

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

Kathleen M. Munley^{a,1}, Andrew P. Hoadley^{a,1}, Beau A. Alward^{a,b,*}

^a Department of Psychology, University of Houston, Houston, TX, USA

^b Department of Biology and Biochemistry, University of Houston, Houston, TX, USA

ARTICLE INFO

Keywords:

Evolution
Fishes
Genome duplication
Paralogs
Steroid hormones
Steroid receptors

ABSTRACT

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 lineage, our analysis revealed several inconsistencies in the nomenclature of steroid receptor paralogs, particularly for sequences encoding estrogen receptor beta (ER β). 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.

1. Introduction

Over the past few decades, teleost fishes have emerged as important model systems in numerous fields, including molecular biology, genetics, evolution, neuroendocrinology, and behavior (reviewed in Glauser and Neuhauss, 2014; Munley and Alward, 2023; Sato and Nishida, 2010; Volff, 2005). Teleost fishes (Class: Actinopterygii, Infraclass: Teleostei) are the largest and most speciose 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 remarkable variability 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. Indeed, the development of high-throughput and rapid DNA sequencing

methods has enabled researchers to sequence the 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), orange clownfish (*Amphiprion percula*; Lehmann et al., 2019), olive flounder (*Paralichthys olivaceus*; Shao 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

* Corresponding author at: Department of Psychology, University of Houston, Houston, TX, USA.

E-mail address: balward@uh.edu (B.A. Alward).

¹ These authors share first authorship.

gene editing.

The immense diversity present in teleosts is likely due, at least in part, to a single whole-genome duplication (WGD) event that occurred approximately 350 mya in the common ancestor of all teleost fishes (Brunet et al., 2006). This event, 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), is often referred to as the teleost-specific WGD [TS-WGD, also known as the third-round WGD or 3R WGD; reviewed in Glasauer and Neuhauss, 2014]. After a WGD event, redundant genes can face one or more of several fates. Frequently, relaxed selective pressure on a duplicated gene leads to the buildup of deleterious mutations, which results in non-functionalization (Ohno, 1970). In some cases, however, one or more gene paralogs may be under positive selection due to the accumulation of mutations conferring novel functions, a process known as neo-functionalization (Force et al., 1999; Ohno, 1970). Alternatively, duplicate genes may undergo sub-functionalization, in which both paralogs are maintained due to the complementary partitioning of ancestral functions between two daughter genes, such that their

complimentary actions comprise the total function of their single ancestral gene (Force et al., 1999). Collectively, these duplicate, redundant genes may serve as an important source of novelty on which natural selection can act, resulting in species diversification and evolutionary innovation within the teleost lineage.

Among the genes that have been duplicated following the TS-WGD are steroid receptors, a family of ligand-dependent nuclear receptors that function as signal transducers and transcription factors (Carson-Jurica et al., 1990, Gronemeyer et al., 2004). Steroid receptors modulate suites of physiological mechanisms, such as homeostatic functions, growth and development, cellular signaling, sexual differentiation and reproduction, and behavior (Szego et al., 2003; Whirledge and Cidlowski, 2019). Steroid receptors consist of five major groups: 1) the androgen receptors (ARs), 2) estrogen receptors (ERs), and 3) progesterone receptors (PRs), which mediate the actions of androgens (e.g., testosterone, dihydrotestosterone, and 11-ketotestosterone), estrogens (e.g., estradiol), and progestins (e.g., progesterone) on the development and maintenance of the reproductive, musculoskeletal, cardiovascular, immune, and central nervous systems (reviewed in Brinton et al., 2008;

