Genome-wide and gene expression analyses of the hagfish olfactory receptors illuminate lineage-specific diversification of olfaction in basal vertebrates

Hirofumi Kariyayama[1,2†], Ooi Yusuke[3†], Hiromu Kashima[4†], Taiki Nakanowatari[5], Riho Harada[3], Yoko Yamaguchi[6], Daichi G. Suzuki[7]*

[1] Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan

[2] Laboratory for AI Biology, RIKEN Center for Biosystems Dynamics Research, Kobe, Japan

[3] Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

[4] Graduate School of Natural Science and Technology, Shimane University, Japan

[5] School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

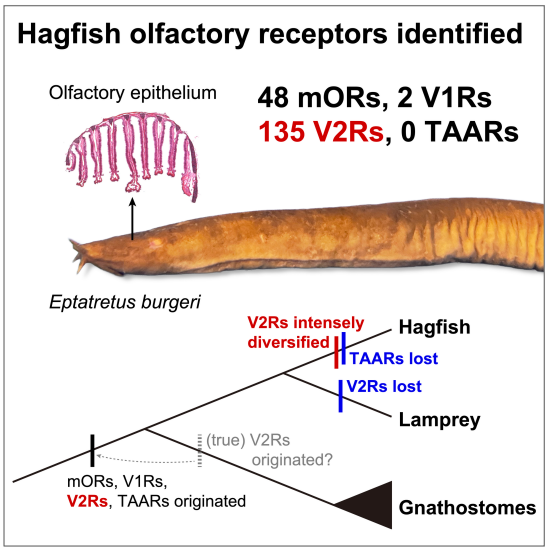
[6] Institute of Agricultural and Life Sciences, Academic Assembly, Shimane University, Japan

[7] Institute of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

*For correspondence: suzuki.daichi.gp@u.tsukuba.ac.jp

[†]These authors contributed equally to this work.

Graphical abstract



Highlights

- Genomic and expression analyses reveal hagfish olfactory receptor repertoire.
- V2Rs are extensively diversified in hagfish.
- The olfactory function of V2Rs is traced back to the common ancestor of vertebrates.
- The olfactory receptor repertoires of basal vertebrates underwent mosaic evolution.

Summary

Olfactory receptors (ORs) of vertebrates have undergone remarkable diversification, but the ancestral repertoire remains incompletely understood. Here, we present a comprehensive genome-wide survey and gene expression analyses of OR genes in hagfish, a cyclostome group that typically inhabits the deep sea with a well-developed olfactory system. We identified 48 main olfactory receptors (mORs), two vomeronasal type 1 receptors (V1Rs), 135 typical vomeronasal type 2 receptors (V2Rs), and no trace amine-associated receptors (TAARs) in the inshore hagfish (*Eptatretus burgeri*). Notably, V2Rs were highly clustered on a single scaffold, suggesting tandem gene duplications as the main mechanism of diversification. Transcriptome analysis confirmed that most of these OR genes are expressed predominantly in the olfactory organ. Using *in situ* hybridization, we further confirmed that at least some of these genes are expressed in olfactory receptor cells. These gene expression analyses support the role of these genes as functional ORs. Our findings challenge the prevailing view that typical olfactory V2Rs originated in gnathostomes and instead suggest that they were present in the common ancestor of all vertebrates but were lost in lampreys. The extensive diversification of V2Rs in hagfish likely reflects adaptations to life in low-light and chemically complex environments. This study highlights the evolution of vertebrate sensory systems in a mosaic manner, underscoring the importance of hagfish as a key animal for reconstructing the evolutionary history of early vertebrates.

Introduction

Olfaction is a vital sense modality for animals. In particular, it plays an even more important role in darkness because visual information is limited in such environments. It is suggested that our mammalian ancestors experienced nocturnal life, in which they heavily relied on olfaction, accompanied by enlargement of the olfactory bulb and cerebral cortex (Rowe et al.

53 2011). Along with this brain evolution, their olfactory receptors diversified explosively
54 (Hughes et al. 2018).
55 A similar evolutionary trend may have occurred to hagfish, which belongs to the
56 basal-most lineage of vertebrates (i.e., the cyclostomes). Hagfish typically live in light-
57 decaying deep water and have a degenerate visual system (Fernholm and Holmberg 1975).
58 Instead, these animals possess a highly developed olfactory bulb and telencephalon (the brain
59 region homologous to the mammalian cerebrum), suggesting that their capacity for olfaction
60 is correspondingly elaborated (Suzuki 2021). However, the repertoire of hagfish olfactory
61 receptors remains largely unknown.
62 Apart from such habitat-specific diversification of olfaction, hagfish may provide
63 crucial information for the early evolution of the vertebrate olfactory receptors (ORs). Within
64 vertebrates, there are four major OR families: main olfactory receptors (mORs, or ORs in the
65 narrow sense), type 1 vomeronasal receptors (V1Rs), type 2 vomeronasal receptors (V2Rs),
66 and trace amine-associated receptors (TAARs). Previous research has shown that the lamprey
67 genome contains 27 mORs (*Petromyzon marinus*; Libants et al. 2009), six V1Rs (*P. marinus*
68 and *Lethenteron camtschaticum*; Kowatschew and Korsching 2022), one (*P. marinus*) or two
69 (*L. camtschaticum*) V2R-like proteins, and 51 (*P. marinus*) or 32 (*L. camtschaticum*) TAARs
70 (Dieris et al. 2021). However, Kowatschew and Korsching (2022) report that the *P. marinus*
71 V2R-like gene is not expressed in the olfactory epithelium. Bi et al. (2021) identified more
72 than 50 V2R-like sequences in hagfish. Still, Zhang et al. (2022) suggest that no typical V2Rs
73 exist in extant agnathans (i.e., lampreys and hagfish) because they did not form a clade with
74 known V2Rs in their phylogenetic analysis. To understand the ancestral repertoire of
75 vertebrate ORs, it is therefore necessary to investigate hagfish ORs in a more thorough and
76 detailed manner.

77 In this study, we conducted a comprehensive genome-wide survey and gene
78 expression analyses of hagfish ORs. We first identified 48 mOR, two V1R, 137 V2R, but no
79 TAAR candidate sequences in the inshore hagfish (*Eptatretus burgeri*) genome. Then,
80 phylogenetic analysis revealed that all mOR candidates are type 1 mOR genes and that the
81 two V1R candidates are also true vertebrate V1Rs. Regarding V2Rs, 135 among
82 137candidates were estimated to be typical V2Rs. We next explored the genomic distribution
83 of the identified OR genes and found that phylogenetically close genes tend to be clustered in
84 the genome. Transcriptome analysis indicated that most of these ORs are expressed in the
85 olfactory organ. Using *in situ* hybridization, we further confirmed that at least some of these
86 genes are expressed in olfactory receptor cells. These results reveal the comprehensive
87 olfactory receptor repertoire of hagfish, suggesting that the common ancestor of vertebrates
88 had functional olfactory V2Rs and that both lampreys and hagfish have experienced lineage-
89 specific diversification of olfactory receptors.
90

91 ## Results

92 **Gene identification and phylogenetic analysis of the hagfish ORs**

93 *mORs*

94 In osteichthyes, including tetrapods, mORs are the biggest gene family among olfactory
95 receptors (see Table 1). We explored hagfish gene models and genome data, and then found
96 48 mOR candidates (for the list of the identified hagfish OR genes, including these mOR
97 candidates, see Supplementary File 1). We also identified 62 lamprey mOR genes that are
98 predicted to be functional from the newest version (Timoshevskaya et al., 2023) of the sea
99 lamprey genome (for the list of the lamprey OR genes identified in this study, see
100 Supplementary File 2). Phylogenetic analysis revealed that all of these hagfish genes form
101 two groups, both belonging to the type 1 mOR subfamily (Fig. 1A). Each group contained

102 hagfish genes alone, showing monophyly with a high support value (100%). One of them
103 (hagfish Group A) included 37 genes, belonging to a larger group with a monophyletic
104 lamprey clade (lamprey Group A) supported by a high bootstrap value. The other (hagfish
105 Group B) consisted of 11 genes and formed a monophyletic group with a sea lamprey gene,
106 with a relatively low support value. No type 2 genes were identified from hagfish, although
107 some lamprey genes were included in this subfamily as previous research suggested (Niimura
108 2009).
109
110 *V1Rs*
111 V1R or olfactory receptor class A (ORA) genes are known to be diversified in
112 sarcopterygians, including tetrapods, in which these receptors are thought to be involved in
113 pheromone detection (Nikaido 2019). We found two V1R candidates in the hagfish genome,
114 while more V1Rs (eight genes) were identified in the sea lamprey genome. It is suggested
115 that vertebrate V1Rs are divided into six subfamilies (ORA1–6; Saraiva and Korsching
116 2007). Through phylogenetic analysis, we found that the hagfish V1R1 candidate forms a
117 monophyletic clade with gnathostome ORA3/4 along with a sea lamprey V1R, supported by
118 a high bootstrap value (99%; Fig. 1B). The hagfish V1R2 candidate was placed as the sister
119 group of the ORA5 family of gnathostomes with a high bootstrap value (97%). In addition,
120 some sea lamprey V1Rs constructed a clade with gnathostome ORA1/2 with a relatively high
121 support value, although the phylogenetic positions of other sea lamprey genes were unclear
122 due to their relatively low support values.
123
124 *V2Rs*
125 V2Rs are chemoreceptors targeting amino acids and peptide pheromones in gnathostomes
126 (Leinders-Zufall et al. 2004; Kimoto et al. 2005; Speca 1999; Koide et al., 2009; Sato et al.,

127 2005). We identified 137 V2R candidates in hagfish. In contrast, only one V2R-like gene was
128 found in the sea lamprey genome, consistent with previous reports (Kowatschew and
129 Korsching 2022). Phylogenetic analysis revealed that one hagfish V2R candidate forms a
130 clade with lamprey and gnathostome calcium-sensing receptor (CaSR) genes, another
131 grouped with the lamprey V2R-like genes, and the remaining 135 genes forms a
132 monophyletic clade together with high bootstrap value (96%), being situated in a basal
133 position to gnathostome V2Rs more closely than the cyclostome V2R-like clade (Fig. 1C).
134 Based on these results, we estimated the first and second candidate gene as hagfish CaSR and
135 V2R-like, respectively, and the last group containing135 candidates as typical V2Rs.
136
137 *TAARs*
138 Previous studies have reported that TAARs are diversified in lamprey (Libants et al. 2009;
139 Dieris et al. 2021). To find hagfish TAARs, we first performed a reciprocal BLAST search
140 against hagfish gene models, but no plausible candidates were found. For further
141 confirmation, we performed phylogenetic analysis of top hit hagfish sequences and known
142 TAARs. Consequently, these hagfish genes did not show affinity to TAARs but to various
143 groups of GPCRs, including the 5-hydroxytryptamine receptor 4 (5-HT4 or HTR4),
144 adrenoceptor beta 1, and dopamine receptor family (Supplementary File 7). These results
145 suggest that hagfish have no TAAR homologs, consistent with previous research (Guo et al.,
146 2022).
147
148 **Genomic distribution of the hagfish ORs**
149 To clarify the diversification pattern of hagfish ORs, we then examined the genomic
150 distribution of the identified genes using the following analyses.

