A phylogenetics-based nomenclature system for steroid receptors in teleost fishes

Andrew P. Hoadley[1†], Kathleen M. Munley[1†], and Beau A. Alward[1,2*]

[1]Department of Psychology, University of Houston, Houston, TX, USA, and [2]Department of Biology and Biochemistry, University of Houston, Houston, TX, USA

[†]These authors share first authorship

***Corresponding Author:**
Beau A. Alward
Department of Psychology and Department of Biology and Biochemistry
University of Houston
4349 Martin Luther King Boulevard
Houston, TX 77204
Telephone: 713-743-4138
E-mail: balward@uh.edu

ABSTRACT (250/250 words)

Teleost fishes have emerged as tractable models for studying the neuroendocrine regulation of social behavior via molecular genetic techniques, such as CRISPR/Cas9 gene editing. Moreover, teleosts provide an opportunity to investigate the evolution of steroid receptors and their functions, as species within this lineage possess novel steroid receptor paralogs that resulted from a teleost-specific whole genome duplication. Although teleost fishes have grown in popularity as models for behavioral neuroendocrinology, there is not a consistent nomenclature system for steroid receptors and their genes, which may impede a clear understanding of steroid receptor paralogs and their functions. Here, we used a phylogenetic approach to assess the relatedness of protein sequences encoding steroid receptor paralogs in 18 species from 12 different orders of the Infraclass Teleostei. While most similarly named sequences grouped based on the established phylogeny of the teleost fish lineage, our analysis revealed several inconsistencies in the nomenclature of steroid receptor paralogs, particularly for sequences encoding estrogen receptor beta (ERβ) and estrogen-related receptor gamma (ERγ). Based on our results, we propose a nomenclature system for teleosts in which Greek symbols refer to proteins and numbers refer to genes encoding different subtypes of steroid receptors within the five major groups of this nuclear receptor subfamily. Collectively, our results bridge a critical gap by providing a cohesive naming system for steroid receptors in teleost fishes, which will serve to improve communication, promote collaboration, and enhance our understanding of the evolution and function of steroid receptors across vertebrates.

Keywords: evolution, fishes, genome duplication, paralogs, steroid hormones, steroid receptors

HIGHLIGHTS (3-5 bullet points, 85 characters or less each)

- Teleosts have emerged as models for studying neuroendocrine regulation of behavior.
- Teleost fishes possess additional steroid receptor paralogs due to the TS-WGD.
- Teleosts lack consistent nomenclature system for steroid receptors and their genes.
- Phylogenetic analyses of protein sequences revealed discrepancies for ERβ and ERγ.
- We propose a novel, cohesive naming system for steroid receptors in teleost fishes.

1. INTRODUCTION

Over the past few decades, teleost fishes have emerged as important model systems in numerous biological fields, including molecular biology, genetics, evolution, neuroendocrinology, and behavior (reviewed in Glasauer and Neuhass, 2014; Sato and Nishida, 2010; Volff, 2005). Teleost fishes (Class: Actinopterygii, Infraclass: Teleostei) are the largest and most diverse group of ray-finned fishes that comprise >95% of all extant fish species and roughly half of all extant vertebrate species (Nelson, 1994; reviewed in Volff, 2005). Consequently, teleosts show a remarkable level of diversity in morphology, ecology, physiology, and behavior, enabling researchers to address evolutionary questions related to various facets of biology. In recent years, the popularity of teleosts in biological research has surged, particularly for their tractability in molecular genetic approaches. The development of high-throughput and rapid DNA sequencing methods has enabled researchers to sequence the whole genomes of several fish species from distinct orders within the Infraclass Teleostei, including channel catfish (*Ictalurus punctatus*; Liu et al., 2016), zebrafish (*Danio rerio*; Howe et al., 2013), Atlantic salmon (*Salmo salar*; Lien et al., 2016), spotted green pufferfish (*Tetraodon nigroviridis*; Jaillon et al., 2004), three-spined stickleback (*Gasterosteus aculeatus*; Nath et al., 2021; Peichel et al. 2017, 2020), Nile tilapia (*Oreochromis niloticus*; Conte et al., 2017), and medaka (*Oryzias latipes*; Kasahara et al., 2007). The availability of whole genome sequences, in addition to the prevalence of external fertilization as a mode of reproduction among many teleost species, makes techniques such as CRISPR/Cas9 relatively straightforward to implement (reviewed in Barman et al., 2017; Lu et al., 2021; Zhu and Ge, 2018). Collectively, these advantages make teleost fishes excellent models for examining the molecular regulation of physiological mechanisms and behavior via gene editing.

Genomic and phylogenetic evidence strongly supports the occurrence of a single whole-genome duplication (WGD) event approximately 350 mya in the common ancestor of all teleosts (Brunet et al., 2006), which took place after the emergence of the non-teleost ray-finned fishes (e.g., gar and bowfin), but before the divergence of the superorder Osteoglossomorpha (Hoegg et al., 2004; Fig. 1). This event is often referred to as the teleost-specific WGD [TS-WGD, also known as 3R WGD or third round WGD;

134 reviewed in Glasauer and Neuhauss (2014)]. After a WGD event, redundant genes can
135 face one or more of several fates. Frequently, relaxed selective pressure on a duplicated
136 gene leads to the buildup of deleterious mutations, which results in non-functionalization
137 (Ohno, 1970). In some cases, however, one or more gene paralogs may be under positive
138 selection due to the accumulation of mutations conferring novel functions, a process
139 known as neo-functionalization (Force et al., 1999; Ohno, 1970). Alternatively, duplicate
140 genes may undergo sub-functionalization, in which both paralogs are maintained due to
141 the complementary partitioning of ancestral functions between two daughter genes, such
142 that their complimentary actions comprise the total function of their single ancestral gene
143 (Force et al., 1999).

144 Among the genes that have been duplicated following the TS-WGD are steroid
145 receptors, a family of ligand-dependent nuclear receptors that function as signal
146 transducers and transcription factors (Carson-Jurica et al., 1990). Steroid receptors
147 modulate suites of physiological mechanisms, such as homeostatic functions, growth and
148 development, cellular signaling, sexual differentiation and reproduction, as well as
149 behavior (Szego et al. 2003; Whirledge and Cidlowski, 2019). Steroid receptors consist
150 of five major groups: 1) the androgen receptors (ARs), 2) estrogen receptors (ERs), and
151 3) progesterone receptors (PRs), which mediate the actions of androgens (e.g.,
152 testosterone, dihydrotestosterone, and 11-ketotestosterone), estrogens (e.g., estradiol),
153 and progestins (e.g., progesterone) on the development and maintenance of the
154 reproductive, musculoskeletal, cardiovascular, immune, and central nervous systems
155 (reviewed in Brinton et al., 2008; Chen et al., 2022; Davey and Grossmann, 2016; Mani
156 et al. 1997), and 4) the glucocorticoid receptors (GRs) and 5) mineralocorticoid receptors
157 (MRs), which facilitate the actions of glucocorticoids (e.g., cortisol) and mineralocorticoids
158 (e.g., aldosterone) on metabolism, growth and development, cardiovascular function, ion
159 transport, and salt balance (reviewed in Pascual-Le Tallec and Lombés, 2005;
160 Timmermans et al., 2019). Prior work suggests that extant steroid receptors evolved from
161 an ancestral estrogen receptor that had estrogen-like functionality and preceded the
162 origin of bilaterally symmetric animals, followed by an ancestral progesterone receptor
163 (Thornton, 2001; Thornton et al., 2003). Specifically, extant steroid receptors evolved
164 novel functions from these ancestral receptors following two large-scale genome

expansions, one before and one after the advent of jawed vertebrates (the 1R-WGD and 2R-WGD, respectively). Because the ancestral steroid receptor had estrogen-like functionality, these findings support a model of ligand exploitation for extant steroid receptors, in which the terminal ligand in a biosynthetic pathway (in this case, estrogens) is the first for which a receptor evolves, thereby indirectly selecting for the synthesis of intermediates within this pathway and enabling duplicated receptors to evolve affinity for these substances (Baker, 2004; reviewed in Eick and Thornton, 2011). Taken together, these studies suggest that ERs and PRs evolved from more ancient lineages, whereas ARs, GRs, and MRs are relatively recent in origin.

The novel steroid receptor paralogs resulting from the TS-GWD present both opportunities and challenges for researchers. For instance, novel steroid receptor paralogs provide opportunities to dissect their functions to discover fundamental roles of steroid hormones across species (reviewed in Alward et al., 2023). Some challenges include difficulty in manipulating the functions of these genes (e.g., two steroid receptors bind the same steroid ligand in many teleosts, whereas in other vertebrates, only one gene product does). Another practical challenge is there is not a cohesive nomenclature system for the steroid receptors. This challenge at first may seem inconsequential, but a lack of a logical nomenclature system for novel teleost steroid receptor genes may hinder a clear understanding of their functions. With the advent of CRISPR/Cas9 gene editing technologies, numerous non-traditional laboratory species, including teleost fishes, are emerging as genetically tractable for a variety of research questions (reviewed in Alward et al., 2023; Juntti, 2019), including those related to steroid hormone function (Alward et al., 2023, 2020; Crowder et al., 2018; Faught and Vijayan, 2019, 2018; Muto et al., 2013; Yin et al., 2017; Yong et al., 2017; Ziv et al., 2013). Therefore, it is important for logical, consistent nomenclature to be used for steroid receptors and their genes across teleost fishes.

