilla*, in a small lake in southwestern Spain. *Environ Biol Fish.* 1987;19(2):111–7.
30. Cox JPL. Hydrodynamic aspects of fish olfaction. *J Royal Soc Interface.* 2008;5(23):575–93.
31. Sharma K, Syed AS, Ferrando S, Mazan S, Korsching SI. The chemosensory receptor repertoire of a true shark is dominated by a single olfactory receptor family. *Genome Biol Evol.* 2019;11(2):398–405.
32. Kishida T. Olfaction of aquatic amniotes. *Cell Tissue Res.* 2021;383:353–65.
33. Der PW. Geruchssinn der Polypteridae (Pisces, Brachiopterygii). *Zeitschrift für vergleichende Physiologie.* 1969;63(2):151–64.
34. Hinaux H, Devos L, Blin M, Elipot Y, Bibliowicz J, Alié A, et al. Sensory evolution in blind cavefish is driven by early embryonic events during gastrulation and neurulation. *Development.* 2016;143(23):4521.
35. Kasumyan A. The olfactory system in fish: structure, function, and role in behavior. *J Ichthyol.* 2004;44:5180–223.
36. Abrahão VP, Pastana M, Marinho M. On a remarkable sexual dimorphic trait in the Characiformes related to the olfactory organ and description of a new miniature species of *Tyttobrycon* Géry (Characiformes: Characidae). *PLoS One.* 2019;14(12):e0226130.
37. Waterhouse RM, Seppely M, Simão FA, Manni M, Ioannidis P, Kliuchnikov G, et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mole Biol Evol.* 2018;35(3):543–8.
38. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2019;35(3):526–8.
39. Azzouzi N, Barloy-Hubler F, Galibert F. Identification and characterization of cichlid TAAR genes and comparison with other teleost TAAR repertoires. *BMC Genom.* 2015;16(1):335.
40. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinform.* 2009;10(1):421.
41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9.
42. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* 2005;6:31.
43. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mole Biol Evol.* 2013;30(4):772–80.
44. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mole Biol Evol.* 2015;32(1):268–74.
45. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007;23(1):127–8.
46. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28(23):3150–2.
47. Bi X, Wang K, Yang L, Pan H, Jiang H, Wei Q, et al. Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. *Cell.* 2021;184(5):1377–91.e14.
48. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol.* 2012;3(2):217–23.
49. Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, et al. Caper: comparative analyses of phylogenetics and Evolution in R. R package version 1.0.1. <https://CRAN.R-project.org/package=caper>. 2018.
50. Niimura Y, Matsui A, Touhara K. Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome Res.* 2014;24(9):1485–96.
51. Policarpo M. Olfactory receptor sequences for 185 ray-finned fishes. figshare [https://figshare.com/articles/dataset/Olfactory\\_receptor\\_sequences\\_for\\_185\\_ray-368finned\\_fishes/17061632\(2021\)](https://figshare.com/articles/dataset/Olfactory_receptor_sequences_for_185_ray-368finned_fishes/17061632(2021))

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