151 First, we defined phylogenetically related clades for mORs and V2Rs (See Materials
152 and Methods). We categorized three major clades for mORs; Group A corresponded entirely
153 to Clade 1, while Group B was subdivided into Clade 2 and 3 (see Supplementary Fig. 1). For
154 V2Rs, nine major clades (Clade 1–9) were distinguished, and the remaining unclassified
155 genes were categorized into "Else" (see Supplementary Fig. 2).
156 Next, we mapped all identified OR genes onto the genome (Fig. 2A). As the current
157 *E. burgeri* genome is not assembled at the telomere-to-telomere level, we arranged scaffolds
158 (> 0.1 Mbp) from longer to shorter. As an overall pattern, the OR genes exhibited a patchy
159 distribution. For example, mORs in the same clade tended to co-occur on the same scaffold.
160 Even in each scaffold, phylogenetically related genes tended to be neighboring closely in the
161 genome as follows. We calculated phylogenetic and genomic distances between each
162 pairwise combination of the identified genes, and then made a scatter plot between these two
163 distances (Fig. 2C). As a result, we found that a higher portion of intraclade pairs (32.2%;
164 140/435 pairs) were located within a shorter genomic distance (<10 Mbp), compared to the
165 interclade pairs (1.0%; 7/693 pairs).
166 V2Rs also exhibited a patchy genomic distribution, and most of them (91.1%;
167 123/135 were concentrated on a single scaffold (FYBX03000002.1). In this scaffold,
168 phylogenetically close genes were arranged in clusters along the genome (Fig. 2B). In
169 addition, scatter plotting between phylogenetic and genomic distances (Fig. 2D) showed that
170 most of the intraclade pairs (69.8%; 273/391 pairs) were located within a shorter genomic
171 distance (<10 Mbp), but fewer interclade pairs (21.5%; 1862/8652 pairs) were within such a
172 close distance.

173 **Organ-specific gene expression of the hagfish ORs**

174 For the examination of whether the hagfish olfactory receptor candidates are indeed
175 expressed in the olfactory organ, we performed organ-specific gene expression analyses using
176 bulk RNA-seq data (the olfactory organ, brain, pituitary gland, gill, intestine, kidney, liver,
177 testis, and ovary). As a result, we confirmed that most of these candidate genes were highly
178 expressed in the olfactory organ.

179 First, 47 genes among 48 mOR candidates showed high expression levels in the
180 olfactory organ (44 genes showed the highest expression levels in the olfactory organ among
181 the nine tested organs), suggesting their function as true olfactory receptors (Fig. 3). Notably,
182 many (19 genes) are also found to be expressed at a certain level in the testis and/or ovary.

183 Second, we confirmed that the two V1R candidates were also expressed at a relatively
184 high level (TPM > 1) in the olfactory organ (Fig. 4A). The Ebu-V1R1 gene was also
185 expressed in the brain at the same expression level, and was slightly weakly expressed in the
186 liver and kidney. The Ebu-V1R2 gene showed the highest expression in the olfactory organ,
187 with lower expression levels in other organs.

188 Last, we detected expression of all 135 hagfish putative V2R genes in the olfactory
189 organ (Fig. 4B). Moreover, almost all of them (134 genes except V2R22) showed the highest
190 expression in this organ. Impressively, expression of these genes was also frequently detected
191 in the testis and/or ovary. In contrast, neither the hagfish CaSR nor V2R-like was expressed
192 in the olfactory organ; instead, the former was expressed in all the other organs tested
193 (especially in the kidney and testis), while expression of the latter was detected highly in the
194 intestine and testis, and weakly in the pituitary gland.

195 **Gene expression patterns of the hagfish ORs in the olfactory epithelium**

196 To investigate whether the olfactory receptor candidates are specifically expressed in
197 olfactory receptor cells, we performed *in situ* hybridization analysis for hagfish mOR
198 (mOR5), V1R (V1R1), and V2R (V2R33), selected based on their expression levels and
199 predicted sequence lengths.

200 The hagfish olfactory organ, or the nasal basket, is situated anterior to the brain (Fig.
201 5A; Marinelli and Strenger 1956; Muramatsu et al., 2024). It consists of a series of seven
202 olfactory lamellae, oriented parallel to the body axis and attached to the dorsal roof of the
203 olfactory cavity (Fig. 5B; Døving 1998).

204 By histological analysis, we first confirmed that the hagfish olfactory lamella exhibits
205 a laminar structure (Fig. 5C), in a manner similar to that of the teleost fish (Hara 1975). Each
206 lamella consists of a medial fibrous layer called the central core and the peripheral olfactory
207 epithelium, which is further subdivided into three layers: from proximal to distal, the basal
208 cell (BC) layer, the olfactory receptor cell (ORC) layer, and the apical cilia/microvilli layer.

209 Based on this observation, we then examined gene expression patterns of hagfish
210 *mOR5*, *V1R1*, and *V2R33*. As a result, we found that all of them were expressed in olfactory
211 cells in a scattered pattern (Fig. 5D–F). Negative controls using sense probes showed no such
212 expression patterns (Supplementary Fig. 3). Among these genes, V1R1-positive cells showed
213 a distinctive pattern, being broadly and relatively densely distributed in the ORC layer (Fig.
214 5E). In addition, some *V1R1*-positive cells were also observed in the BC layer. In contrast,
215 the olfactory cells expressing the other two (*mOR5* and *V2R33*) were more sparsely
216 distributed and found in a relatively superficial part of the ORC layer (Fig. 5D, F).

217 ## Discussion

218 **Hagfish OR repertoire**

219 In this study, we have investigated the olfactory receptor repertoire in hagfish. By combining
220 genome-wide and gene expression analyses, we identified putatively functional olfactory
221 receptor genes as follows (see also Table 1 and Fig. 6).

222

223 *mORs*

224 We identified 48 hagfish mOR genes, all of which fall into type 1 mORs (Fig. 1A). In
225 addition, almost all of them (47 genes) are highly expressed in the olfactory organ (Fig. 2),
226 and at least one of them (*mOR5*) in olfactory receptor cells (Fig. 5D), suggesting that these
227 genes encode functional (i.e., "true") mORs. As these hagfish mORs form a monophyletic
228 group together, these genes are considered to have duplicated through hagfish-specific
229 diversification.

230 Nevertheless, no hagfish mOR genes were classified as type 2 subfamily. In contrast,
231 the lamprey has both type 1 and type 2 subfamilies (this study; Niimura 2009), suggesting
232 that the common ancestor of all vertebrates had both types, but hagfish has lost type 2 mORs.

233

234 *V1Rs*

235 In this study, we found only two hagfish V1R genes, which show high affinity to the ORA3/4
236 subfamily and ORA5 subfamily, respectively (Fig. 1B). We also demonstrated that these
237 genes are expressed in the olfactory organ, and at least one of them (*V1R1*) is expressed
238 particularly in olfactory receptor cells (Fig. 4A, 5E). These results suggest that hagfish have
239 true V1Rs.

240 Still, we identified more V1Rs (eight genes) from the latest version (Timoshevskaya
241 et al., 2023) of the sea lamprey genome. These genes include two additional ones that were

242 unreported in previous research (Kowatschew and Korsching 2022). Among them, two
243 (lamprey V1R1 and V1R2) show affinity to the gnathostome ORA1/2, two (lamprey V1R3
244 and V1R7) to the gnathostome ORA3/4, and one (lamprey V1R6) to ORA6. The remaining
245 three (lamprey V1R4, V1R5, and V1R8) were placed in the basal position of the gnathostome
246 ORA5/6, despite showing low support values. Combining these results with previous
247 research (Kowatschew and Korsching 2022), it is suggested that the common ancestor of all
248 vertebrates had at least three subtypes of V1Rs, ORA1/2, ORA3/4, and ORA5/6, and that the
249 hagfish has lost the ORA1/2 gene(s).
250
251 *V2Rs*
252 Most notably, we revealed that V2Rs show the highest diversification among hagfish
253 olfactory receptors. We identified typical V2Rs (135 genes) that form a monophyletic clade
254 basal to the gnathostome V2Rs. As the closest outgroup to these V2Rs, we found a pair of
255 V2R-like genes in hagfish and lamprey. Consistent with a previous suggestion that lamprey
256 V2Rs do not function as olfactory receptors (Kowatschew and Korsching 2022), our
257 transcriptome analysis indicates that the hagfish V2R-like gene is not expressed in the
258 olfactory organ at a functional level (Fig. 3B).
259 In contrast, we confirmed that all of the true V2Rs are expressed in the olfactory
260 organ (Fig. 3B), particularly and at least one of them (*V2R33*) in olfactory receptor cells (Fig.
261 5F), suggesting that these genes do serve as olfactory receptors. This finding challenges the
262 conventional view that the olfactory function of the V2Rs was established in the gnathostome
263 lineage (Kowatschew and Korsching 2022; Zhang et al. 2022). Instead, we suggest that true
264 V2Rs had already been acquired in the common ancestor of all vertebrates; these genes
265 underwent extensive diversification in hagfish, but they have been lost specifically in the
266 lamprey lineage.

267 *TAARs*

268 We found no TAAR gene candidates in the current version of the hagfish genome, consistent
269 with previous research (Guo et al., 2022). In contrast, previous studies (Libants et al. 2009;
270 Dieris et al. 2021) reported that the lamprey has dozens of TAARs. Therefore, it is suggested
271 that TAARs can be traced back to the common ancestor of all vertebrates and that they have
272 been lost specifically in the hagfish lineage.

273

274 **Importance of assembly quality for comparative genomics**

275 To obtain a more complete gene list for comparison, we re-surveyed lamprey ORs using the
276 latest version (Timoshevskaya et al., 2023) of the sea lamprey genome. As a result, we
277 identified a greater number of mOR and V1R genes than previously reported (Table 1),
278 suggesting that better genome assemblies provide more reliable information for comparative
279 genomics.

280 As the current version of the *E. burgeri* genome used in this study is not completely
281 assembled at the telomere-to-telomere level (Yu et al., 2024), future high-quality assemblies
282 may reveal additional OR genes that are not detected in this study. In addition, if a complete
283 genome assembly were available, our analysis of the genomic distribution of hagfish OR
284 genes would be more precise. Still, the overall trends in OR gene diversification, loss, and
285 genomic distribution are clearly demonstrated.