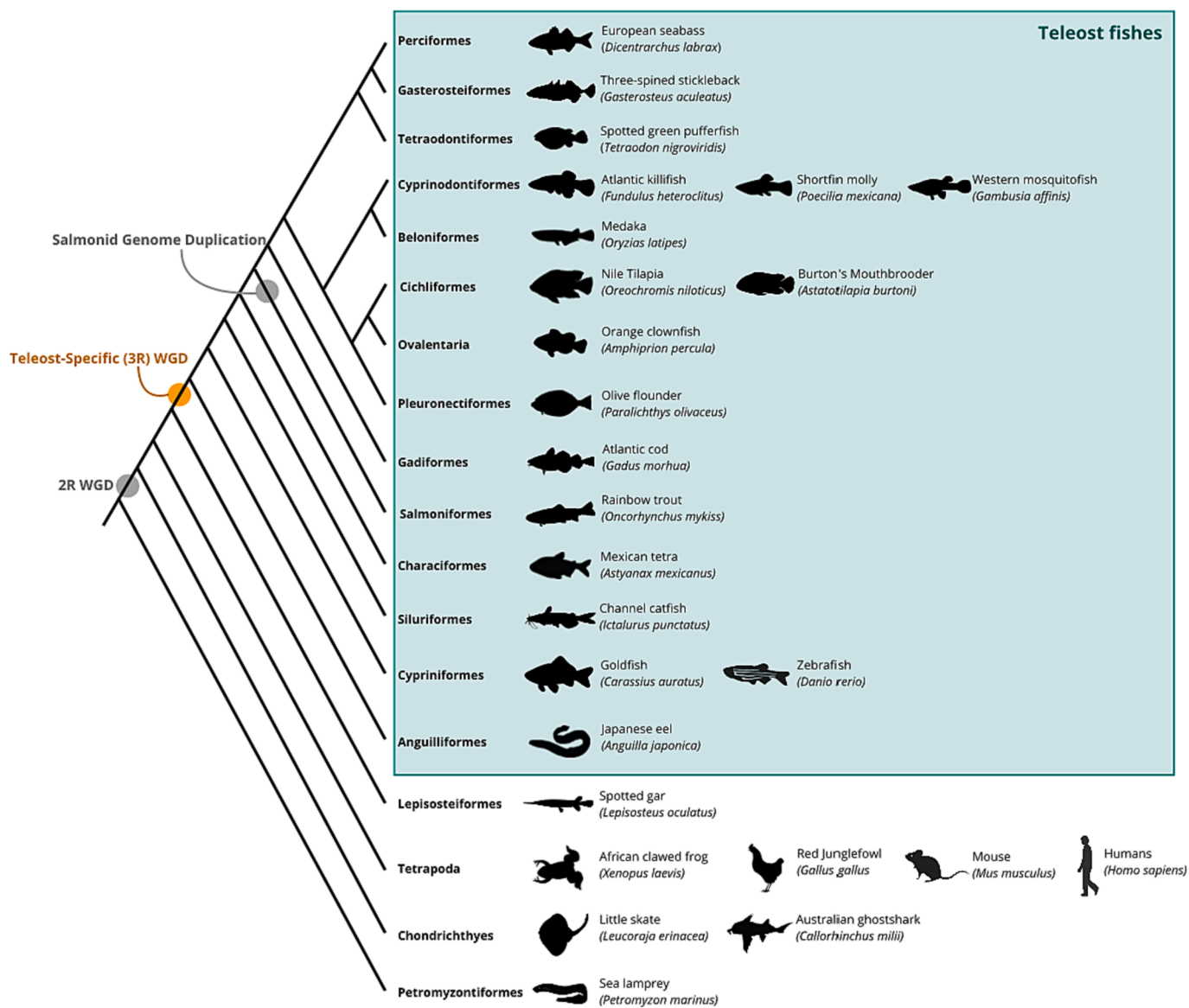


Fig. 1. Phylogeny identifying the evolutionary relationships between taxa included in this study [after Betancur-R et al. (2017) and Hughes et al. (2018)]. Gray and orange circles indicate genome duplication events. Abbreviations: 2R, second-round; 3R, third-round; WGD, whole-genome duplication. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Chen et al., 2022; Davey and Grossmann, 2016; Mani et al., 1997), and 4) the glucocorticoid receptors (GRs) and 5) mineralocorticoid receptors (MRs), which facilitate the actions of glucocorticoids (e.g., cortisol) and mineralocorticoids (e.g., aldosterone) on metabolism, growth and development, cardiovascular function, ion transport, and salt balance (reviewed in Pascual-Le Tallec and Lombès, 2005; Timmermans et al., 2019; Wada, 2008). Prior work suggests that extant steroid receptors evolved from an ancestral estrogen receptor that preceded the origin of bilaterally symmetric animals through the process of ligand exploitation (Baker, 2019; Dube et al., 2023; Eick and Thornton, 2011, Hochberg et al., 2020; Thornton, 2001; Thornton et al., 2003). More specifically, these studies suggest that ERs evolved from a more ancient lineage, whereas PRs, ARs, GRs, and MRs are more recent in origin.