The goal of the current study was to use a quantitative approach to propose a nomenclature system for steroid receptors in teleosts. To achieve this, we used phylogenetics to determine the relatedness of protein sequences encoding teleost steroid receptor paralogs from the five major groups of this nuclear receptor subfamily in vertebrates: ARs, ERs, PRs, GRs, and MRs. This study focused solely on steroid

196 receptors, as the nomenclature issue across teleosts is especially apparent for this class
197 of molecules. We predicted that steroid receptor protein sequences would generally group
198 based on the established phylogeny of the teleost fish lineage, but that there would be a
199 few instances in which protein sequences of certain steroid receptor paralogs show
200 greater similarity to a different group of related paralogs than for that which the sequence
201 is currently named. Together, our findings bridge a critical gap among researchers who
202 study teleost fishes by providing a consistent nomenclature system for steroid receptors,
203 clarifying scientific communication among and between fields and promoting the
204 progression of teleost fish research during rapid technological advancements in genetic
205 technology. We hope that our approach may act as a model to benefit other researchers
206 to establish a consistent nomenclature for other species and/or other molecular systems
207 with an inconsistent nomenclature system.
208

209 **2. MATERIALS AND METHODS**

210 ***2.1. Species selection***

211 To assess the relatedness of steroid receptor protein sequences among teleost
212 fishes, we chose 18 species representing 12 different orders from the Infraclass Teleostei
213 [Order Cypriniformes (carps, minnows, and loaches), Order Cyprinodontimformes
214 (toothcarps), Order Characiformes (tetras and piranhas), Order Siluriformes (catfish),
215 Order Salmoniformes (salmonids), Order Gadiformes (cods), Order Beloniformes
216 (medakas, ricefishes, and needlefishes), Order Cichliformes (cichlids and convict
217 blennies), Order Perciformes (perches, darters, basses, and groupers), Order
218 Gasterosteiformes (sticklebacks, pipefishes, and seahorses), Order Pleuronectiformes
219 (flatfishes), Order Tetraodontiformes (puffers and filefishes), and family Pomacentridae
220 in the subseries Ovalentaria (damselfishes and clownfishes)]. Specifically, we selected
221 the following species for phylogenetic analysis: *Anguilla japonica* (Japanese eel), *Danio*
222 *rerio* (zebrafish), *Carassius auratus* (goldfish), *Fundulus heteroclitus* (Atlantic killifish),
223 *Poecilia mexicana* (shortfin molly), *Gambusia affinis* (western mosquitofish), *Astyanax*
224 *mexicanus* (Mexican tetra), *Ictalurus punctatus* (channel catfish), *Oncorhynchus mykiss*
225 (rainbow trout), *Gadus morhua* (Atlantic cod), *Oryzias latipes* (medaka), *Astatotilapia*
226 (*Haplochromis) burtoni* (Burton's mouthbrooder), *Oreochromis niloticus* (Nile tilapia),

227 *Amphiprion percula* (orange clownfish), *Dicentrarchus labrax* (European seabass),
228 *Gasterosteus aculeatus* (three-spined stickleback), *Paralichthys olivaceus* (olive
229 flounder), and *Tetraodon nigroviridis* (spotted green pufferfish). These species were
230 chosen based on the availability of protein sequences for steroid receptors, many of which
231 were acquired via whole-genome sequencing projects, and to represent the diversity of
232 fishes within the teleost lineage.

233 In addition, we included 1 non-teleost ray-finned fish [*Lepisosteus oculatus*
234 (spotted gar), Infraclass: Holostei], which diverged from teleost fishes before the TS-
235 WGD; and 2 cartilaginous fishes [*Leucoraja erinacea* (little skate), Order: Rajiformes;
236 *Callorhinchus milli* (Australian ghostshark), Order: Chimaeriformes] and 1 jawless fish
237 [*Petromyzon marinus* (sea lamprey), Order: Petromyzontiformes], which are ancestral to
238 ray-finned fishes. These species were selected to examine potential differences between
239 steroid receptor protein sequences from more ancestral fish species (*P. marinus*, *L.*
240 *erinacea*, and *C. milli*) and from a species that diverged from teleosts prior to the TS-
241 WGD (*L. oculatus*) to steroid receptor paralogs within the teleost lineage. Finally, we
242 chose 4 additional species that represent traditional animal models across biological
243 fields [*Xenopus laevis* (African clawed frog), *Gallus gallus* (red junglefowl), *Mus musculus*
244 (house mouse), and *Homo sapiens* (humans)] as outgroups for our phylogenetic
245 analyses. A phylogeny showing established relationships between the teleost fishes, non-
246 teleost fishes, and outgroup species included in our phylogenetic analyses are shown in
247 Figure 1.

248

249 ***2.2. Protein sequence collection and criteria***

250 Protein sequences for androgen, estrogen, progesterone, glucocorticoid, and
251 mineralocorticoid receptors, including paralogs and isoforms resulting from the TS-WGD,
252 were acquired from the National Center for Biotechnology Information (NCBI) Protein and
253 UniProt databases. This search yielded a total of 211 protein sequences across 26
254 species (36 androgen receptor sequences, 95 estrogen receptor sequences, 27
255 progesterone receptor sequences, 30 glucocorticoid receptor sequences, and 23
256 mineralocorticoid receptor sequences). Sequences for the following steroid receptors,
257 paralogs, and isoforms were included in our analysis: ARα, ARβ, ARa, ARb, ERα, ERβ,

258 ERβ1, ERβ2, estrogen-related receptor gamma (ERγ), estrogen-related receptor gamma
259 1 (ERγ1), estrogen-related receptor gamma 2 (ERγ2), PR, PR1, PR2, GR, GR1, GR2,
260 GR2a, GR2b, and MR. Sequences that were described as alternative splice variants or
261 as partial sequences were excluded from analysis. Protein sequences were downloaded
262 from NCBI Protein and UniProt in FASTA format for inclusion in phylogenetic analysis
263 (see ***Section 2.3***). A complete list of protein sequences that were used for phylogenetic
264 analysis, along with their sources, accession numbers, and references, is available in the
265 Supplementary Material (Tables S1-S5).
266

267 ***2.3. Phylogenetic analyses***

268 For each major group of steroid receptors (excluding ERγ, which was run
269 separately from the other ER subtypes because it is highly conserved among teleosts),
270 available or inferred full-length amino acid sequences were used to generate a molecular
271 phylogeny. Construction of multiple alignments of receptor protein sequences with the
272 neighbor-joining method were performed using the MUSCLE Alignment version 3.8.425
273 (Edgar, 2004; Saitou and Nei, 1987). To estimate molecular phylogenies, trees were
274 constructed using the Jukes-Cantor genetic distance model (Jukes and Cantor, 1969).
275 Trees were rooted using the most basal organism for which we had an available sequence
276 [*P. marinus* (sea lamprey), *L. erinacea* (little skate), *C. milii* (Australian ghostshark), or *H.*
277 *sapiens* (human)]. Bootstrap probabilities for the neighbor-joining analyses were
278 calculated using bootstrap resampling with 1000 replicates (Felsenstein, 1985). All
279 analyses were performed using Geneious Prime version 2022.2.2.
280

281 ***2.4 Approach for steroid receptor nomenclature***

282 Our approach to determining a nomenclature system for protein and gene
283 sequences encoding steroid receptors was based on clustering within our phylogenies.
284 We hypothesized that collections of closely related genes across the species would form
285 large clusters that generally correspond to the current understanding of the proteins
286 present in teleosts in which the genes and/or proteins have been identified. For example,
287 we predicted that there would be two distinct clusters for the ARs because two distinct

288 paralogs that have been discovered for these genes in several species of teleosts (Lorin
289 et al., 2015; Ogino et al., 2016).
290 Protein names were assigned for each cluster based on majority rule, in which the
291 name proposed for a given cluster reflected the current protein name that was listed for
292 most teleost species included in our analyses. Once protein names were assigned, we
293 proposed names for their corresponding genes (see Table 1 for a basic consensus of all
294 steroid receptor proteins and genes and Supplementary Material, Tables S1-S5 for a full
295 list of proposed protein and gene names for each individual species). We propose a new
296 system for naming novel paralogous steroid receptor genes using a combination of the
297 existing nomenclature system in teleosts as well as the model of steroid receptor gene
298 nomenclature in mice (*M. musculus*).
299
300 **3. RESULTS**
301 Our phylogenetic approach led to clear classifications of a variety of steroid
302 receptor protein sequences across species. The proposed names for the steroid receptor
303 proteins are shown in each figure for the receptors in distinct, broad clusters delineated
304 by differently colored, transparent rectangles. The corresponding gene name for each
305 protein is provided in Table 1. Based on clusters in the phylogenies, we propose names
306 across defined clusters for each steroid receptor.
307
308 ***3.1. Androgen receptors***
309 Our analysis of AR strongly resolves the split between the teleost specific ARα and
310 ARβ paralogs (Fig. 2). Overall, protein sequences for AR, ARα, and ARβ clustered into
311 two well-supported clades (90% bootstrap value) that generally agreed with established
312 phylogenetic relationships in the teleost lineage (Fig. 1). The first clade primarily
313 contained ARα sequences ("Androgen Receptor Alpha" group), whereas the second
314 clade mostly consisted of ARβ sequences ("Androgen Receptor Beta" group; Fig. 2).
315 These results agree with previous work showing that many species of teleosts possess
316 two distinct AR paralogs (Lorin et al., 2015; Ogino et al., 2016). Like the current naming
317 system for AR genes in mice (Juntti et al., 2010), we propose referring to the genes for

318 an AR protein with an *ar* base. Thus, genes encoding ARα will be called *ar1*, and genes
319 encoding ARβ will be called *ar2* (Table 1).
320 The AR sequences for the tetrapod species clustered with *L. erinacea* (little skate)
321 at the base of the AR phylogenetic tree (100% bootstrap value; Fig. 2). Most ARα
322 sequences clustered into the "Androgen Receptor Alpha" group (100% bootstrap value),
323 while the majority of ARβ sequences grouped into the "Androgen Receptor Beta" group
324 (94% bootstrap value). Importantly, we identified several discrepancies in how ARs are
325 currently named in the NCBI Protein and UniProt databases. Specifically, sequences
326 labeled "androgen receptor" for *G. morhua* (Atlantic cod) and *A. burtoni* (Burton's
327 mouthbrooder) clustered with the ARα sequences in the "Androgen Receptor Alpha"
328 group (≥ 96% bootstrap values), whereas similarly named sequences for *T. nigroviridis*
329 (spotted green pufferfish), *D. rerio* (zebrafish), *C. auratus* (goldfish), *A. mexicanus*
330 (Mexican tetra), *D. labrax* (European seabass), *A. percula* (orange clownfish), *F.*
331 *heteroclitus* (Atlantic killifish), and *G. affinis* (western mosquitofish) clustered with ARβ
332 sequences in the "Androgen Receptor Beta" group (>90% bootstrap values).
333 Interestingly, the ARα and ARβ sequences for *O. mykiss* (rainbow trout) and *A. japonica*
334 (Japanese eel) clustered in the "Androgen Receptor Beta" group, suggesting that the ARα
335 sequences in these species are more similar to ARβ sequences in other teleost species
336 than to other ARα sequences.
337
338 ***3.2. Estrogen receptors***
339 Our phylogenetic analysis for ERs revealed several inconsistencies in the
340 nomenclature of estrogen receptor paralogs in teleost fishes based on protein sequence
341 similarities, particularly for ERβ and ERγ (Fig. 3). Protein sequences for ER, ERα, and
342 ERβ clustered into two well-supported clades (100% bootstrap values). The first clade
343 largely consisted of ERα sequences ("Estrogen Receptor Alpha" group), whereas the
344 second clade primarily contained ERβ sequences ("Estrogen Receptor Beta" group).
345 Furthermore, there were two additional clusters within the "Estrogen Receptor Beta"
346 group that were supported by our analysis (70% bootstrap value): 1) the "Estrogen
347 Receptor Beta 2" group, which was mostly comprised of ERβ2 sequences (93% bootstrap
348 value), and 2) the "Estrogen Receptor Beta 1" group, which mainly included ERβ1

349 sequences (60% bootstrap value; Fig. 3). These results are consistent with the
350 expression of two distinct ERβ genes, *esr2a* and *esr2b*, in medaka (*O. latipes*; Nishiike et
351 al., 2021) and Burton's mouthbrooder (*A. burtoni*; Maruska et al., 2020). We propose
352 calling the genes for ERα, ERβ1, and ERβ2, and ERγ *esr1, esr2a, esr2b,* and *esr3*,
353 respectively (Table 1).