286

287 **Gene duplication of hagfish ORs**

288 Our analysis of the genomic mapping of hagfish OR genes revealed that phylogenetically
289 close mOR and V2R genes tend to reside close in the hagfish genome. Particularly in the
290 scaffold FYBX03000002.1, phylogenetically close V2R genes were organized in clusters
291 along the genome. However, a given genomic cluster often contains genes from different

292 clades (Fig. 2B). This distribution pattern suggests that these genes have been diversified
293 through tandem gene duplications and chromosomal rearrangements, as has been observed in
294 human OR gene evolution (Niimura and Nei 2005).
295 It is suggested that whole-genome duplication (WGD) may contribute to the genetic
296 diversity of the ORs (Wang et al. 2021). Recently, the cyclostomes were proposed to have
297 experienced lineage-specific WGD or triplication (Marlétaz et al., 2024; Yu et al., 2024).
298 Nevertheless, this event appears to have had no or a small contribution to the evolution of
299 ORs, especially V2Rs. In fact, the scaffold FYBX03000002.1 includes most of the hagfish
300 V2R genes (91%), indicating that these genes have duplicated at the intrachromosomal level.
301 Another possible mechanism is retrotransposon-mediated duplication. However, this
302 scenario is unlikely in the diversification of hagfish V2Rs because retrotransposon-encoded
303 reverse transcriptase typically generates intronless genes (Lallemand 2020), and all hagfish
304 V2Rs contain introns except for a few exceptions whose domains were only partially
305 predicted (Supplementary File 1).
306 Taken together, the most plausible mechanism for the extensive diversification of
307 hagfish V2Rs is tandem gene duplications followed by chromosomal rearrangements.
308
309 **Expression specificity of the ORs**
310 In mammals, it is widely believed that each olfactory receptor neuron expresses only one OR
311 gene (Bystrova & Kolesnikov, 2021). This “one neuron–one receptor” hypothesis was
312 originally proposed based on studies in mice and rats (e.g., Chess et al., 1994; Malnic et al.,
313 1999). Nevertheless, it has yet to be conclusively demonstrated (Mombaerts, 2004). In fact, a
314 previous study showed that two functional odor receptors are expressed in one neuron in the
315 fruit fly (Goldman et al., 2005).

316 Particularly, whether the “one neuron–one receptor” rule also applies to basal
317 vertebrates remains an open question. The broad expression pattern of the V1R gene in the
318 hagfish olfactory epithelium (Fig. 5E) imply the possibility that V1Rs are coexpressed with
319 other ORs in a single receptor neuron. To test this hypothesis in cyclostomes, further
320 analyses—such as single-cell RNA-seq—are necessary.
321
322 **Cell morphology and distribution of the OR-expressing neurons**
323 Certain vertebrates exhibit lineage-specific distribution patterns of olfactory receptor cells.
324 For example, mammals have two distinctive olfactory organs: the main olfactory organ and
325 the vomeronasal organ. The former consists of mOR- and TAAR-expressing ciliated cells,
326 while the latter is specialized for pheromone detection and contains V1R- and V2R-
327 expressing—and additionally in rodents (Boillat et al., 2021), formyl peptide receptor (FPR)-
328 expressing—microvillous cells (Poncelet and Shimeld, 2020). In teleost fish, there are four
329 types of olfactory receptor cells: mOR- and TAAR-expressing ciliated cells, V1R- and V2R-
330 expressing microvillous cells, V1R-expressing crypt neurons, and kappe neurons with
331 unidentified receptors (Poncelet & Shimeld, 2020). Among them, the mOR-expressing
332 ciliated cells are found in the deep layer of the olfactory epithelium, while the V1R/V2R-
333 expressing microvillous cells, along with crypt neurons, are located in the superficial layer
334 (Hansen et al., 2004; Sato et al., 2005).
335 In hagfish, ciliated and microvillous cell types have been reported to be present
336 (Døving, 1998). However, the relationship between these cell types and OR expression
337 patterns, as well as the spatial distribution of olfactory receptor cells, remain unresolved.
338 Although our gene expression analysis suggests that both mOR- and V2R-expressing cells
339 are located in the relatively superficial layer (Fig. 5D, F), further detailed analyses, including
340 spatial RNA-seq, would provide key information for these questions.

341 **Lineage-specific diversification of olfaction in early vertebrates**

342 One of the most notable findings of this study is that functional V2Rs are present and
343 extensively diversified in the cyclostome hagfish. While V2Rs are reported to bind to water-
344 soluble molecules such as the peptide ligands of MHC class I molecules and exocrine gland
345 peptides (ESPs) in mammals (Leinders-Zufall et al. 2004; Kimoto et al. 2005), they are also
346 expected to detect amino acids and their derivatives, eliciting feeding behaviors in teleosts
347 (Speca 1999; Koide et al., 2009; Sato et al., 2005). The massive diversification of hagfish
348 V2Rs may thus reflect an adaptation to scavenging life in low-light environments, with
349 various functions such as the detection of food sources and mating partners. To determine the
350 functions of hagfish V2Rs, it is necessary to examine these receptors from multiple
351 disciplinary perspectives, including ligand binding assays, electrophysiological recordings,
352 and behavioral experiments using candidate odorants.

353 Furthermore, our phylogenetic analyses revealed that the OR repertoires in basal
354 vertebrates have evolved in a mosaic manner (Fig. 6). On the one hand, hagfish have
355 experienced extensive diversification of V2Rs and loss of TAARs. On the other hand,
356 lampreys appear to have lost true V2Rs. As previous studies have relied on lampreys only, it
357 has been mistakenly suggested that true V2Rs originated in the common ancestor of
358 gnathostomes. However, this study pushes the evolutionary origin of the V2Rs back to the
359 common ancestor of all vertebrates. This finding calls attention to the importance of hagfish,
360 which has often been overlooked, yet holds significant potential to illuminate critical aspects
361 of early vertebrate evolution.

362

363 **Conclusion**

364 In this study, we conducted a comprehensive genome-wide survey and gene expression
365 analyses of ORs in hagfish, a basal vertebrate that typically inhabits the deep sea. We

366 identified 48 mORs, two V1Rs, 135 typical V2Rs, and no TAARs in the inshore hagfish (*E.*
367 *burgeri*). Phylogenetic analysis revealed that all mORs belong to the type 1 subfamily. The
368 two V1Rs form a clade with known vertebrate ORA subfamilies, and the 135 V2Rs form a
369 distinct monophyletic clade basal to gnathostome V2Rs. Genomic mapping revealed large
370 clusters of the hagfish V2Rs on a single scaffold, likely resulting from tandem gene
371 duplications and chromosome rearrangements. Gene expression profiling confirmed that
372 nearly all of these receptors are specifically expressed in the olfactory organ, and *in situ*
373 hybridization showed that at least some are localized in olfactory receptor cells.
374 These findings revise the current evolutionary scenario of vertebrate olfactory
375 systems. While lampreys lack typical V2Rs, our data indicate that these receptors were
376 already present in the common ancestor of vertebrates and were subsequently lost in the
377 lamprey lineage. The large-scale expansion of V2Rs in hagfish likely reflects ecological
378 adaptation to their scavenging niche in dark environments. Our results elucidate the
379 previously unrecognized evolutionary diversity in the olfactory system of basal vertebrates,
380 highlighting hagfish as a key animal for understanding the early vertebrate evolution.
381

382 **Acknowledgment**

383 We thank Drs. Hiroshi Wada and Yoshiaki Morino for their valuable comments, and Dr.
384 Haruka Ozaki for providing computer resources. We appreciate Susumu Hatanaka (Shinsho
385 Maru; Fujisawa, Kanagawa, Japan) and Masanori Nishizaki and Dr. Masa-aki Yoshida (Oki
386 Marine Biological Station, Shimane University) for providing the *E. burgeri* specimens.
387 Computations were partially performed on the NIG supercomputer at ROIS National Institute
388 of Genetics. This work was supported by the Grant-in-Aid for the Japan Society for the
389 Promotion of Science (JSPS; Grant Numbers JP20K15855, JP22K15164, JP24K09556, and

390 JP24H01538 to D.G.S, and JP19K16178 to Y.Y.) and by the Sasakawa Scientific Research
391 Grant from The Japan Science Society (Grant Number 2023-4098 to H.K.).
392

393 Materials and methods

394 Animals

395 Inshore hagfish (*E. burgeri*) specimens, both male and female, were collected from Sagami
396 Bay, Kanagawa Prefecture in 2021–2025 or from Kamo Bay near Oki Marine Biological
397 Station, Shimane University (Oki, Shimane, Japan) in 2021. As for the animals from Sagami
398 Bay, they were transported to the laboratory at the University of Tsukuba, Japan, where they
399 were kept in 160 L tanks filled with circulating artificial seawater at 12°C ± 0.5 °C before
400 dissection. As for the animals from Kamo Bay, they were transported to a laboratory at
401 Matsue campus, Shimane University (Matsue, Shimane, Japan). They were maintained in a
402 260 L tank filled with circulating artificial seawater at 15 ± 0.5 °C and sacrificed for tissue
403 sampling. All procedures in this study were performed in compliance with the guidelines for
404 animal use of the Animal Care Committees at the University of Tsukuba and Shimane
405 University (specific approval is not required for experimentation on fishes under the Japanese
406 law, Act on Welfare and Management of Animals). During the investigation, every effort was
407 made to minimize suffering and to reduce the number of animals used.
408

409 Tissue sampling

410 For RNAseq of the olfactory organ, the tissue was collected from the decapitated animal
411 following ice anesthesia, and was frozen using liquid nitrogen and kept at −80 °C until
412 analyzed. For molecular cloning, histological analysis, and *in situ* hybridization, animals
413 were deeply anesthetized with MS-222 (100 mg/L; Sigma, A5040) and euthanized by

414 decapitation. Then, the decapitated heads were dissected to harvest the olfactory organs,
415 which were served for each experiment described below.
416
417 **RNAseq and differential gene expression analysis**
418 RNA-seq of the olfactory organ was newly performed in this study, following the method for
419 other organs described previously (Yamaguchi et al., 2023). Total RNAs were extracted from
420 the frozen tissue from a single individual (total length 55 cm, weight 226.3 g, male; as for
421 other tissues, the pituitary glands from five individuals of mixed sexes and sizes were pooled,
422 while other tissues were collected from a single individual, as reported by Yamaguchi et al.,
423 2023). Raw RNA-Seq data are available at the public repository (BioProject accession ID:
424 PRJDB35770, DRR Run: DRR707969–DRR707972, DRR707975-DRR707977,
425 DRR707979, and DRR707980).
426 The adaptor sequences and low-quality sequences were then removed using
427 Trimmomatic version 0.38 (Bolger et al., 2014). We used FastQC version 0.11.9 (Andrews,
428 2010) to evaluate the sequence data quality. The gene models of *E. burgeri* were retrieved
429 from Ensembl Release 104 and were used as a reference for calculating expression levels. To
430 prepare a more comprehensive reference dataset, we added sequences identified from the
431 genome by TBLASTN searches. The expression amount was quantified as transcripts per
432 kilobase million (TPM) using Salmon version 1.10.3 (Patro et al., 2017). For visualizing the
433 expression amount of olfactory receptor genes, we used the tidyverse and ggplot2 packages in
434 R. TPMs were log10-transformed after adding a pseudocount of 0.001.