The novel steroid receptor paralogs resulting from the TS-WGD present both opportunities and challenges for researchers. For instance, novel steroid receptor paralogs provide opportunities to dissect their functions and discover their fundamental roles across species (reviewed in Alward et al., 2023; Jackson et al., 2023). Some challenges include difficulty in manipulating the functions of these genes (e.g., two steroid receptors bind the same ligand in many teleosts, whereas only one steroid receptor does in other vertebrates). Another practical challenge is there is not a cohesive naming system for steroid receptors, despite previous efforts to unify the nomenclature of this nuclear receptor subfamily (Nuclear Receptors Nomenclature Committee, 1999). This challenge at first may seem inconsequential, but the lack of a logical nomenclature system for novel teleost steroid receptor paralogs 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; Jackson et al., 2023; Juntti, 2019; Munley and Alward, 2023), including those related to steroid hormone function (Alward et al., 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 naming 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. Specifically, 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 receptors, as the nomenclature issue across teleosts is especially apparent for this class of molecules. We predicted that steroid receptor protein sequences would generally group based on the established phylogeny of the teleost fish lineage, but that there would be a few instances in which protein sequences of certain steroid receptor paralogs show greater similarity to a different group of related paralogs than that for which the sequence is currently named. Together, our findings bridge a critical gap among researchers who study teleost fishes by providing a consistent nomenclature system for steroid receptors, clarifying scientific communication and promoting the progression of teleost research during a time of rapid advancements in genetic technology. We hope that our approach will build on similar efforts (e.g., Ocampo Daza et al., 2022; Olender et al., 2020; Theofanopoulou et al., 2021; Tsai, 2018) to establish consistent naming systems for species and/or molecular systems with inconsistent nomenclature.

2. Materials and methods

2.1. Species selection

To assess the relatedness of steroid receptor protein sequences among teleost fishes, we chose 18 species representing 12 different orders from the Infraclass Teleostei [Order Cypriniformes (carps, minnows, and loaches), Order Cyprinodontiformes (toothcarps), Order

Characiformes (tetras and piranhas), Order Siluriformes (catfish), Order Salmoniformes (salmonids), Order Gadiformes (cods), Order Belontiiformes (medakas, ricefishes, and needlefishes), Order Cichliformes (cichlids and convict blennies), Order Perciformes (perches, darters, basses, and groupers), Order Gasterosteiformes (sticklebacks, pipefishes, and seahorses), Order Pleuronectiformes (flatfishes), Order Tetraodontiformes (puffers and filefishes), and family Pomacentridae in the subseries Ovalentaria (damsel fishes and clownfishes)]. Specifically, we selected the following species for phylogenetic analysis: *Anguilla japonica* (Japanese eel), *Danio rerio* (zebrafish), *Carassius auratus* (goldfish), *Fundulus heteroclitus* (Atlantic killifish), *Poecilia mexicana* (shortfin molly), *Gambusia affinis* (western mosquitofish), *Astyanax mexicanus* (Mexican tetra), *Ictalurus punctatus* (channel catfish), *Oncorhynchus mykiss* (rainbow trout), *Gadus morhua* (Atlantic cod), *Oryzias latipes* (medaka), *Astatotilapia (Haplochromis) burtoni* (Burton's mouthbrooder), *Oreochromis niloticus* (Nile tilapia), *Amphiprion percula* (orange clownfish), *Dicentrarchus labrax* (European seabass), *Gasterosteus aculeatus* (three-spined stickleback), *Paralichthys olivaceus* (olive flounder), and *Tetraodon nigroviridis* (spotted green pufferfish). These species were chosen based on: 1) the availability of protein sequences for steroid receptors, many of which were acquired via whole-genome sequencing projects and, thus, are from teleost species which are commonly studied in a research setting, and 2) to represent the diversity of fishes within the teleost lineage.

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

2.2. Protein sequence collection and criteria

Protein sequences for androgen, estrogen, progesterone, glucocorticoid, and mineralocorticoid receptors, including paralogs and isoforms resulting from the TS-WGD, were acquired from the National Center for Biotechnology Information (NCBI) Protein and UniProt databases. This search yielded a total of 213 protein sequences across 26 species (36 androgen receptor sequences, 93 estrogen receptor sequences, 27 progesterone receptor sequences, 31 glucocorticoid receptor sequences, and 26 mineralocorticoid receptor sequences). Sequences for the following steroid receptors, paralogs, and isoforms were included in our analyses: AR α , AR β , AR β 1, AR β 2, AR α , AR β , ER, ER α , ER β , ER β 1, ER β 2, ER2 α , ER2 β , PR, PR1, PR2, GR, GR1, GR2, and MR. Sequences that were described as alternative splice variants or as partial sequences were excluded from analysis. Available or inferred full-length amino acid sequences were downloaded from NCBI Protein and UniProt in FASTA format for inclusion in phylogenetic analysis (see Section 2.3). All sequences were checked manually for the inclusion of 3' and 5' untranslated regions to ensure that they represented full-length amino acid sequences, and partial sequences were excluded from phylogenetic analysis. A complete list of protein sequences that were used for phylogenetic analysis, along with their sources, accession numbers, and references, is available in the Supplementary Material (Tables S1-S5).