354 The ERα sequence for *P. marinus* (sea lamprey) clustered by itself at the base of
355 the ERα and ERβ phylogenetic tree (100% bootstrap value; Fig. 3). All other ERα
356 sequences clustered into a single clade ("Estrogen Receptor Alpha," 100% bootstrap
357 value), including those of *X. laevis* (African clawed frog; "estrogen receptor alpha 1" and
358 "estrogen receptor alpha 2") and the two cartilaginous fish species [*L. erinacea* (little
359 skate) and *C. milli* (Australian ghostshark)], and generally agreed with established
360 phylogenetic relationships within teleost fishes (Fig. 1). Notably, all sequences labeled
361 "estrogen receptor" in NCBI Protein and UniProt, including those of *D. rerio* (zebrafish),
362 *C. auratus* (goldfish), *A. mexicanus* (Mexican tetra), *O. mykiss* (rainbow trout), *O. latipes*
363 (medaka), and *A. percula* (orange clownfish), as well as "estrogen receptor 1" for *A.*
364 *japonica* (Japanese eel) and the "type I estrogen receptor" for *O. niloticus* (Nile tilapia),
365 clustered with the ERα sequences in the "Estrogen Receptor Alpha" group (≥ 99%
366 bootstrap values; Fig. 3).

367 Conversely, we found some discrepancies in nomenclature based on protein
368 sequence alignment for the "Estrogen Receptor Beta" clade (Fig. 3). The tetrapod
369 outgroup sequences for ERβ ["estrogen receptor beta 1" and "estrogen receptor beta 2"
370 for *X. laevis* (African clawed frog)] clustered with *L. erinacea* (little skate), *C. milli*
371 (Australian ghostshark), and *L. oculatus* (spotted gar), species that diverged from teleosts
372 prior to the TS-WGD, and these sequences were basal to all teleost ERβ sequences that
373 were included in our analysis (100% bootstrap value). The majority of ERβ2 sequences,
374 which are labeled "estrogen receptor beta 2" and "estrogen receptor 2b," clustered into a
375 single well-supported clade ("Estrogen Receptor Beta 2," 93% bootstrap value) that was
376 consistent with known phylogenetic relationships in the teleost lineage (Fig. 1).
377 Importantly, there was one sequence within the "Estrogen Receptor Beta 2" group that
378 was unexpected: the "estrogen receptor beta 1" sequence for *D. rerio* (zebrafish, 100%
379 bootstrap value). The sequences "estrogen receptor type beta" for *I. punctatus* (channel

380 catfish) and “estrogen receptor BetaB” for *F. heteroclitus* (Atlantic killifish) also clustered
381 with the ERβ2 sequences in the “Estrogen Receptor β2” group (≥ 96% bootstrap values).
382 Similarly, most ERβ1 sequences, which are labeled “estrogen receptor beta 1” or
383 “estrogen receptor 2a,” clustered into a single clade (60% bootstrap value) that typically
384 followed known phylogenetic relationships in teleost fishes (Fig. 1). Moreover, all
385 sequences labeled “estrogen receptor beta” in NCBI Protein and UniProt, including those
386 of *C. auratus* (goldfish), *O. mykiss* (rainbow trout), *O. latipes* (medaka), and *P. olivaceus*
387 (olive flounder), as well as “estrogen receptor BetaA” for *F. heteroclitus* (Atlantic killifish)
388 and “type II estrogen receptor” for *O. niloticus* (Nile tilapia), clustered with the ERβ
389 sequences in the “Estrogen Receptor Beta 1” group (≥ 80% bootstrap values; Fig. 3).
390 Of the ER paralogs examined in this study, ERγ showed the highest level of protein
391 sequence similarity across teleost species (Fig. 4). Because the ERγ protein sequence is
392 highly conserved among teleosts relative to the protein sequences of other ER paralogs,
393 we performed our ERγ phylogenetic analysis separately from that of ERα and ERβ.
394 Interestingly, “estrogen-related receptor gamma” for *P. marinus* (sea lamprey) and
395 “estrogen-related receptor beta” for *A. percula* (orange clownfish) clustered in a single
396 group that was located at the base of the ERγ phylogenetic tree (100% bootstrap values).
397 Overall, ERγ sequences clustered into two clades, one that primarily consisted of ERγb
398 sequences (“Estrogen-Related Receptor Gamma B” group; 92% bootstrap value) and
399 one that was mostly comprised of ERγa sequences (“Estrogen-Related Receptor Gamma
400 A” group; 98% bootstrap value). Notably, there were more discrepancies between the
401 relatedness of ERγ protein sequences and established phylogenetic relationships in
402 teleost fishes compared to the phylogenetic tree generated from ERα and ERβ
403 sequences (Figs. 1 and 3). The “Estrogen-Related Receptor Gamma B” group included
404 sequences labeled “estrogen-related receptor gamma b” with a few exceptions:
405 “estrogen-related receptor gamma type 2” for *O. latipes* (medaka), “estrogen-related
406 receptor alpha” for *T. nigroviridis* (spotted green pufferfish) and *A. percula* (orange
407 clownfish), and “estrogen-related receptor gamma-like” for *C. auratus* (goldfish; 100%
408 bootstrap values). Conversely, the “Estrogen-Related Receptor Gamma A” group
409 consisted of sequences labeled “estrogen-related receptor gamma a” as well as most
410 sequences labeled “estrogen-related receptor gamma” in NCBI Protein or UniProt,

411 including those of *C. milli* (Australian ghostshark), *L. oculatus* (spotted gar), *D. rerio*
412 (zebrafish), *I. punctatus* (channel catfish), *O. mykiss* (rainbow trout), *F. heteroclitus*
413 (Atlantic killifish), *G. morhua* (Atlantic cod), *P. mexicana* (shortfin molly), *O. niloticus* (Nile
414 tilapia), and *A. burtoni* (Burton's mouthbrooder; ≥ 98% bootstrap values). In addition, the
415 "estrogen-receptor gamma type 1" sequence from *O. latipes* (medaka) clustered with the
416 ERγa sequences in the "Estrogen-Related Receptor Gamma A" group (76% bootstrap
417 value; Fig. 4). To our knowledge, it is unknown whether ERγa and Eryb are encoded by
418 distinct genes in the species in which distinct ERγ proteins have been discovered.
419

420 ***3.3. Progesterone receptors***

421 Our analysis of PR suggests that there was no widespread retention of orthologous
422 PR genes following the TS-WGD, consistent with previous studies (Ren et al. 2019; Fig.
423 5). The tetrapod sequences for PR clustered as a sister group to all teleost sequences
424 (100% bootstrap value), and all teleost PR sequences clustered into a single well-
425 supported clade (100% bootstrap value) that mostly agreed with established phylogenetic
426 relationships in this lineage (Fig. 1). Importantly, our search yielded two PR sequences in
427 *C. auratus* (goldfish; PR1 and PR2), which clustered next to each other (86% bootstrap
428 value; Fig. 5). No additional PR sequences have been discovered in closely-related
429 teleost species (e.g., zebrafish; Tang et al., 2016), suggesting that these sequences might
430 represent a gene duplication that is specific to the goldfish lineage. The current name for
431 the PR gene is *pgr*, and this nomenclature was consistent across all teleost species
432 examined in our study.
433

434 ***3.4. Glucocorticoid receptors***

435 Our analysis revealed several inconsistencies in the nomenclature of GR (Fig. 6).
436 Protein sequences for GR, GR1, and GR2 clustered into two clades (78% bootstrap
437 value) that were consistent with known phylogenetic relationships in teleosts (Fig. 1). The
438 first clade mostly consisted of GR2 sequences ("Glucocorticoid Receptor 2" group), while
439 the second clade primarily contained GR1 sequences ("Glucocorticoid Receptor 1" group;
440 Fig. 6), suggesting that two glucocorticoid receptors persisted in the teleost lineage
441 following the TS-WGD. To remain consistent with the naming systems for the other

442 steroid receptor proteins, we propose calling “glucocorticoid receptor 1” GRα and
443 “glucocorticoid receptor 2” GRβ with gene names of *nr3c1a* and *nr3c1b* (Table 1).
444 The tetrapod outgroup sequences for GR clustered with that of *L. oculatus* (spotted
445 gar) at the base of the GR phylogenetic tree (100% bootstrap value; Fig. 6). Interestingly,
446 most GR sequences used in our analysis were labeled “glucocorticoid receptor,” yet these
447 sequences still grouped into two separate clades. Specifically, sequences labeled
448 “glucocorticoid receptor” for *D. rerio* (zebrafish), *T. nigroviridis* (spotted green pufferfish),
449 *F. heteroclitus* (Atlantic killifish), *O. latipes* (medaka), *G. aculeatus* (three-spined
450 stickleback), and *A. burtoni* (Burton’s mouthbrooder) clustered in the “Glucocrticoid
451 Receptor 2” group (100% bootstrap value), whereas sequences labeled “glucocorticoid
452 receptor” for *A. japonica* (Japanese eel), *A. mexicanus* (Mexican tetra), *O. mykiss*
453 (rainbow trout), *A. percula* (orange clownfish), *O. niloticus* (Nile tilapia), *D. labrax*
454 (European seabass), and *P. olivaceus* (olive flounder), in addition to “glucocorticoid
455 receptor-like” for *G. morhua* (Atlantic cod) and *P. mexicana* (shortfin molly), clustered in
456 the “Glucocorticoid Receptor 1” group (93% bootstrap value). A few species in our
457 analysis had two different GR sequences, including *I. punctatus* (channel catfish;
458 “glucocorticoid receptor 1” and “glucocorticoid receptor 2”), *A. burtoni* [Burton’s
459 mouthbrooder; “glucocorticoid receptor” and “glucocorticoid receptor (2)”] and *O. mykiss*
460 (rainbow trout; “glucocorticoid receptor” and “glucocorticoid receptor 2”). Although these
461 sequences generally clustered with other similarly named sequences, the “glucocorticoid
462 receptor (2)” sequence for *A. burtoni* clustered with sequences in the “Glucocorticoid
463 Receptor 1” group, suggesting that this sequence should be called GRα. Because
464 duplicate GR sequences were discovered in highly derived clades (e.g., *A. burtoni*), our
465 results suggest that GR paralogs may exist, but remain undescribed in some teleost
466 species.
467
468 ***3.5. Mineralocorticoid receptors***
469 Our analysis of MR indicates no widespread retention of orthologous MR genes
470 following the TS-WGD (Fig. 7). The tetrapod outgroup sequences for MR clustered with
471 *P. marinus* (sea lamprey) and were basal to all teleost MR sequences (100% bootstrap
472 value). MR protein sequences from teleost fishes clustered into a single well-supported