435 **Gene search**

436 To identify main olfactory receptor (mOR) gene candidates, we performed BLASTP searches
437 on the hagfish and sea lamprey (*P. marinus*) gene models using mOR sequences of
438 gnathostomes as queries. The gnathostomes included the following species: human, mouse,
439 zebrafish, and African clawed frog. We retrieved hagfish gene models and genome sequences
440 from ENSEMBL (Release 104). As a new version of the sea lamprey genome was reported
441 recently (Timoshevskaya et al., 2023), we re-surveyed mORs of this animal. The amino acid
442 (AA) sequences of the gnathostomes were retrieved from Niimura and Nei (2005). We
443 extracted hagfish and sea lamprey sequences from the BLAST search result with E values <
444 1E-3 and an alignment length of ≥ 250 AA. To confirm the homology, we also conducted
445 BLASTP searches against the human, zebrafish, and African clawed frog gene models
446 (ENSEMBL release 104) using the hagfish and sea lamprey sequences as queries with the
447 same thresholds. For gene extraction, we used SeqKit ver. 2.8.2 (Shen et al., 2024). To
448 narrow down the list to more plausible sequences, we performed a domain search using
449 HMMER 3 (ver. 3.2.1) with the Pfam database (ver. 34.0). As a result, we found 35 genes
450 from the hagfish and 61 genes from the sea lamprey gene models.

451 Recently, a new version of the *E. burgeri* genome was published (Yu et al., 2024), but
452 gene models based on this genome data are not officially provided. To extract potential mOR
453 gene candidates that were missed in the gene models based on the old version, we searched
454 for mOR candidate genes directly from the new one. To identify potential mOR gene
455 candidates, TBLASTN was performed on the hagfish genome version 4.0
456 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_900186335.3/). We used the same
457 query sequences as the analysis using gene models. The threshold for homologous sequences
458 was E value < 1E-10 and alignment length ≥ 250AA as described in the previous research
459 (Niimura and Nei, 2005). To confirm their homology, we also conducted BLASTP searches

460 as described above. The coding sequences were predicted from TBLASTN hit regions with
461 900 bp of flanking sequences using GeneWise (version 2.4.1) with default parameters. We
462 used the same query sequences for this TBLASTN search as those used for the BLASTP
463 search described above. Next, the possibly functional sequences were extracted from the
464 predicted sequences using HMMER 3 and the Pfam database. 29 sequences were matched to
465 the mOR candidate genes extracted from the gene models, and 13 sequences were newly
466 identified from the genome sequences.

467 We collected candidate sequences of hagfish V1R and V2R genes in the same way as
468 for mORs. In the reciprocal BLAST search of gene models, homologous genes of V2R,
469 CaSR, taste receptor type 1 (TAS1R), and G protein-coupled receptor family C group 6
470 (GPRC6) were extracted as the candidate sequences for V2R. The plausible sequences were
471 extracted with at least 200 AA of a 7-transmembrane domain (PF00003, TM7_3) by domain
472 searches for V2R. As a result of BLAST searches and domain searches against the gene
473 models, two and 38 gene models were identified as candidates of V1R and V2R, respectively.
474 The one and 99 additional candidate sequences of V1R and V2R were found from the hagfish
475 genome in a similar way to identify mOR sequences. To predict coding sequences by
476 GeneWise, we used sequences found by TBLASTN against the TM7 domain. As V2Rs
477 contain another domain in their N-terminal region, we also included this domain and the
478 intron sequences for prediction if a TBLASTN hit sequence is found to contain it (on the
479 same strand and intersequences between the TBLASTN hits < 100,000 bp). In the other case,
480 that is, even if a TBLASTN hit sequence does not include the N-terminal region, we also
481 used any such sequence for our analysis because there is a possibility that its N-terminal
482 domain is just not detected. In addition, we identified 11 more plausible sequences that
483 correspond to the gene models. Thus, we used 110 predicted genes, including 11 sequences
484 for downstream analysis.

485 For lamprey V1R genes, we initially identified four genes from the sea lamprey (*P.*
486 *marinus*) gene models. While three genes were previously reported (Kowatschew and
487 Korsching, 2022), the other one (Pma-V1R7) was newly identified. Because the detected
488 genes were fewer than V1Rs reported in the previous study, we further conducted a
489 TBLASTN search against the sea lamprey genome. As a result, we detected seven V1R
490 candidate sequences (six known V1R genes and Pma-V1R7). Additionally, we searched V1R
491 sequences against the arctic lamprey (*L. camtschaticum*) genome and gene models. We
492 detected the seven genes, of which five correspond to the genes identified in the previous
493 research. Although another gene corresponded to Pma-V1R7, it was truncated. One of the
494 seven genes (Lca-V1R8) was newly identified, and we also found the orthologous gene
495 (Pma-V1R8) from the sea lamprey genome. In summary, we identified eight lamprey V1R
496 genes for the phylogenetic analysis.
497
498 **Phylogenetic analysis**
499 We retrieved mOR sequences of the gnathostomes and the amphioxus from Niimura (2009),
500 V1R and TAS2R sequences from the ENSEMBL database and previous research
501 (Kowatschew and Korsching, 2022, Sharma et al., 2019, Zapilko and Korsching, 2016, Date-
502 Ito et al., 2008), and V2R, TAS1R, GPRC6, CaSR, glutamate receptor metabotropic (GRM),
503 and gamma-aminobutyric acid type B receptor (GABBR) sequences from previously
504 published papers and the ENSEMBL database (Zhang et al., 2022, Kowatschew and
505 Korsching, 2022). These sequences were aligned using MAFFT version 7.481 (Katoh and
506 Standley 2013) with the default option and trimmed to poorly aligned regions using TrimAL
507 version 1.4. rev15 (Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009) with the
508 "gappyout" option. Phylogenetic analysis was conducted using RAxML version 8.2.12

509 (Stamatakis 2014) with the options of "-f a -x 12345 -p 12345 -# 500 -m
510 PROTGAMMAAUTO --auto-prot=aic -T 16."
511
512 **Genomic mapping and distribution analysis**
513 To classify subgroups of mOR genes and V2R genes, we focused on clades containing at
514 least five genes with their bootstrap value ≥ 85%. We aimed to obtain three or more clades
515 for mORs and V2Rs, respectively, trying to define as larger clades as possible at the same
516 time. Consequently, we obtained three clades for mORs and nine clades for V2Rs
517 (Supplementary Fig.1 and 2). The V2R genes that were not classified into any of these clades
518 were categorized into "Else" (Supplementary 2). We then visualized the genomic distribution
519 of the OR genes using ggbio version 1.46.0 within R version 4.2.2 (R Center For Statistical
520 Computing, Vienna, Austria), based on this classification.
521 For scatter plotting, we calculated the phylogenetic and genomic distances for all
522 interclade and intraclade pairs of the mOR genes and of the V2R genes, respectively, based
523 on the clade classification. We included the pairs of "Else" V2R genes into the interclade
524 pairs. The phylogenetic distances were obtained using the "cophenetic.phylo" function in the
525 ape version 5.7.1. Accordingly, we visualized the relationship between the genomic and
526 phylogenetic distances using the tidyverse version 2.0.0 and ggplot2 version 3.5.2.
527
528 **Molecular cloning and probe synthesis**
529 Total RNA was extracted from the harvested olfactory organs of adult *E. burgeri* using
530 TRIzol Reagent (Invitrogen, 15596026). These RNAs were reverse transcribed into cDNAs
531 using PrimeScript II 1st strand cDNA Synthesis Kit (Takara, 6210A) and were used as
532 templates in the polymerase chain reaction (PCR). Hagfish genes of interest were amplified
533 by PCR using the primers listed in Supplementary Table 1. We used reverse primers

534 including the T3 promoter sequence (20 bp) to synthesize DIG-labeled RNA probes for *in*
535 *situ* hybridization. As a negative control, we instead used forward primers including the T7
536 promoter sequence (20 bp) to synthesize DIG-labeled RNA probes for *in situ* hybridization
537 (sequences not shown). DIG-labeled RNA probes were transcribed from PCR products using
538 T3 and T7 RNA polymerase (Roche, 11031163001 and 12352204, respectively).
539
540 **Hematoxylin-eosin (HE) staining and section *in situ* hybridization**
541 To perform histological analysis and *in situ* hybridization, the harvested olfactory organs
542 were fixed with 4% PFA/PBS for 1 hour at room temperature (RT) or overnight at 4 C°. The
543 fixed specimen was then rinsed with PBS, dehydrated in a graded methanol series (25%,
544 50%, 75%, and 100%), and stored at −20 °C.
545 For cryosectioning, the stored samples were rehydrated in a graded methanol series
546 (75%, 50%, and 25% in PBS) and PBS. They were replaced with 10% and then 30% sucrose.
547 Subsequently, the specimens were embedded in Tissue-Tek Optimum Cutting Temperature
548 (O.C.T.) Compound (Sakura Finetek Japan, 4583). Frozen sections in 20 μm were prepared
549 using a cryostat (Leica, CM1860) and mounted on MAS-coated slide glasses (Matsunami
550 Glass Ind., SMAS-01).
551 HE staining was performed according to the conventional method after sections were
552 washed with PBS to remove the O.C.T. compound. Namely, the sections were incubated in
553 Mayer's hematoxylin solution (Wako 131-09665) for 15 min and then rinsed with running tap
554 water for 30 min. Afterward, the sections were incubated in 1% Eosin Y solution (Wako 051-
555 06515). The stained sections were dehydrated in a graded methanol series (25%, 50%, 75%,
556 95%, and 100%) and xylene-equivalent solvent G-NOX (Genostuff, GN04), and mounted
557 using xylene-free mounting medium PARAmount-D (Falma, 308-500-1).