2.3. Phylogenetic analyses

For each major group of steroid receptors, protein sequences were used to generate a molecular phylogeny. Construction of multiple alignments of steroid receptor amino acid sequences with the neighbor-joining method were performed using the MUSCLE Alignment version 3.8.425 (Edgar, 2004; Saitou and Nei, 1987). Sequence alignments were reviewed manually to ensure that registered sequences contained canonical ligand and DNA binding domain(s). To estimate molecular phylogenies, trees were constructed with PhyML 3.3.20180214 (Guindon et al., 2010) using the LG substitution model, which had proportion of invariable sites fixed at 0, an estimated gamma distribution parameter, and was optimized for topology/length/rate. Unrooted trees are shown as rooted using the most basal organism for which a protein sequence was available [*P. marinus* (sea lamprey), *L. erinacea* (little skate), *C. milii* (Australian ghostshark), or *H. sapiens* (human)]. Branch support for the PhyML analyses were calculated using bootstrap resampling with 100 replicates. All analyses were performed using Geneious Prime version 2022.2.2 (<https://www.geneious.com>).

2.4. Approach for steroid receptor nomenclature

Our approach for determining a nomenclature system for protein and gene sequences encoding steroid receptors was based on clustering within our phylogenies. We hypothesized that collections of closely related protein sequences across species would form large clusters that generally correspond to the current understanding of the steroid receptor paralogs present in teleosts. For example, we predicted that there would be two distinct clusters for the ARs because two distinct paralogs have been identified for these genes in several species of teleosts (Lorin et al., 2015; Ogino et al., 2016).

Protein names were assigned for each cluster based on majority rule, in which the name proposed for a given cluster reflected the current protein name that was listed for most teleost species included in our analyses. Once protein names were assigned, we proposed names for their corresponding genes (see Table 1 for a basic consensus of all steroid receptor proteins and genes and Supplementary Material, Tables S1-S5 for a full list of proposed protein and gene names for each species). We propose a new system for naming novel paralogous steroid receptor genes using a combination of the existing nomenclature system in teleosts and the steroid receptor gene nomenclature system in mice (*M. musculus*) as a model.

3. Results

Our phylogenetic approach led to clear classifications of a variety of steroid receptor protein sequences across species. The proposed names for the steroid receptor proteins are shown in each figure in distinct, broad clusters delineated by differently colored, transparent rectangles. The corresponding gene name for each protein is provided in Table 1. Based on clusters in the phylogenies, we propose names across defined clusters for each steroid receptor.

3.1. Androgen receptors

Our analysis of AR strongly resolves the split between the teleost-specific AR α and AR β paralogs (Fig. 2). Overall, protein sequences for AR, AR α , and AR β clustered into two well-supported clades (96 % bootstrap value) that generally agreed with established phylogenetic relationships in the teleost lineage (Fig. 1). The first clade primarily contained AR α sequences (“Androgen Receptor Alpha” group), whereas the second clade mostly consisted of AR β sequences (“Androgen Receptor Beta” group; Fig. 2). These results agree with previous work showing that many species of teleosts possess two distinct AR paralogs (Lorin et al., 2015; Ogino et al., 2016). Like the current naming system for AR genes in mice (Juntti et al., 2010), we propose referring to the

Table 1

Proposed nomenclature for proteins and genes encoding steroid receptors in teleost fishes.

Protein	Gene	Other Gene Aliases	Figure
AR		<i>nr3c4</i> , <i>ar</i>	
AR α	<i>ar1</i>	<i>ara</i>	Fig. 2
AR β	<i>ar2</i>	<i>arb</i>	Fig. 2
ER			
ER α	<i>esr1</i>	<i>nr3a1</i> , <i>era</i> , <i>esr</i>	Fig. 3
ER β 1	<i>esr2a</i>	<i>erb1</i>	Fig. 3
ER β 2	<i>esr2b</i>	<i>erb2</i>	Fig. 3
PR	<i>pgr</i>	<i>nr3c3</i> , <i>pr</i>	
PR α	<i>pgr1</i>	N/A	Fig. 4
PR β	<i>pgr2</i>	N/A	Fig. 4
GR	<i>nr3c1</i>	<i>gr</i> , <i>gcr</i>	
GR α	<i>nr3c1a</i>	<i>gr1</i>	Fig. 5
GR β	<i>nr3c1b</i>	<i>gr2</i>	Fig. 5
MR	<i>nr3c2</i>	<i>mr</i> , <i>mcr</i>	Fig. 6