473 clade (100% bootstrap value) that was consistent with known phylogenetic relationships
474 in this lineage (Fig. 1). While one of the sequences included in our analysis was labeled
475 "mineralocorticoid receptor-like" [from *P. mexicana* (shortfin molly)], this sequence
476 confidently clustered with other MRs in our analysis (100% bootstrap value), suggesting
477 that this protein sequence encodes MR. The current name for the MR gene is *nr3c2*, and
478 this nomenclature was consistent in all teleost species included in our study.
479

480 **4. DISCUSSION**

481 As the most specious clade in the animal kingdom, teleost fishes exhibit profound
482 variation in life-history traits, especially those related to the actions of steroid hormones,
483 such as reproductive and physiological plasticity (reviewed in Godwin, 2010). As such,
484 steroid production and signaling mechanisms have been studied extensively in many
485 teleost fishes, presenting both opportunities and challenges in making comparative
486 insights in the field of behavioral neuroendocrinology. Because the genome editing
487 revolution has enhanced research progress and genetic tractability in diverse non-
488 traditional species like teleosts (reviewed in Alward et al., 2023; Juntti, 2019), it is
489 important that basic tools, such as the names of widely-studied proteins and genes, are
490 consistently and logically applied. To facilitate both ongoing and future collaboration,
491 researchers in this diverse field must endeavor to work from the same set of facts. As an
492 initial step in that direction, we used a phylogenetic approach to propose a consistent
493 naming scheme for steroid receptors and their genes.

494 In our analysis, we identified several instances of supporting evidence for novel
495 steroid receptors, including PRs in goldfish, which possess two distinct PRs encoded by
496 distinct genes, and ERγ, for which a few species had two copies. Future work
497 investigating whether other species possess extra PR genes and whether the different
498 ERγ proteins are encoded by distinct genes is warranted. Several intriguing observations
499 were also made for the ARs. For example, in the Japanese eel (*A. japonica*), both
500 identified AR sequences clustered with ARβ sequences in the "Androgen Receptor Beta"
501 group, suggesting this species has two ARβ receptors. It is unclear whether these two
502 receptors are encoded by distinct genes, but unlike in other teleosts, no significant
503 differences in the androgen-binding capacity of the two eel receptors has been detected

504 (Peñaranda et al., 2014). These results suggest that the eel ARα and ARβ known in the
505 literature (Ikeuchi et al., 1999; Todo et al., 1999) are either splice variants or paralogs of
506 the teleost ARβ. It is intriguing to speculate whether these sequences represent evidence
507 of early loss of function (e.g., non-functionalization) of the ARα receptor in this lineage.
508 The AR sequences in rainbow trout *(O. mykiss)* also clustered with ARβ sequences in the
509 "Androgen Receptor Beta" group, which is consistent with prior work suggesting that
510 these paralogs resulted from a lineage specific WGD occurred in the Salmoniformes
511 (Allendorf et al., 1984; Takeo and Yamashita, 1999). The framework we propose can
512 provide researchers with novel hypotheses to test, enabling them to disentangle AR
513 evolution and selection in teleost fishes and vertebrates more broadly.
514 Most of the name changes we propose are straightforward, aligning
515 inconsistencies between species with what is otherwise a consensus within the field.
516 Proposing any universal naming system for certain proteins and genes, however, remains
517 a challenge. For example, some teleosts, such as *D. rerio,* only have one described AR
518 protein and one AR gene. Traditionally, animals with a single AR use the name AR (*ar*).
519 Our analysis shows, however, that the AR protein sequence in *D. rerio* clusters with ARβ
520 sequences. Thus, this sequence should be referred to as ARβ (*ar2*) based on our
521 proposed nomenclature system in teleosts, which takes into account that the vast majority
522 of teleost species possess two ARs (ARα and ARβ). We propose that for such cases,
523 including widely studied models like zebrafish and goldfish, a phylogenetic approach
524 should be used to unambiguously identify that the gene in question belongs to the ARβ
525 (*ar2*) lineage of receptors. When, as is the case of this example, an extensive body of
526 literature exists that retains the ambiguous naming convention, we suggest that explicit
527 reference is made to the protein or gene of interest's evolutionary context. By rendering
528 synonymous the original name AR or *ar* and the more nuanced ARβ or *ar2*, we believe
529 future research will facilitative collaboration as well as yield comparative insights within
530 the diverse teleost clade.
531
532 **5. CONCLUSIONS**
533 Within the most popular genetic model systems, such as house mice, there is
534 broad consensus on names of some of the most widely-studied genes that are relevant

535 to numerous subfields of biology. This is clearly the case for steroid receptors, in which
536 the names of proteins and genes encoding steroid receptors are consistent across the
537 literature. For teleost fishes, where studies of steroid signaling functions are extensive
538 and broadly studied, the names used for these proteins and genes have been
539 inconsistently assigned. Our study addresses this challenge by providing a cohesive
540 nomenclature system for steroid receptors in teleosts, an essential step for enhancing
541 communication and promoting collaborative efforts. More broadly, we hope our analysis
542 and recommendations will propagate a shift towards consistency in the names used for
543 these widely studied proteins and genes, enabling us to better understand the evolution
544 and function of steroid receptors across taxa.
545

546 **DECLARATION OF COMPETING INTEREST**

547 The authors declare no competing or financial interests.
548

549 **DATA AVAILABILITY**

550 A complete list of sources and accession numbers for the protein sequences used
551 in our phylogenetic analyses are available in the Supplementary Material (Tables S1-S5).
552

553 **ACKNOWLEDEGMENTS**

554 We thank members of the Alward Lab for their intellectual contributions to the ideas
555 discussed in this paper. This research was funded by NIGMS grant R35GM142799 (to
556 B.A.A.), University of Houston – National Research University Fund grant R0503962 (to
557 B.A.A.), and a Beckman Young Investigator Award (to B.A.A.).
558

559 **AUTHOR CONTRIBUTIONS**

560 A.P.H. conceived of the original idea for the manuscript. B.A.A., A.P.H., and
561 K.M.M. conceived of the approach to analyses and final framing of the ideas presented
562 in the manuscript. A.P.H. and K.M.M. chose the species to include in the analyses. A.P.H.
563 conducted the analyses, generated the phylogenies, and made the figures. K.M.M. did
564 the protein sequence collection, made the table in the main text, and made the
565 supplementary material. B.A.A., A.P.H., and K.M.M. edited final versions of the

supplementary tables. A.P.H. and K.M.M. wrote the methods and the results sections; B.A.A., A.P.H., and K.M.M. wrote the first draft of the introduction; and B.A.A. wrote the first draft of the discussion and conclusion. K.M.M. wrote the abstract, highlights, and references. K.M.M. edited significant portions of all sections of subsequent drafts. All authors reviewed, edited, and approved the final versions of the manuscript and supplementary files.

REFERENCES

Alward, B.A., Hoadley, A.P., Jackson, L.R., Lopez, M.S., 2023. Genetic dissection of steroid-hormone modulated social behavior: Novel paralogous genes are a boon for discovery. Horm Behav 147, 105295. https://doi.org/10.1016/j.yhbeh.2022.105295

Alward, B.A., Laud, B., Skalnik, C.J., York, R.A., Juntii, S., Fernald, R.D., 2020. Modular Genetic Control of Social Status in a cichlid fish. PNAS 1–30.

Baker, M.E., 2004. Evolution of Steroid Receptors, in: Martini, L. (Ed.) Encyclopedia of Endocrine Diseases. Academic Press, pp. 312-319.

Barman, H.K., Rasal, K.D., Chakrapani, V., Ninawe, A.S., Vengayil, D.T., Asrafuzzaman, S., Sundaray, J.K., Jayasankar, P., 2017. Gene editing tools: state-of-the-art and the road ahead for the model and non-model fishes. Transgenic Res. 26, 577-589.

Betancur-R, R., Wiley, E.O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., Ortí, G., 2017. Phylogenetic classification of bony fishes. BMC Evol Biol 17, 162. https://doi.org/10.1186/s12862-017-0958-3.

Brinton, R.D., Thompson, R.F., Foy, M.R., Baudry, M., Wang, J., Finch, C.E., Morgan, T.E., Pike, C.J., Mack, W.J., Stanczyk, F.Z., Nilsen, J., 2008. Progesterone receptors: Form and function in brain. Front. Neuroendocrinol. 29(2), 313-339.

Brunet, F.G., Roest Crollius, H., Paris, M., Aury, J.-M., Gibert, P., Jaillon, O., Laudet, V., Robinson-Rechavi, M., 2006. Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. Mol Biol Evol 23, 1808–1816. https://doi.org/10.1093/molbev/msl049

Carson-Jurica, M.A., Schrader, W.T., O'Malley, B.W., 1990. Steroid receptor family: Structure and functions. Endocr Rev 11, 201–220. https://doi.org/10.1210/edrv-11-2-201

Chen, P., Li, B., Ou-Yang, L., 2022. Role of estrogen receptors in health and disease. Front. Endocrinol. 13, 839005.

Conte, M.A., Gammerdinger, W.J., Bartie, K.L., Penman, D.J., Kocher, T.D., 2017. A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. BMC Genomics 18(1), 341.

Crowder, C.M., Lassiter, C.S., Gorelick, D.A., 2018. Nuclear Androgen Receptor Regulates Testes Organization and Oocyte Maturation in Zebrafish. Endocrinology 159, 980–993. https://doi.org/10.1210/en.2017-00617

Davey, R.A., Grossmann, M., 2016. Androgen receptor structure, function and biology: From bench to bedside. Clin. Biochem. Rev. 37(1), 3-15.

608 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high
609 throughput. Nucleic Acids Res 32, 1792–1797.
610 https://doi.org/10.1093/nar/gkh340

611 Eick, G.N., Thornton, J.W., 2011. Evolution of steroid receptors from an estrogen-
612 sensitive ancestral receptor. Mol. Cell. Endocrinol. 334(1-2), 31-38.