558 For *in situ* hybridization, the sections were first washed with PBS to remove the
559 O.C.T. compound and postfixed for 20 min with 4% PFA/PBS and rinsed with PBS.
560 Afterward, they were prehybridized in hybridization buffer (50% Formamide, 5× SSC, 5×
561 Denhardt's solution, 50 μg/ml yeast tRNA, and 50 μg/ml heparin in deionized-distilled H_2O)
562 and then incubated in the hybridization buffer with 0.1 mg/ml probe overnight at 60 °C.
563 Subsequently, the specimens were washed with the following solutions: 50% Formamide and
564 2× SSC (for 30 min at 60 °C, twice), 2× SSC (for 30 min at 60 °C, twice), 0.2× SSC (for 30
565 min at 60 °C, twice), and PBS (5 min at RT). For signal detection, the specimens were
566 blocked with 0.5% blocking reagent (Roche, 11277073910) in PBS for 1 hour at RT and then
567 incubated with 0.5 μl/ml alkaline phosphatase (AP)-conjugated anti-digoxigenin Fab
568 fragments (diluted in 0.5% blocking reagent/PBS; Roche, 11093274910) overnight at 4 °C.
569 After washing with tris-buffered saline (TBS) four times for 30 min each at RT, alkaline
570 phosphatase signals were detected with 20 μl/ml NBT/BCIP in color development buffer
571 (100 mM Tris HCl pH 9.5, 100 mM NaCl, and 50 mM $MgCl_2$). The stained specimens were
572 postfixed in 4% PFA/PBS, dehydrated in a graded methanol series (25%, 50%, 75%, 95%,
573 and 100%) and G-NOX, and mounted using PARAmount-D. The prepared slides were
574 examined under a microscope (Nikon, ECLIPSE Ni) and photographed using a microscope
575 digital camera (Nikon, DS-Ri1).
576

577 **Reference**

578 Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.
579 Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
580 Bi X, Wang K, Yang L, Pan H, Jiang H, Wei Q, Fang M, Yu H, Zhu C, Cai Y, He Y, Gan X,
581 Zeng H, Yu D, Zhu Y, Jiang H, Qiu Q, Yang H, Zhang YE, Wang W, Zhu M, He S,

582 Zhang G (2021) Tracing the genetic footprints of vertebrate landing in non-teleost
583 ray-finned fishes. *Cell* **184**(5):1377–1391.e14. doi:10.1016/j.cell.2021.01.046.
584 Boillat M, Carleton A, Rodriguez I (2021) From immune to olfactory expression:
585 neofunctionalization of formyl peptide receptors. *Cell Tissue Res* **383**(1):387–393.
586 doi:10.1007/s00441-020-03393-5.
587 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina
588 sequence data. *Bioinformatics* **30**(15):2114–2120. doi:10.1093/bioinformatics/btu170.
589 Bystrova MF, Kolesnikov SS (2021) The "one neuron–one receptor" rule in the physiology
590 and genetics of olfaction. *Neurosci Behav Physiol* **51**:1008–1017.
591 doi:10.1007/s11055-021-01159-2.
592 Chess A, Simon I, Cedar H, Axel R (1994) Allelic inactivation regulates olfactory receptor
593 gene expression. *Cell* **78**(5):823–834. doi:10.1016/s0092-8674(94)90562-2.
594 Date-Ito A, Ohara H, Ichikawa M, Mori Y, Hagino-Yamagishi K (2008) Xenopus V1R
595 vomeronasal receptor family is expressed in the main olfactory system. *Chem Sense*
596 **33**(4):339–346. doi:10.1093/chemse/bjm090.
597 Dieris M, Kowatschew D, Korsching SI (2021) Olfactory function in the trace amine-
598 associated receptor family (TAARs) evolved twice independently. *Sci Rep* **11**:7807.
599 doi:10.1038/s41598-021-87236-5.
600 Døving KB (1998). The Olfactory System of Hagfishes. In: Jørgensen JM, Lomholt JP,
601 Weber RE, Malte H (eds.) *The Biology of Hagfishes.* Springer, pp. 533–540.
602 doi:10.1007/978-94-011-5834-3_33.
603 Fernholm B, Holmberg K (1975) The eyes in three genera of hagfish (*Eptatretus*,
604 *Paramyxine* and *Myxine*)─a case of degenerative evolution. *Vision Res* **15**(2):253–
605 259. doi:10.1016/0042-6989(75)90215-1.

606 Goldman AL, Van der Goes van Naters W, Lessing D, Warr CG, Carlson JR. Coexpression
607 of two functional odor receptors in one neuron. *Neuron* **45**(5):661–666.
608 doi:10.1016/j.neuron.2005.01.025.
609 Guo L, Dai W, Xu Z, Liang Q, Miller ET, Li S, Gao X, Baldwin MW, Chai R, Li Q (2022)
610 Evolution of brain-expressed biogenic amine receptors into olfactory trace amine-
611 associated receptors. *Mol Biol Evol* 2022 **39**(3):msac006. doi:
612 10.1093/molbev/msac006. Erratum in: *Mol Biol Evol* **39**(5):msac072.
613 doi:10.1093/molbev/msac072.
614 Hansen A, Anderson KT, Finger TE (2004) Differential distribution of olfactory receptor
615 neurons in goldfish: structural and molecular correlates. *J Comp Neurol* **477**(4):347–
616 359. doi:10.1002/cne.20202.
617 Hara TJ (1975) Olfaction in fish. *Prog Neurobiol* **5**(4):271–335. doi:10.1016/0301-
618 0082(75)90014-3.
619 Hughes GM, Boston ESM, Finarelli JA, Murphy WJ, Higgins DG, Teeling EC (2018) The
620 birth and death of olfactory receptor gene families in mammalian niche adaptation.
621 *Mol Biol Evol* **35**(6):1390–1406. doi:10.1093/molbev/msy028.
622 Hussain A, Saraiva LR, Korsching SI (2009) Positive Darwinian selection and the birth of an
623 olfactory receptor clade in teleosts. *PNAS* **106**(11):4313–4318.
624 doi:10.1073/pnas.0803229106.
625 Kimoto H, Haga S, Sato K, Touhara K (2005) Sex-specific peptides from exocrine glands
626 stimulate mouse vomeronasal sensory neurons. *Nature* **437**:898–901.
627 doi:10.1038/nature04033.
628 Koide T, Miyasaka N, Morimoto K, Asakawa K, Urasaki A, Kawakami K, Yoshihara Y
629 (2009) Olfactory neural circuitry for attraction to amino acids revealed by transposon-

630 mediated gene trap approach in zebrafish. *PNAS* **106**(24):9884–9889.
631 doi:10.1073/pnas.0900470106.
632 Kowatschew D, Korsching SI (2022) Lamprey possess both V1R and V2R olfactory
633 receptors, but only V1Rs are expressed in olfactory sensory neurons. *Chem Senses*
634 **47**:bjac007. doi:10.1093/chemse/bjac007.
635 Lallemand T, Leduc M, Landès C, Rizzon C, Lerat E (2020) An overview of duplicated gene
636 detection methods: why the duplication mechanism has to be accounted for in their
637 choice. *Genes* **11**(9):1046. doi:10.3390/genes11091046.
638 Libants S, Carr K, Wu H, Teeter JH, Chung-Davidson Y-W, Zhang Z, Wilkerson C, Li W
639 (2009) The sea lamprey *Petromyzon marinus* genome reveals the early origin of
640 several chemosensory receptor families in the vertebrate lineage. *BMC Evol Biol*
641 **9**:180. doi:10.1186/1471-2148-9-180.
642 Leinders-Zufall T, Brennan P, Widmayer P, S PC, Maul-Pavicic A, Jäger M, Li XH, Breer H,
643 Zufall F, Boehm T (2004) MHC class I peptides as chemosensory signals in the
644 vomeronasal organ. *Science* **306**(5698):1033–1037. doi: 10.1126/science.1102818.
645 Malnic B, Hirono J, Sato T, Buck LB. Combinatorial receptor codes for odors. *Cell*
646 **96**(5):713–723. doi:10.1016/s0092-8674(00)80581-4.
647 Marinelli W, Strenger A. (1956). *Vergleichende Anatomie und Morphologie der Wirbeltiere,*
648 *Band 2. Myxine glutinosa.* Franz Deuticke.
649 Marlétaz F, Timoshevskaya N, Timoshevskiy VA, Parey E, Simakov O, Gavriouchkina D,
650 Suzuki M, Kubokawa K, Brenner S, Smith JJ, Rokhsar DS. The hagfish genome and
651 the evolution of vertebrates. *Nature* **627**(8005):811–820. doi:10.1038/s41586-024-
652 07070-3.

Mombaerts P (2004) Odorant receptor gene choice in olfactory sensory neurons: the one receptor-one neuron hypothesis revisited. *Curr Opin Neurobiol* **14**(1):31–36. doi:10.1016/j.conb.2004.01.014.

Muramatsu B, Suzuki DG, Suzuki M, Higashiyama H (2024) Gross anatomy of the Pacific hagfish, *Eptatretus burgeri*, with special reference to the coelomic viscera. *Anat Rec* **307**(1):155–171. doi:10.1002/ar.25208.

Niimura Y (2009) On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. *Genome Biol Evol* **1**:34–44. doi:10.1093/gbe/evp003.

Niimura Y, Nei M (2005) Evolutionary changes of the number of olfactory receptor genes in the human and mouse lineages. *Gene* **346**:23–28. doi:10.1016/j.gene.2004.09.027.

Nikaido M (2019) Evolution of V1R pheromone receptor genes in vertebrates: diversity and commonality. *Genes Genet Syst* **94**(4):141–149. doi:10.1266/ggs.19-00009.

Patro R, Duggal G, Love M, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**:417–419. doi:10.1038/nmeth.4197.

Poncelet G, Shimeld SM (2020) The evolutionary origins of the vertebrate olfactory system. *J Comp Neurol* **528**(13):2239–2253. doi:10.1002/cne.24887.

Rowe TB, Macrini TE, Luo Z-X (2011) Fossil evidence on origin of the mammalian brain. *Science* **332**:955–957. doi:10.1126/science.1203117.

Saraiva LR, Korsching SI (2007) A novel olfactory receptor gene family in teleost fish. *Genome Res* **17**(10):1448–1457. doi:10.1101/gr.6553207.

Sato Y, Miyasaka N, Yoshihara Y (2005) Mutually exclusive glomerular innervation by two distinct types of olfactory sensory neurons revealed in transgenic zebrafish. *J Neurosci* **25**(20):4889–4897. doi:10.1523/JNEUROSCI.0679-05.2005.

678 Shen W, Sipos B, Zhao L (2024) SeqKit2: A Swiss army knife for sequence and alignment
679 processing. *Imeta* **3**(3):e191. doi: 10.1002/imt2.191.
680 Sharma K, Syed AS, Ferrando S, Mazan S, Korsching SI (2019) The chemosensory receptor
681 repertoire of a true shark is dominated by a single olfactory receptor family. *Genome*
682 *Biol Evol* **11**(2):398–405. doi:10.1093/gbe/evz002.
683 Speca DJ, Lin DM, Sorensen PW, Isacoff EY, Ngai J, Dittman AH (1999) Functional
684 identification of a goldfish odorant receptor. *Neuron* **23**(3):487–498. doi:
685 10.1016/s0896-6273(00)80802-8.
686 Suzuki DG (2021) Consciousness in jawless fishes. *Frontiers Syst Neurosci* **15**:751876.
687 doi:10.3389/fnsys.2021.751876.
688 Timoshevskaya N, Eşkut KI, Timoshevskiy VA, Robb SMC, Holt C, Hess JE, Parker HJ,
689 Baker CF, Miller AK, Saraceno C, Yandell M, Krumlauf R, Narum SR, Lampman
690 RT, Gemmell NJ, Mountcastle J, Haase B, Balacco JR, Formenti G, Pelan S, Sims Y,
691 Howe K, Fedrigo O, Jarvis ED, Smith JJ (2023) An improved germline genome
692 assembly for the sea lamprey Petromyzon marinus illuminates the evolution of
693 germline-specific chromosomes. *Cell Rep* **42**(3):112263.
694 doi:10.1016/j.celrep.2023.112263.
695 Wang H, Chen L, Dong C, Chen B, Li B, Li X, Xu P (2021) Genome-wide identification and
696 characterization of olfactory receptor genes in common carp (Cyprinus carpio). *Gene*
697 **777**:145468. doi:10.1016/j.gene.2021.145468.
698 Yamaguchi Y, Takagi W, Kaiya H, Konno N, Yoshida MA, Kuraku S, Hyodo S (2023)
699 Phylogenetic and functional properties of hagfish neurohypophysial hormone
700 receptors distinct from their jawed vertebrate counterparts. *Gen Comp Endocrinol*
701 **336**:114257. doi:10.1016/j.ygcen.2023.114257.

702 Young JM, Massa HF, Hsu L, Trask BJ (2009) Extreme variability among mammalian V1R
703 gene families. *Genome Res* **20**(1):10–18. doi:10.1101/gr.098913.109.
704 Yu D, Ren Y, Uesaka M, Beavan AJS, Muffato M, Shen J, Li Y, Sato I, Wan W, Clark JW,
705 Keating JN, Carlisle EM, Dearden RP, Giles S, Randle E, Sansom RS, Feuda R,
706 Fleming JF, Sugahara F, Cummins C, Patricio M, Akanni W, D'Aniello S, Bertolucci
707 C, Irie N, Alev C, Sheng G, de Mendoza A, Maeso I, Irimia M, Fromm B, Peterson
708 KJ, Das S, Hirano M, Rast JP, Cooper MD, Paps J, Pisani D, Kuratani S, Martin FJ,
709 Wang W, Donoghue PCJ, Zhang YE, Pascual-Anaya J. (2024) Hagfish genome
710 elucidates vertebrate whole-genome duplication events and their evolutionary
711 consequences. *Nat Ecol Evol* **8**(3):519–535. doi:10.1038/s41559-023-02299-z.
712 Zapilko V, Korsching SI (2016) Tetrapod V1R-like ora genes in an early-diverging ray-
713 finned fish species: the canonical six ora gene repertoire of teleost fish resulted from
714 gene loss in a larger ancestral repertoire. *BMC Genomics* **17**:83. doi:10.1186/s12864-
715 016-2399-6.
716 Zhang Z, Sakuma A, Kuraku S, Nikaido M (2022) Remarkable diversity of vomeronasal type
717 2 receptor (OlfC) genes of basal ray-finned fish and its evolutionary trajectory in
718 jawed vertebrates. *Sci Rep* **12**:6455 (2022). doi:10.1038/s41598-022-10428-0.
719

720 **Figure Legend**

721 **Figure 1.** Phylogenetic analysis of mOR, V1R, and V2R genes. (A) Phylogenetic tree of
722 mORs. White triangles indicate clades of type 1 mOR subgroups of gnathostomes, type 2
723 mOR subgroups of gnathostomes, amphioxus mORs, and outgroup genes. (B) Phylogenetic
724 tree of V1R. White triangles indicate subgroups of gnathostome V1R genes and outgroup
725 genes. (C) Phylogenetic tree of V2R. White triangles indicate subgroups of gnathostome V2R
726 genes, gnathostome CaSR genes, gnathostome TAS1R/GPRC6 genes, Ambulacraria CaSR

genes, GRMs, and outgroup genes. For all trees, red triangles and edges denote clades of hagfish genes, while blue triangles and edges mark clades of lamprey genes. The range of bootstrap support values is shown as the circles at each node (black > 75, 75 ≥ gray ≥ 50, white < 50). The numbers described on some nodes indicate exact bootstrap values. Scale bar represents the number of amino acid substitutions per site. For the list of genes used, see Supplementary File 3. For more detailed trees, see Supplementary File 4–6.

Figure 2. Genomic distribution of the hagfish OR genes. (A, B) Genomic mapping of the OR genes to the scaffolds larger than 0.1 Mbp (A) and to a single scaffold, FYBX03000002.1 (B). Each clade of mOR and V2R genes (for identification of clades, see text) is differently colored. Scale interval: 10 Mbp for (A). (C, D) Scatter plots between phylogenetic and genomic distances, which are calculated between each pairwise combination of the identified genes: mORs (C) and V2Rs (D).

Figure 3. Gene expression profile of mOR candidates across various organs. Expression levels are shown as TPM values, log-transformed. Genes on the y-axis are ordered according to the phylogenetic tree shown in Fig. 1. The mOR gene used for *in situ* hybridization analysis (hagfish mOR5, abbreviated as Ebu-mOR5) is boldfaced.

Figure 4. Gene expression profile of organ-specific differential expression analysis for V1R and V2R candidates across various organs. Expression levels are shown as TPM values, log-transformed. Genes on the y-axis are ordered according to the phylogenetic tree shown in Fig. 1. The genes used for *in situ* hybridization analysis (hagfish V1R1 and V2R33, abbreviated as Ebu-V1R1 and Ebu-V2R33, respectively) are boldfaced.

752 **Figure 5.** Histological analysis of the hagfish olfactory organ and gene expression patterns of
753 olfactory receptor candidates. (A) Overview of the hagfish olfactory organ and brain, dorsal
754 view. The skin and dorsal roof of the braincase are removed. The section plane for (B) is
755 indicated by the dashed line. (B) Cross-section of the olfactory organ, hematoxylin-eosin
756 (HE) staining. (C) Magnified image of (B), showing the layered structure of the olfactory
757 lamina. (D–E) Gene expression patterns of hagfish *mOR5* (D), *V1R1* (E), and *V2R33* (F).
758 Abbreviations: BC, basal cells; Die, diencephalon; Mes, mesencephalon; ORC, olfactory
759 receptor cell; Rho, rhombencephalon; Tel, telencephalon. The scale bar for (C–F) is shown in
760 (C).
761
762 **Figure 6**. Diversification of ORs in vertebrate evolution. Silhouettes from Phylopic (public
763 domain).
764
765 **Supplementary Figure 1.** Classification for mOR genes into Clade 1–3.
766
767 **Supplementary Figure 2.** Classification for V2R genes into Clade 1–9.
768
769 **Supplementary Figure 3.** Negative controls for *in situ* hybridization using sense probes of
770 *mOR5* (A), *V1R1* (B), and *V2R33* (C). The scale bar for (A–C) is shown in (A).

771 **Table**

772 **Table 1.** Comparison of OR repertoires in vertebrates.

	mOR	V1R	V2R(-like)	TAAR
Inshore hagfish (*Eptatretus burgeri*)	48[a]	2[a]	135[a] (1[a])	0[a]
Sea lamprey (*Petromyzon marinus*)	62[a]/27[b]	8[a]/4[b]/6[c]	0[b] (1[a,c])	28[b]/23[d]
Elephant shark (*Callorhinchus milii*)	8[e]	6[e]	30[f]	5[d]
Zebrafish (*Danio rerio*)	154[g]	7[h]	72[f]	18[d]
Western clawed frog (*Xenopus tropicalis*)	824[g]	15[h]	691[f]	5[i]
Mouse (*Mus musculus*)	1035[g]	239[j]	154[f]	9[i]

773

774 *Source:* (a) this study; (b) Libants et al. (2009); (c) Kowatschew & Korsching (2022); (d)

775 Dieris et al. (2021); (e) Sharma et al. (2019); (f) Zhang et al. (2022); (g) Niimura (2009); (h)

776 Saraiva & Korsching (2007); (i) Hussain et al. (2009); (j) Young et al. (2010).

777

778 **Supplementary Table 1**. Primer sequences for RNA probes.

Gene	Primers Forward	Reverse + T3 (in lowercase)	Ensembl ID
mOR5	CGATGTATTCTTCCTCAAAA CCATGTTCCA	attaaccctcactaaagggaATGGGGTTGCTGA TAGGAGG	ENSEBUT00000003975
V1R1	GATCATGCACCTGGCGTTGG TAAATCTGGT	attaaccctcactaaagggaCCGAGAGGCTTCA CATCTCC	ENSEBUT00000028488
V2R33	GCATGGAAGTATCGAACAA TCACAAAGCAT	attaaccctcactaaagggaAAATGAAGCACAG ACCTTGC	ENSEBUT00000023919

779

780 **Supplementary Files**

781 **Supplementary File 1.** Details of identified hagfish genes.

782

783 **Supplementary File 2.** Details of identified sea lamprey genes.

784

785 **Supplementary File 3.** Gene list used for phylogenetic analysis.

786 **Supplementary File 4.** Detailed phylogenetic tree of mOR genes.

787

788 **Supplementary File 5.** Detailed phylogenetic tree of V1R genes.

789

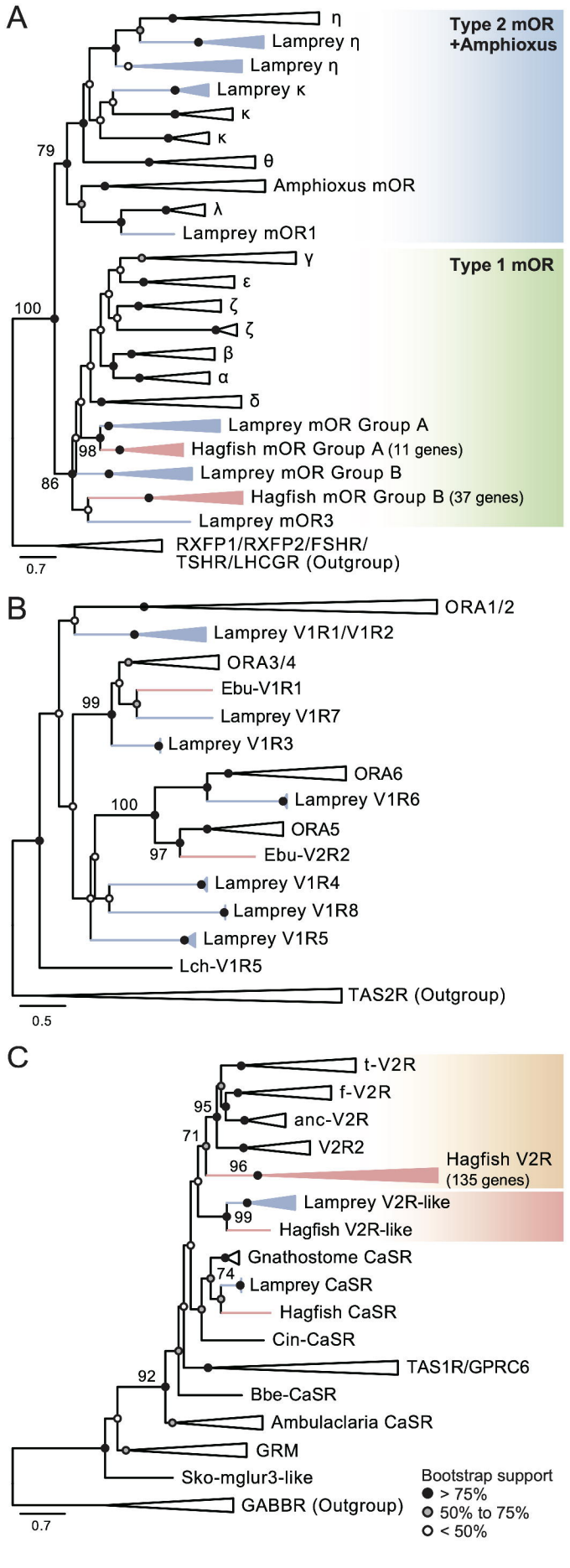
790 **Supplementary File 6.** Detailed phylogenetic tree of V2R genes.