613 Faught, E., Vijayan, M.M., 2019. Loss of the glucocorticoid receptor in zebrafish improves
614 muscle glucose availability and increases growth. Am J Physiol Endocrinol Metab
615 316, E1093–E1104. https://doi.org/10.1152/ajpendo.00045.2019

616 Faught, E., Vijayan, M.M., 2018. The mineralocorticoid receptor is essential for stress
617 axis regulation in zebrafish larvae. Sci Rep 8, 1–11.
618 https://doi.org/10.1038/s41598-018-36681-w

619 Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y., Postlethwait, J., 1999.
620 Preservation of Duplicate Genes by Complementary, Degenerative Mutations.
621 Genetics 151, 1531–1545. https://doi.org/10.1093/genetics/151.4.1531

622 Glasauer, S.M.K., Neuhauss, S.C.F., 2014. Whole-genome duplication in teleost fishes
623 and its evolutionary consequences. Mol. Genet. Genomics 289(6), 1045-1060.

624 Godwin, J. (2010). Neuroendocrinology of sexual plasticity in teleost fishes. Frontiers in
625 Neuroendo. 31, 203-216.

626 Hoegg, S., Brinkmann, H., Taylor, J.S., Meyer, A., 2004. Phylogenetic timing of the fish-
627 specific genome duplication correlates with the diversification of teleost fish. J
628 Mol Evol 59, 190–203. https://doi.org/10.1007/s00239-004-2613-z

629 Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E.,
630 Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M.,
631 Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.-J., White, S., Chow, W.,
632 Kilian, B., Quintais, L.T., Guerra-Assunção, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel,
633 J.-H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire,
634 S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King,
635 J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton,
636 J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle,
637 C., Elliott, D., Threadgold, G., Harden, G., Ware, D., Begum, S., Mortimore, B.,
638 Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C.,
639 Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P.,
640 Barker, N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J.,
641 Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen,
642 C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire,
643 S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead,
644 S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N.,
645 Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-
646 Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N.,
647 Ellwood, M., Woodmansey, R., Clark, G., Cooper, J.D., Tromans, A., Grafham, D.,

Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Ürün, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberländer, M., Rudolph-Geiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Schuster, S.C., Carter, N.P., Harrow, J., Ning, Z., Herrero, J., Searle, S.M.J., Enright, A., Geisler, R., Plasterk, R.H.A., Lee, C., Westerfield, M., de Jong, P.J., Zon, L.I., Postlethwait, J.H., Nüsslein-Volhard, C., Hubbard, T.J.P., Crollius, H.R., Rogers, J., Stemple, D.L., 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature 496(7446), 498-503.

Hughes, L.C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C.C., Thompson, A.W., Arcila, D., Betancur-R., R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhou, Z., Huang, H., Chen, S., Venkatesh, B., Shi, Q., 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proc. Natl. Acad. Sci. U.S.A. 115, 6249–6254. https://doi.org/10.1073/pnas.1719358115

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C., Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N., Castellano, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Biémont, C., Skalli, Z., Cattolico, L., Poulain, J., De Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.-P., Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J., McEwan, P., Bosak, S., Kellis, M., Volff, J.-N., Guigó, R., Zody, M.C., Mesirov, J., Lindblad-Toh, K., Birren, B., Nusbaum, C., Kahn, D., Robinson-Rechavi, M., Laudet, V., Schachter, V., Quétier, F., Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J., Crollius, H.R., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431(7011), 946-957.

Jukes, T.H., Cantor, C.R., 1969. Evolution of Protein Molecules, in: Mammalian Protein Metabolism. Elsevier, pp. 21–132. https://doi.org/10.1016/B978-1-4832-3211-9.50009-7

Juntti, S., 2019. The future of gene-guided neuroscience research in non-traditional model organisms. Brain Behav Evol. https://doi.org/10.1159/000500072

Juntti, S.A., Tollkuhn, J., Wu, M.V., Fraser, E.J., Soderborg, T., Tan, S., Honda, S.-I., Harada, N., Shah, N.M., 2010. The androgen receptor governs the execution, but not programming, of male sexual and territorial behaviors. Neuron 66, 260–72. https://doi.org/10.1016/j.neuron.2010.03.024

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., N., S., Hashimoto, S., Yang, J., Lee, Y., Matsushima, K., Sugano, S., Sakaizumi, M.,

Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K., Morishita, T., Endo, T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y., 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447(7145), 714-719.

Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Olav Vik, J., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Inge Våge, D., de Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturra, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S.W., Davidson, W.S., 2016. The Atlantic salmon genome provides insights into rediploidization. Nature 533(7602), 200-205.

Liu, Z., Liu, S., Yao, J., Bao, L., Zhang, J., Li, Y., Jiang, C., Sun, L., Wang, R., Zhang, Y., Zhou, T., Zeng, Q., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y., Zimin, A., Xu, P., Phillippy, A.M., Geng, X., Song, L., Sun, F., Li, C., Wang, X., Chen, A., Jin, Y., Yuan, Z., Yang, Y., Tan, S., Peatman, E., Lu, J., Qin, Z., Dunham, R., Li, Z., Sonstegard, T., Feng, J., Danzmann, R.G., Schroeder, S., Scheffler, B., Duke, M.V., Ballard, L., Kucuktas, H., Kaltenboeck, L., Liu, H., Armbruster, J., Xie, Y., Kirby, M.L., Tian, Y., Flanagan, M.E., Mu, W., Waldbieser, G.C., 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. Nat. Commun. 7, 11757.

Lorin, T., Salzburger, W., Böhne, A., 2015. Evolutionary fate of the androgen receptor-Signaling pathway in ray-finned fishes with a special focus on cichlids. G3: Genes, Genomes, Genetics 5, 2275–2283. https://doi.org/10.1534/g3.115.020685

Lu, J., Fang, W., Huang, J., Li, S., 2021. The application of genome editing technology in fish. Mar. Life Sci. Technol. 3, 326-346.

Mani, S.K., Blaustein, J.D., O'Malley, B.W., 1997. Progesterone receptor function from a behavioral perspective. Horm. Behav. 31(3), 244-255.

Maruska, K.P., Butler, J.M., Anselmo, C., Tandukar, G., 2020. Distribution of aromatase in the brain of the African cichlid fish Astatotilapia burtoni: Aromatase expression, but not estrogen receptors, varies with female reproductive-state. Journal of Comparative Neurology 528, 2499–2522. https://doi.org/10.1002/cne.24908

Muto, A., Taylor, M.R., Suzawa, M., Korenbrot, J.I., Baier, H., 2013. Glucocorticoid receptor activity regulates light adaptation in the zebrafish retina. Front Neural Circuits 7, 1–8. https://doi.org/10.3389/fncir.2013.00145

Nath, S., Shaw, D.E., White, M.A., 2021. Improved contiguity of the threespine stickleback genome using long-read sequencing. G3 (Bethesda) 11(2), jkab007.

Nelson, J.S., 1994. Fishes of the World, 3 ed. John Wiley and Sons, New York.

Nishiike, Y., Miyazoe, D., Togawa, R., Yokoyama, K., Nakasone, K., Miyata, M., Kikuchi, Y., Kamei, Y., Todo, T., Ishikawa-Fujiwara, T., Ohno, K., Usami, T., Nagahama, Y., Okubo, K., 2021. Estrogen receptor 2b is the major determinant of sex-typical mating behavior and sexual preference in medaka. Current Biology 31, 1699-1710.e6. https://doi.org/10.1016/j.cub.2021.01.089

Ogino, Y., Kuraku, S., Ishibashi, H., Miyakawa, H., Sumiya, E., Miyagawa, S., Matsubara, H., Yamada, G., Baker, M.E., Iguchi, T., 2016. Neofunctionalization of androgen receptor by Gain-of-Function mutations in teleost fish lineage. Mol Biol Evol 33, 228–244. https://doi.org/10.1093/molbev/msv218

Ohno, S., 1970. The Enormous Diversity in Genome Sizes of Fish as a Reflection of Nature's Extensive Experiments with Gene Duplication. Transactions of the American Fisheries Society 99, 120–130. https://doi.org/10.1577/1548-8659(1970)99<120:TEDIGS>2.0.CO;2

Pascual-Le Tallec, L., Lombès, M., 2005. The mineralocorticoid receptor: A journey exploring its diversity and specificity of action. Mol. Endocrinol. 19(9), 2211-2221.

Peichel, C.L., McCann, S.R., Ross, J.A., Naftaly, A.F.S., Urton, J.R., Cech, J.N., Grimwood, J., Schmutz, J., Myers, R.M., Kingsley, D.M., White, M.A., 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. Genome Biol. 21(1), 177.

Peichel, C.L., Sullivan, S.T., Liachko, I., White, M.A., 2017. Improvement of the threespine stickleback genome using a Hi-C-based proximity-guided assembly. J. Hered. 108(6), 693-700.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4, 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Sato, Y., Nishida, M., 2010. Teleost fish with specific genome duplication as unique models of vertebrate evolution. Environ. Biol. Fishes 88, 169-188.

Szego, C.M., Nemere, I., Pietras, R.J., 2003. Membrane Steroid Receptors, in: Henry, H.L., Norman, A.W. (Eds.), Encyclopedia of Hormones. Academic Press, pp. 657-671.

Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc. Natl. Acad. Sci. USA 98(10), 5671-5676.

Thornton, J.W., Need, E., Crews, D., 2003. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 301(5640), 1714-1717.

Timmermans, S., Souffriau, J., Libert, C., 2019. A general introduction to glucocorticoid biology. Front. Immunol. 10, 1545.

765 Volff, J.-N., 2005. Genome evolution and biodiversity in teleost fish. Heredity 94(3), 280-
766 294.

767 Whirledge, S., Cidlowski, J.A., 2019. Chapter 5 - Steroid Hormone Action, in: Strauss III,
768 J.F., Barbieri, R.L. (Eds.), Yen & Jaffe's Reproductive Endocrinology: Physiology,
769 Pathophysiology, and Clinical Management. Elsevier, pp. 115-131.e114.

770 Yin, Y., Tang, H., Liu, Y., Chen, Y., Li, G., Liu, X., Lin, H., 2017. Targeted disruption of
771 aromatase reveals dual functions of cyp19a1a during sex differentiation in
772 zebrafish. Endocrinology 158, 3030–3041. https://doi.org/10.1210/en.2016-1865

773 Yong, L., Thet, Z., Zhu, Y., 2017. Genetic editing of the androgen receptor contributes to
774 impaired male courtship behavior in zebrafish. Journal of Experimental Biology
775 220, 3017–3021. https://doi.org/10.1242/jeb.161596

776 Ziv, L., Muto, A., Schoonheim, P.J., Meijsing, S.H., Strasser, D., Ingraham, H.A., Schaaf,
777 M.J.M., Yamamoto, K.R., Baier, H., 2013. An affective disorder in zebrafish with
778 mutation of the glucocorticoid receptor. Mol Psychiatry 18, 681–691.
779 https://doi.org/10.1038/mp.2012.64

780 Zhu, B., Ge, W., 2018. Genome editing in fishes and their applications. Gen. Comp.
781 Endocrinol. 257, 3-12.

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798 **TABLES**

799

800 **Table 1.** Proposed nomenclature for proteins and genes encoding steroid receptors in
801 teleost fishes.