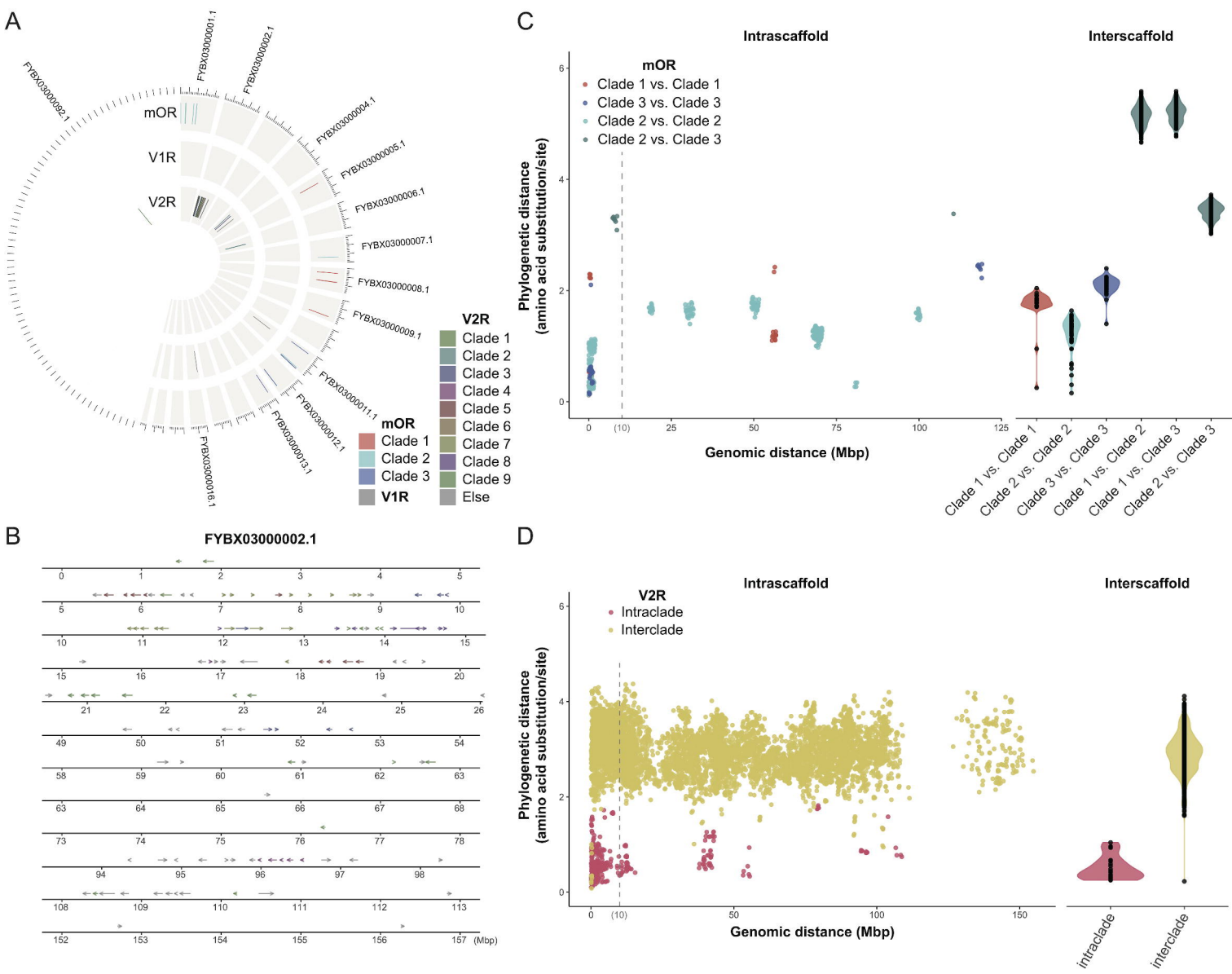
791

792 **Supplementary File 7.** Phylogenetic tree of known TAAR genes with their close GPCR

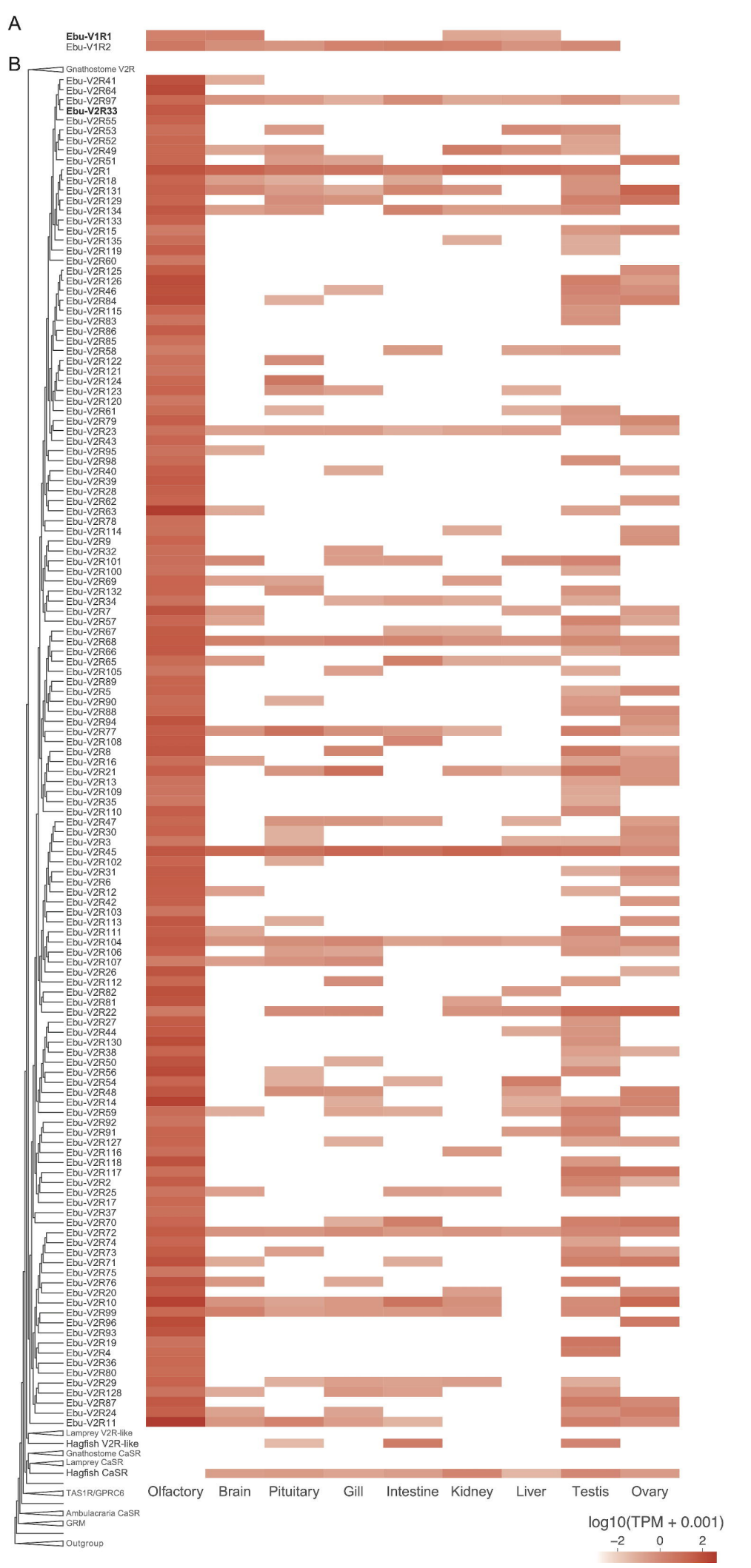
793 families, including the 5-hydroxytryptamine receptor 4 (5-HT4 or HTR4), adrenoceptor beta

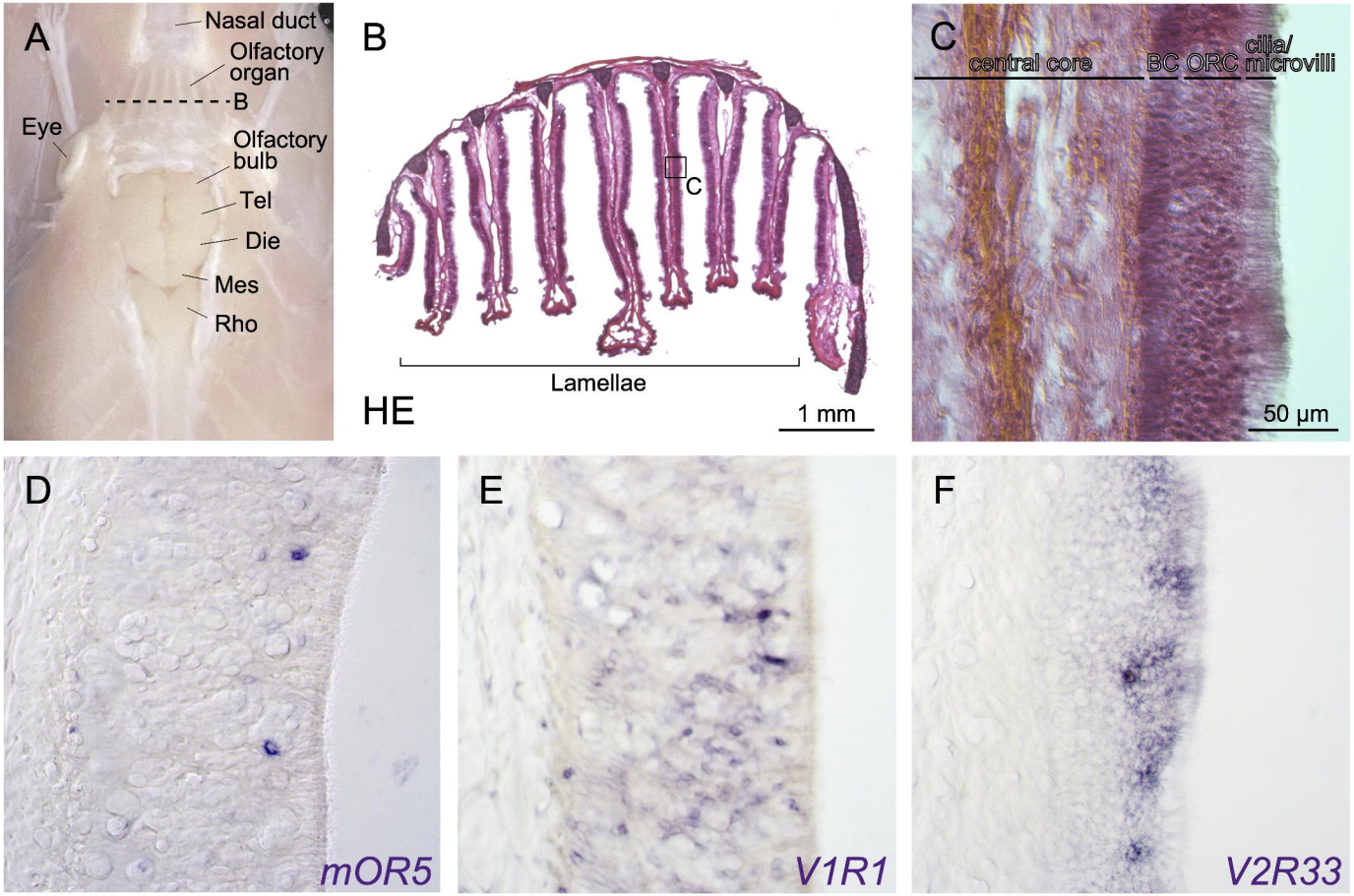
794 1, and dopamine receptor genes.