802

Protein	Gene	Figure
AR		
ARα	*ar1*	Fig. 2
ARβ	*ar2*	Fig. 2
ER		
ERα	*esr1*	Fig. 3
ERβ1	*esr2a*	Fig. 3
ERβ2	*esr2b*	Fig. 3
ERγ	*esr3*	Fig. 4
ERγ1	*esr3*	Fig. 4
ERγ2	*esr3*	Fig. 4
PR	*pgr*	
PRα	*pgr1*	Fig. 5
PRβ	*pgr2*	Fig. 5
GR	*nr3c1*	
GRα	*nr3c1a*	Fig. 6
GRβ	*nr3c1b*	Fig. 6
MR	*nr3c2*	Fig. 7

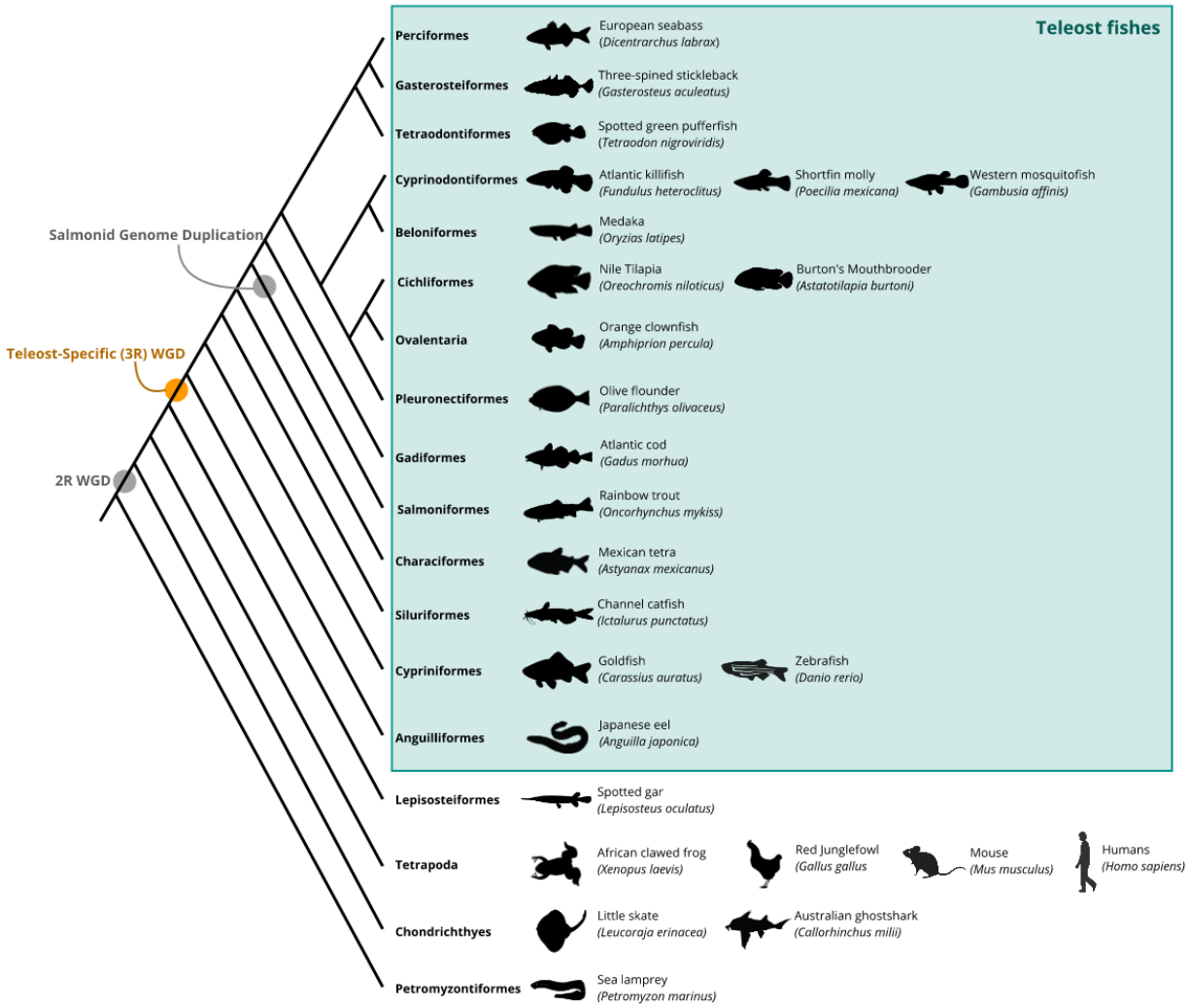
803
804 Proposed protein and gene names for androgen receptors, estrogen receptors, progesterone receptors,
805 glucocorticoid receptors, and mineralocorticoid receptors in teleost fishes based on phylogenetic analysis.
806 The location of figures containing phylogenies associated with each steroid receptor subtype are also
807 shown. *Abbreviations: AR, androgen receptor; ar1, androgen receptor 1 gene; ARα, androgen receptor*
808 *alpha protein; ar2, androgen receptor 2 gene; ARβ, androgen receptor beta protein; ER, estrogen*
809 *receptor; ERα, estrogen receptor alpha protein; ERβ, estrogen receptor beta protein; ERβ1, estrogen*
810 *receptor beta 1 protein; ERβ2, estrogen receptor beta 2 protein; ERγ, estrogen-related receptor gamma*
811 *protein; ERγ1, estrogen-related receptor gamma 1 protein; ERγ2, estrogen-related receptor gamma 2*
812 *protein; esr1, estrogen receptor 1 gene; esr2a, estrogen receptor 2a gene; esr2b, estrogen receptor 2b*
813 *gene; esr3, estrogen receptor 3 gene; GR, glucocorticoid receptor protein; GRα, glucocorticoid receptor*
814 *alpha protein; GRβ, glucocorticoid receptor beta protein; MR, mineralocorticoid receptor; nr3c1, nuclear*
815 *receptor subfamily 3 group C member 1 gene; nr3c1a, nuclear receptor subfamily 3 group C member 1a*
816 *gene; nr3c1b, nuclear receptor subfamily 3 group C member 1b gene; nr3c2, nuclear receptor subfamily*

817 *3 group C member 2 gene; pgr, progesterone receptor gene; pgr1, progesterone receptor 1 gene; pgr2,*
818 *progesterone receptor 2 gene; PR, progesterone receptor; PRα, progesterone receptor alpha protein;*
819 *PRβ, progesterone receptor beta protein.*
820
821 **FIGURE LEGENDS**
822
823 **Figure 1**: Phylogeny identifying the evolutionary relationships between taxa considered
824 in this paper [after Betancur-R et al. (2017) and Hughes et al. (2018)]. Gray and orange
825 circles indicate genome duplication events. *Abbreviations: 2R, second-round; 3R, third-*
826 *round; WGD, whole-genome duplication.*
827
828 **Figure 2:** Consensus phylogenetic tree of androgen receptors. Numbers shown at each
829 branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic
830 distance (0.1 substitutions/site). The orange circle represents the approximate position of
831 the teleost-specific whole-genome duplication event (TS-WGD).
832
833 **Figure 3:** Consensus phylogenetic tree of estrogen receptors alpha (α) and beta (β).
834 Numbers shown at each branch node indicate bootstrap values (%), and the scale bar
835 indicates phylogenetic distance (0.1 substitutions/site). Orange circles represent the
836 approximate position of the teleost-specific whole-genome duplication event (TS-WGD),
837 which impacted the independent duplication of ERβ, but not ERα.
838
839 **Figure 4**: Consensus phylogenetic tree of estrogen-related receptor gamma (ERγ).
840 Numbers shown at each branch node indicate bootstrap value (%), and the scale bar
841 indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the
842 approximate position of the teleost-specific whole-genome duplication event (TS-WGD).
843 **Figure 5:** Consensus phylogenetic tree of progesterone receptor. Numbers shown at
844 each branch node indicate bootstrap value (%), and the scale bar indicates phylogenetic
845 distance (0.1 substitutions/site). The orange circle represents the approximate position of
846 the teleost-specific whole-genome duplication event (TS-WGD).
847

848 **Figure 6:** Consensus phylogenetic tree of glucocorticoid receptors. Numbers shown at
849 each branch node indicate bootstrap value (%), and the scale bar indicates phylogenetic
850 distance (0.1 substitutions/site). The orange circle represents the approximate position of
851 the teleost-specific whole-genome duplication event (TS-WGD).
852 **Figure 7:** Consensus phylogenetic tree of mineralocorticoid receptor. Numbers shown at
853 each branch node indicate bootstrap value (%), and the scale bar indicates phylogenetic
854 distance (0.1 substitutions/site). The orange circle represents the approximate position of
855 the teleost-specific whole-genome duplication event (TS-WGD).
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878