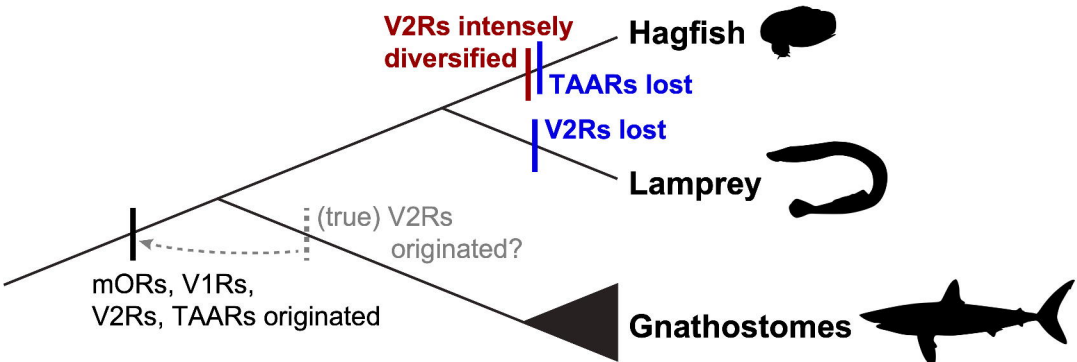

A
Type 2 mOR
+Amphioxus
η
Lamprey η
Lamprey η
Lamprey κ
κ
κ
79
θ
Amphioxus mOR
λ
Lamprey mOR1
Type 1 mOR
γ
ε
100
ζ
ζ
β
α
δ
Lamprey mOR Group A
98
Hagfish mOR Group A (11 genes)
86
Lamprey mOR Group B
Hagfish mOR Group B (37 genes)
Lamprey mOR3
RXFP1/RXFP2/FSHR/
TSHR/LHCGR (Outgroup)
0.7
B
ORA1/2
Lamprey V1R1/V1R2
ORA3/4
Ebu-V1R1
99
Lamprey V1R7
Lamprey V1R3
ORA6
100
Lamprey V1R6
ORA5
97
Ebu-V2R2
Lamprey V1R4
Lamprey V1R8
Lamprey V1R5
Lch-V1R5
TAS2R (Outgroup)
0.5
C
t-V2R
f-V2R
95
anc-V2R
71
V2R2
96
Hagfish V2R
(135 genes)
Lamprey V2R-like
99
Hagfish V2R-like
Gnathostome CaSR
74
Lamprey CaSR
Hagfish CaSR
Cin-CaSR
TAS1R/GPRC6
92
Bbe-CaSR
Ambulaclaria CaSR
GRM
Sko-mglur3-like
GABBR (Outgroup)
0.7
Bootstrap support
> 75%
50% to 75%
< 50%


A
mOR
V1R
V2R
FYBX03000001.1
FYBX03000002.1
FYBX03000004.1
FYBX03000005.1
FYBX03000006.1
FYBX03000007.1
FYBX03000008.1
FYBX03000009.1
FYBX03000011.1
FYBX03000012.1
FYBX03000013.1
FYBX03000016.1
FYBX03000092.1
V2R
Clade 1
Clade 2
Clade 3
Clade 4
Clade 5
Clade 6
Clade 7
Clade 8
Clade 9
Else
mOR
Clade 1
Clade 2
Clade 3
V1R
B
FYBX03000002.1
(Mbp)
C
Intrascaffold
Interscaffold
mOR
Clade 1 vs. Clade 1
Clade 3 vs. Clade 3
Clade 2 vs. Clade 2
Clade 2 vs. Clade 3
Phylogenetic distance
(amino acid substitution/site)
Genomic distance (Mbp)
Clade 1 vs. Clade 1
Clade 2 vs. Clade 2
Clade 3 vs. Clade 3
Clade 1 vs. Clade 2
Clade 1 vs. Clade 3
Clade 2 vs. Clade 3
D
Intrascaffold
Interscaffold
V2R
Intraclade
Interclade
Phylogenetic distance
(amino acid substitution/site)
Genomic distance (Mbp)
intraclade
interclade

Gnathostome type 2 & amphioxus
Gnathostome Type 1 mOR
Lamprey Group A
Ebu-mOR44
Ebu-mOR7
Ebu-mOR34
Ebu-mOR15
Ebu-mOR40
Ebu-mOR3
Ebu-mOR19
Ebu-mOR37
Ebu-mOR5
Ebu-mOR27
Ebu-mOR18
Lamprey Group B
Ebu-mOR14
Ebu-mOR23
Ebu-mOR28
Ebu-mOR35
Ebu-mOR20
Ebu0mOR45
Ebu-mOR33
Ebu-mOR4
Ebu-mOR47
Ebu-mOR21
Ebu-mOR46
Ebu-mOR24
Ebu-mOR12
Ebu-mOR39
Ebu-mOR38
Ebu-mOR16
Ebu-mOR36
Ebu-mOR10
Ebu-mOR43
Ebu-mOR42
Ebu-mOR41
Ebu-mOR13
Ebu-mOR25
Ebu-mOR22
Ebu-mOR31
Ebu-mOR11
Ebu-mOR29
Ebu-mOR1
Ebu-mOR17
Ebu-mOR32
Ebu-mOR30
Ebu-mOR2
Ebu-mOR48
Ebu-mOR8
Ebu-mOR6
Ebu-mOR26
Ebu-mOR9
Outgroup
Group A
Group B
Olfactory
Brain
Pituitary
Gill
Intestine
Kidney
Liver
Testis
Ovary
log10(TPM + 0.001)
−2
0
2


A
Ebu-V1R1
Ebu-V1R2
B
Gnathostome V2R
Ebu-V2R41
Ebu-V2R64
Ebu-V2R97
Ebu-V2R33
Ebu-V2R55
Ebu-V2R53
Ebu-V2R52
Ebu-V2R49
Ebu-V2R51
Ebu-V2R1
Ebu-V2R18
Ebu-V2R131
Ebu-V2R129
Ebu-V2R134
Ebu-V2R133
Ebu-V2R15
Ebu-V2R135
Ebu-V2R119
Ebu-V2R60
Ebu-V2R125
Ebu-V2R126
Ebu-V2R46
Ebu-V2R84
Ebu-V2R115
Ebu-V2R83
Ebu-V2R86
Ebu-V2R85
Ebu-V2R58
Ebu-V2R122
Ebu-V2R121
Ebu-V2R124
Ebu-V2R123
Ebu-V2R120
Ebu-V2R61
Ebu-V2R79
Ebu-V2R23
Ebu-V2R43
Ebu-V2R95
Ebu-V2R98
Ebu-V2R40
Ebu-V2R39
Ebu-V2R28
Ebu-V2R62
Ebu-V2R63
Ebu-V2R78
Ebu-V2R114
Ebu-V2R9
Ebu-V2R32
Ebu-V2R101
Ebu-V2R100
Ebu-V2R69
Ebu-V2R132
Ebu-V2R34
Ebu-V2R7
Ebu-V2R57
Ebu-V2R67
Ebu-V2R68
Ebu-V2R66
Ebu-V2R65
Ebu-V2R105
Ebu-V2R89
Ebu-V2R5
Ebu-V2R90
Ebu-V2R88
Ebu-V2R94
Ebu-V2R77
Ebu-V2R108
Ebu-V2R8
Ebu-V2R16
Ebu-V2R21
Ebu-V2R13
Ebu-V2R109
Ebu-V2R35
Ebu-V2R110
Ebu-V2R47
Ebu-V2R30
Ebu-V2R3
Ebu-V2R45
Ebu-V2R102
Ebu-V2R31
Ebu-V2R6
Ebu-V2R12
Ebu-V2R42
Ebu-V2R103
Ebu-V2R113
Ebu-V2R111
Ebu-V2R104
Ebu-V2R106
Ebu-V2R107
Ebu-V2R26
Ebu-V2R112
Ebu-V2R82
Ebu-V2R81
Ebu-V2R22
Ebu-V2R27
Ebu-V2R44
Ebu-V2R130
Ebu-V2R38
Ebu-V2R50
Ebu-V2R56
Ebu-V2R54
Ebu-V2R48
Ebu-V2R14
Ebu-V2R59
Ebu-V2R92
Ebu-V2R91
Ebu-V2R127
Ebu-V2R116
Ebu-V2R118
Ebu-V2R117
Ebu-V2R2
Ebu-V2R25
Ebu-V2R17
Ebu-V2R37
Ebu-V2R70
Ebu-V2R72
Ebu-V2R74
Ebu-V2R73
Ebu-V2R71
Ebu-V2R75
Ebu-V2R76
Ebu-V2R20
Ebu-V2R10
Ebu-V2R99
Ebu-V2R96
Ebu-V2R93
Ebu-V2R19
Ebu-V2R4
Ebu-V2R36
Ebu-V2R80
Ebu-V2R29
Ebu-V2R128
Ebu-V2R87
Ebu-V2R24
Ebu-V2R11
Lamprey V2R-like
Hagfish V2R-like
Gnathostome CaSR
Lamprey CaSR
Hagfish CaSR
TAS1R/GPRC6
Ambulacraria CaSR
GRM
Outgroup
Olfactory
Brain
Pituitary
Gill
Intestine
Kidney
Liver
Testis
Ovary
log10(TPM + 0.001)
−2
0
2


A
Nasal duct
Olfactory organ
B
Eye
Olfactory bulb
Tel
Die
Mes
Rho
B
C
Lamellae
HE
1 mm
C
central core
BC
ORC
cilia/ microvilli
50 µm
D
mOR5
E
V1R1
F
V2R33


V2Rs intensely diversified
Hagfish
TAARs lost
V2Rs lost
Lamprey
(true) V2Rs originated?
mORs, V1Rs, V2Rs, TAARs originated
Gnathostomes