879 **FIGURES**

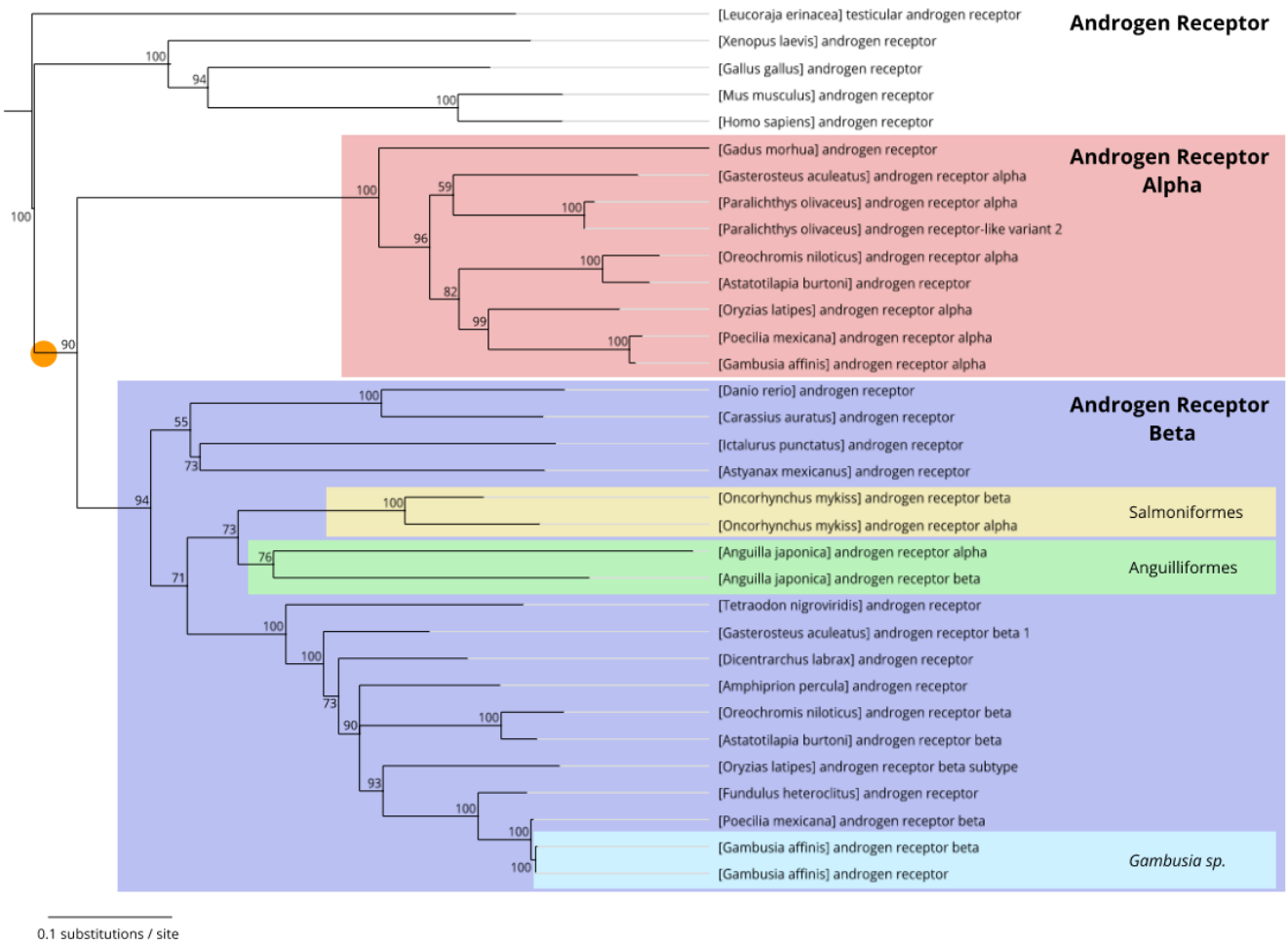
880 **Figure 1.**



881
882
883
884
885
886
887
888
889
890
891

892 **Figure 2.**

893



894

895

896

897

898

899

900

901

902

903

904

905

Figure 3.

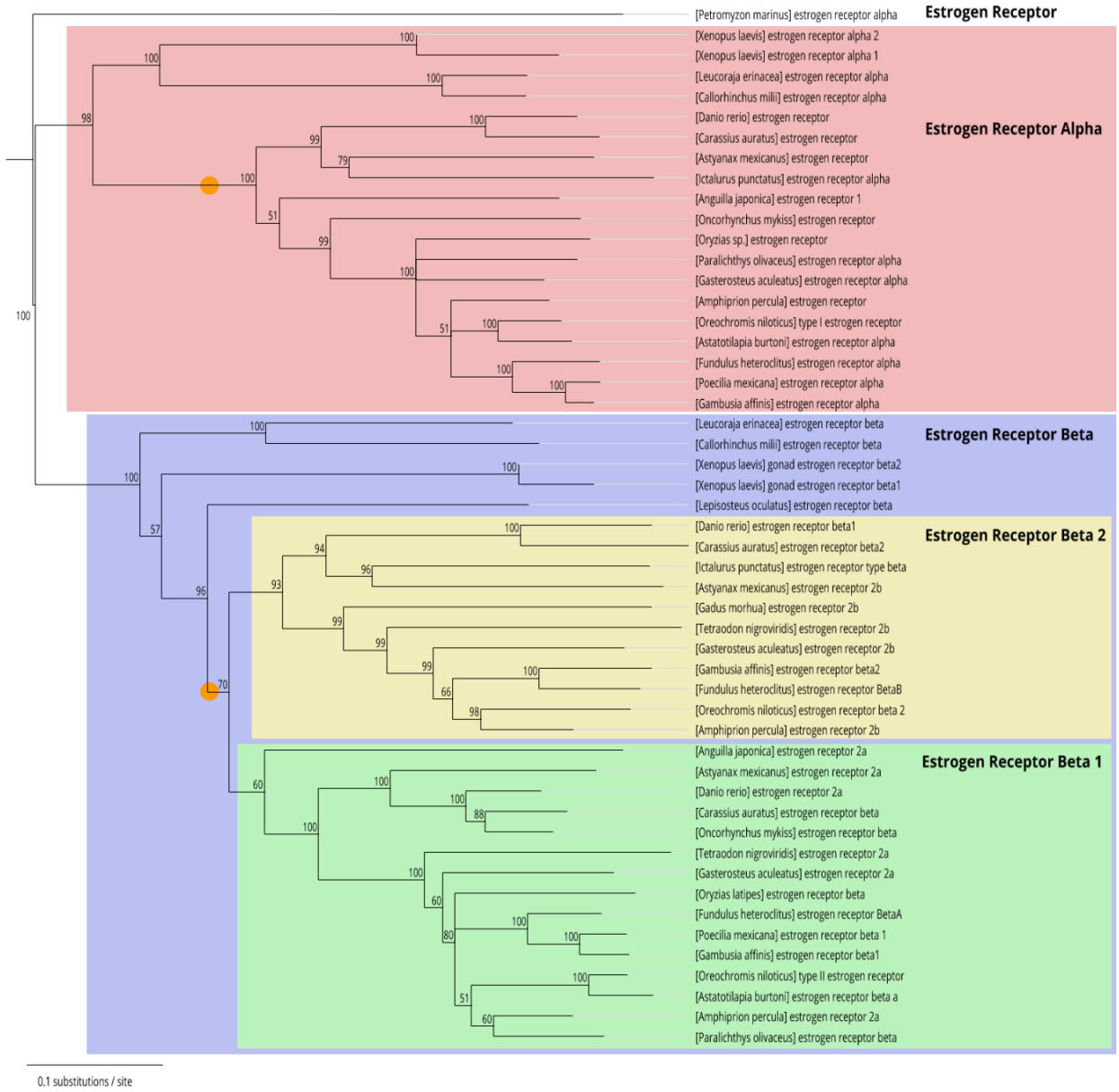
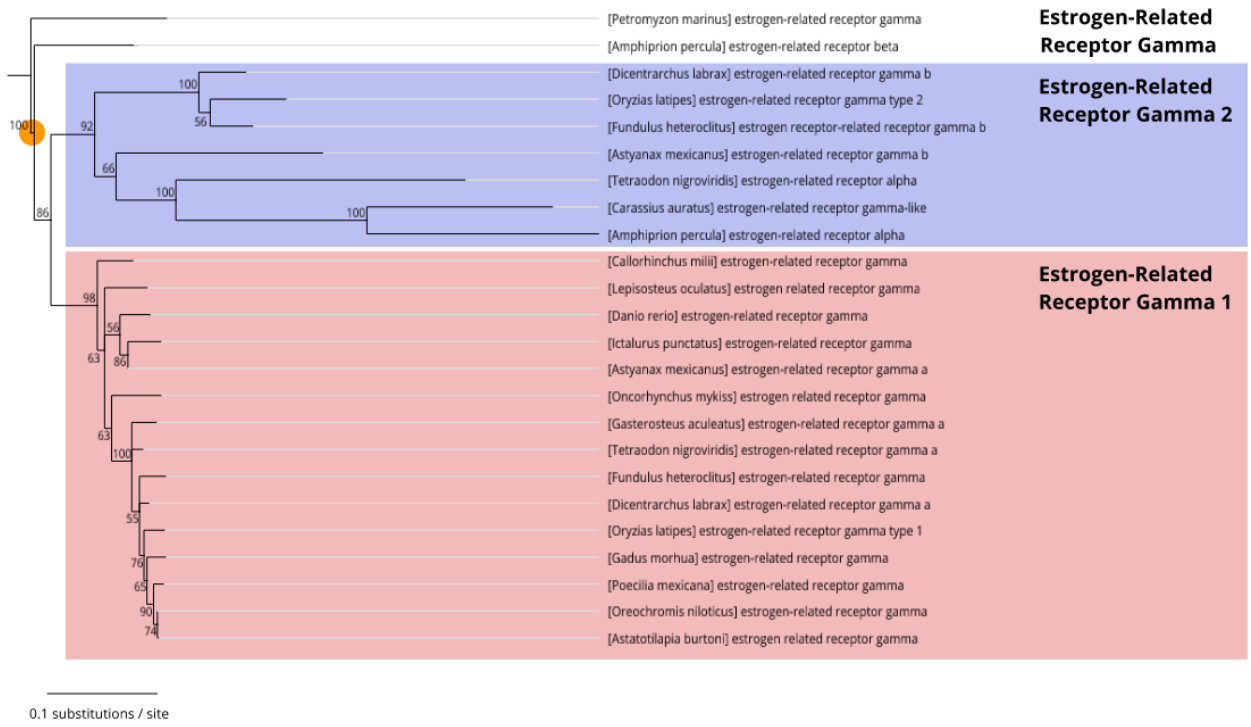
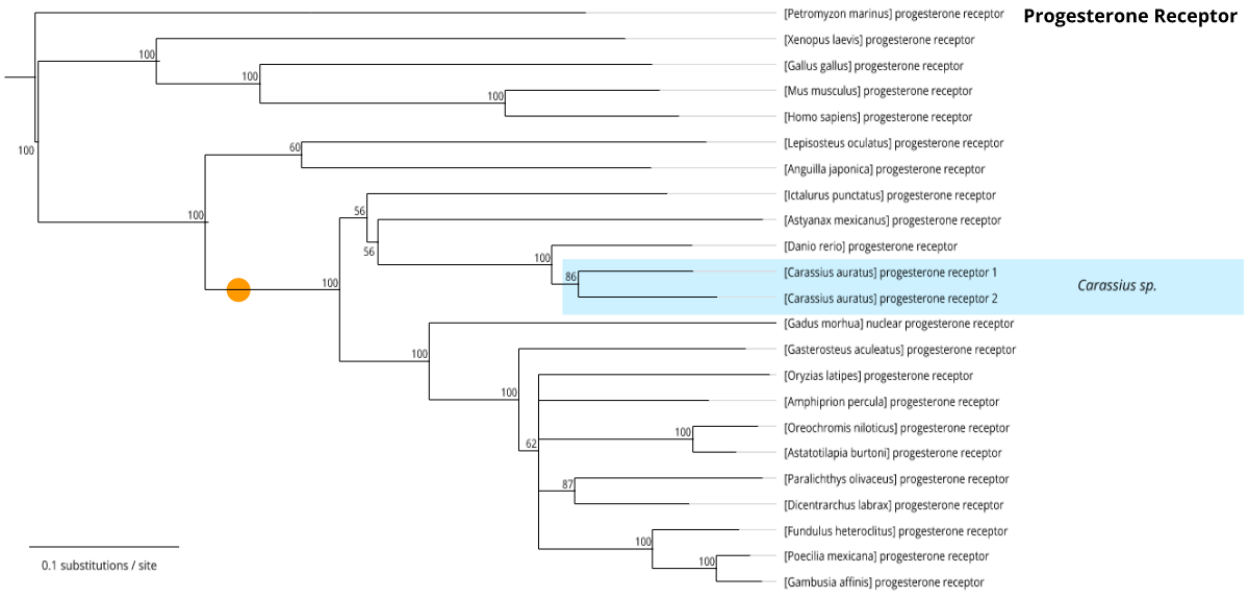

[Petromyzon marinus] estrogen receptor alpha
Estrogen Receptor
[Xenopus laevis] estrogen receptor alpha 2
[Xenopus laevis] estrogen receptor alpha 1
[Leucoraja erinacea] estrogen receptor alpha
[Callorhinchus milii] estrogen receptor alpha
[Danio rerio] estrogen receptor
[Carassius auratus] estrogen receptor
[Astyanax mexicanus] estrogen receptor
[Ictalurus punctatus] estrogen receptor alpha
[Anguilla japonica] estrogen receptor 1
[Oncorhynchus mykiss] estrogen receptor
[Oryzias sp.] estrogen receptor
[Paralichthys olivaceus] estrogen receptor alpha
[Gasterosteus aculeatus] estrogen receptor alpha
[Amphiprion percula] estrogen receptor
[Oreochromis niloticus] type I estrogen receptor
[Astatotilapia burtoni] estrogen receptor alpha
[Fundulus heteroclitus] estrogen receptor alpha
[Poecilia mexicana] estrogen receptor alpha
[Gambusia affinis] estrogen receptor alpha
Estrogen Receptor Alpha
[Leucoraja erinacea] estrogen receptor beta
[Callorhinchus milii] estrogen receptor beta
[Xenopus laevis] gonad estrogen receptor beta2
[Xenopus laevis] gonad estrogen receptor beta1
[Lepisosteus oculatus] estrogen receptor beta
Estrogen Receptor Beta
[Danio rerio] estrogen receptor beta1
[Carassius auratus] estrogen receptor beta2
[Ictalurus punctatus] estrogen receptor type beta
[Astyanax mexicanus] estrogen receptor 2b
[Gadus morhua] estrogen receptor 2b
[Tetraodon nigroviridis] estrogen receptor 2b
[Gasterosteus aculeatus] estrogen receptor 2b
[Gambusia affinis] estrogen receptor beta2
[Fundulus heteroclitus] estrogen receptor BetaB
[Oreochromis niloticus] estrogen receptor beta 2
[Amphiprion percula] estrogen receptor 2b
Estrogen Receptor Beta 2
[Anguilla japonica] estrogen receptor 2a
[Astyanax mexicanus] estrogen receptor 2a
[Danio rerio] estrogen receptor 2a
[Carassius auratus] estrogen receptor beta
[Oncorhynchus mykiss] estrogen receptor beta
[Tetraodon nigroviridis] estrogen receptor 2a
[Gasterosteus aculeatus] estrogen receptor 2a
[Oryzias latipes] estrogen receptor beta
[Fundulus heteroclitus] estrogen receptor BetaA
[Poecilia mexicana] estrogen receptor beta 1
[Gambusia affinis] estrogen receptor beta1
[Oreochromis niloticus] type II estrogen receptor
[Astatotilapia burtoni] estrogen receptor beta a
[Amphiprion percula] estrogen receptor 2a
[Paralichthys olivaceus] estrogen receptor beta
Estrogen Receptor Beta 1
0.1 substitutions / site

Figure 4.


[Petromyzon marinus] estrogen-related receptor gamma
[Amphiprion percula] estrogen-related receptor beta
Estrogen-Related Receptor Gamma
[Dicentrarchus labrax] estrogen-related receptor gamma b
[Oryzias latipes] estrogen-related receptor gamma type 2
[Fundulus heteroclitus] estrogen receptor-related receptor gamma b
[Astyanax mexicanus] estrogen-related receptor gamma b
[Tetraodon nigroviridis] estrogen-related receptor alpha
[Carassius auratus] estrogen-related receptor gamma-like
[Amphiprion percula] estrogen-related receptor alpha
Estrogen-Related Receptor Gamma 2
[Callorhinchus milii] estrogen-related receptor gamma
[Lepisosteus oculatus] estrogen related receptor gamma
[Danio rerio] estrogen-related receptor gamma
[Ictalurus punctatus] estrogen-related receptor gamma
[Astyanax mexicanus] estrogen-related receptor gamma a
[Oncorhynchus mykiss] estrogen related receptor gamma
[Gasterosteus aculeatus] estrogen-related receptor gamma a
[Tetraodon nigroviridis] estrogen-related receptor gamma a
[Fundulus heteroclitus] estrogen-related receptor gamma
[Dicentrarchus labrax] estrogen-related receptor gamma a
[Oryzias latipes] estrogen-related receptor gamma type 1
[Gadus morhua] estrogen-related receptor gamma
[Poecilia mexicana] estrogen-related receptor gamma
[Oreochromis niloticus] estrogen-related receptor gamma
[Astatotilapia burtoni] estrogen related receptor gamma
Estrogen-Related Receptor Gamma 1
0.1 substitutions / site

933 **Figure 5.**

934



935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

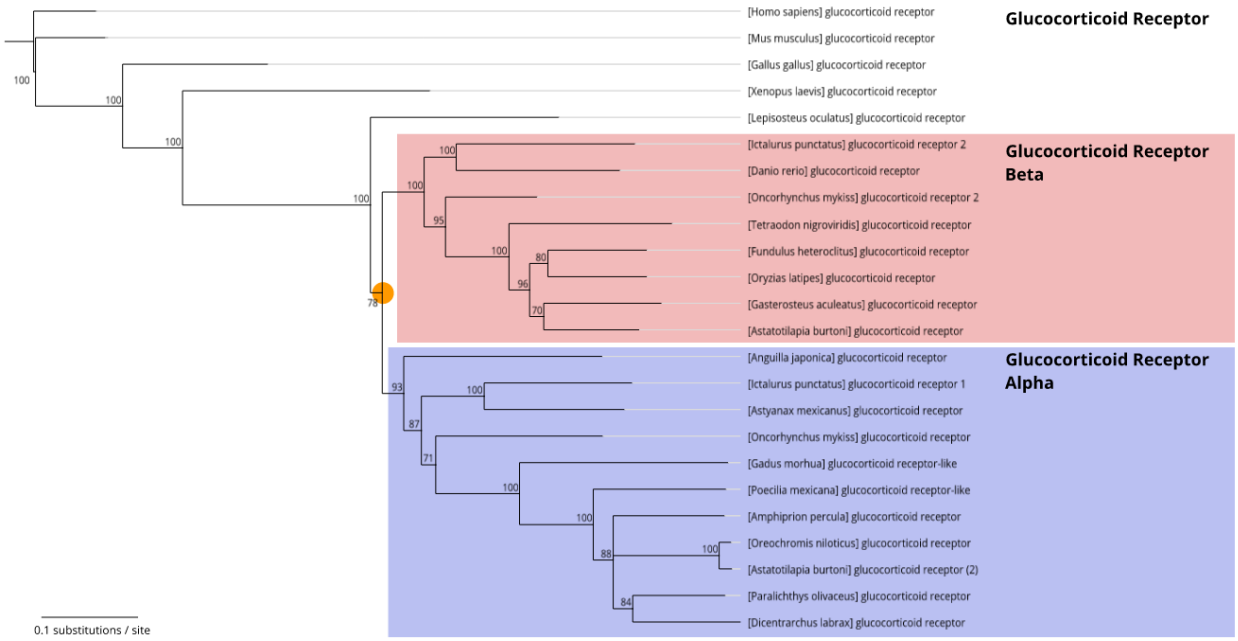
952

953

954

955

956 **Figure 6.**



957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

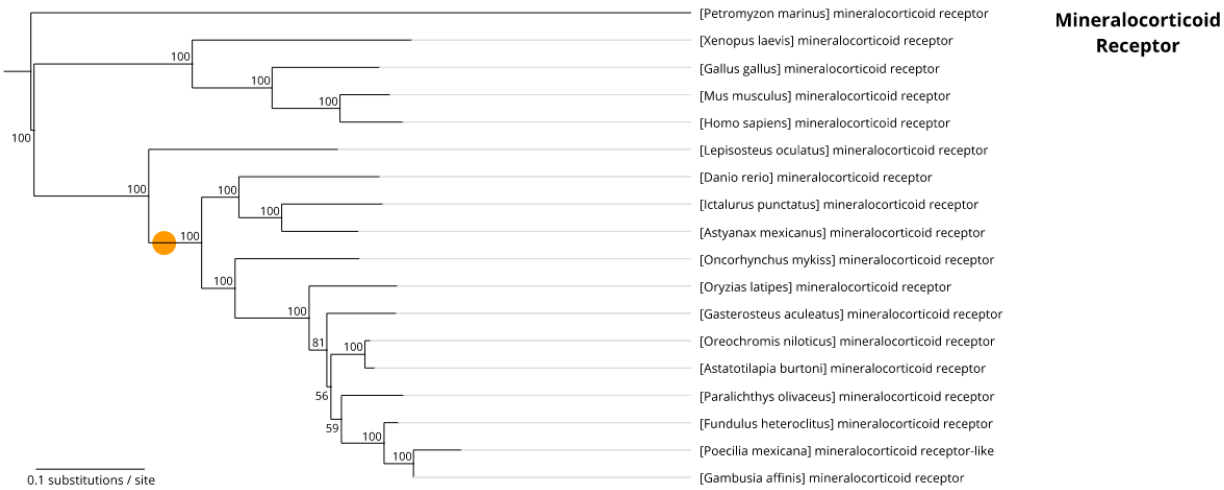
972

973

974

975 **Figure 7.**

976



977

978

979

980