Proposed protein and gene names for androgen receptors, estrogen receptors, progesterone receptors, glucocorticoid receptors, and mineralocorticoid receptors in teleost fishes based on phylogenetic analysis. Alternative aliases for each gene based on existing nomenclature systems and the location of figures containing phylogenies associated with each steroid receptor subtype are also shown. Abbreviations: *ar*, androgen receptor gene; AR, androgen receptor protein; *ar1*, androgen receptor 1 gene; *ar2*, androgen receptor 2 gene; *ara*, androgen receptor alpha gene; AR α , androgen receptor alpha protein; *arb*, androgen receptor beta gene; AR β , androgen receptor beta protein; ER, estrogen receptor protein; ER α , estrogen receptor alpha protein; ER β , estrogen receptor beta protein; ER β 1, estrogen receptor beta 1 protein; ER β 2, estrogen receptor beta 2 protein; *era*, estrogen receptor a gene; *erb1*, estrogen receptor b1 gene; *erb2*, estrogen receptor b2 gene; *esr*, estrogen receptor gene; *esr1*, estrogen receptor 1 gene; *esr2a*, estrogen receptor 2a gene; *esr2b*, estrogen receptor 2b gene; *gcr*, glucocorticoid receptor gene; *gr*, glucocorticoid receptor gene; GR, glucocorticoid receptor protein; *gr1*, glucocorticoid receptor 1 gene; *gr2*, glucocorticoid receptor 2 gene; GR α , glucocorticoid receptor alpha protein; GR β , glucocorticoid receptor beta protein; *mcr*, mineralocorticoid receptor gene; *mr*, mineralocorticoid receptor protein; *nr3a1*, nuclear receptor subfamily 3 group A member 1 gene; *nr3c1*, nuclear receptor subfamily 3 group C member 1 gene; *nr3c1a*, nuclear receptor subfamily 3 group C member 1a gene; *nr3c1b*, nuclear receptor subfamily 3 group C member 1b gene; *nr3c2*, nuclear receptor subfamily 3 group C member 2 gene; *nr3c3*, nuclear receptor subfamily 3 group C member 3 gene; *nr3c4*, nuclear receptor subfamily 3 group C member 4 gene; *pgr*, progesterone receptor gene; *pgr1*, progesterone receptor 1 gene; *pgr2*, progesterone receptor 2 gene; *pr*, progesterone receptor gene; PR, progesterone receptor protein; PR α , progesterone receptor alpha protein; PR β , progesterone receptor beta protein.

genes for an AR protein with an *ar* base. Thus, genes encoding AR α will be called *ar1*, and genes encoding AR β will be called *ar2* (Table 1).

The AR sequences for the tetrapod species clustered with *L. erinacea* (little skate) at the base of the AR phylogenetic tree (100 % bootstrap value; Fig. 2). Most AR α sequences clustered into the “Androgen Receptor Alpha” group (100 % bootstrap value), while the majority of AR β sequences grouped into the “Androgen Receptor Beta” group (62 % bootstrap value). Importantly, we identified several discrepancies in how ARs are currently named in the NCBI Protein and UniProt databases. Specifically, sequences labeled “androgen receptor” for *G. morhua* (Atlantic cod) and *A. burtoni* (Burton’s mouthbrooder) clustered with the AR α sequences in the “Androgen Receptor Alpha” group (96 % bootstrap values), whereas similarly named sequences for *T. nigroviridis* (spotted green pufferfish), *D. rerio* (zebrafish), *C. auratus* (goldfish), *A. mexicanus* (Mexican tetra), *D. labrax* (European seabass), *A. percula* (orange clownfish), *F. heteroclitus* (Atlantic killifish), and *G. affinis* (western mosquitofish) clustered with AR β sequences in the “Androgen Receptor Beta” group (96 % bootstrap values). Interestingly, the AR α and AR β sequences for *O. mykiss* (rainbow trout) and *A. japonica* (Japanese eel) clustered in the “Androgen Receptor Beta” group, indicating that sequences referred to as AR α in these species are more similar to AR β sequences in other teleost species than to other AR α sequences. These

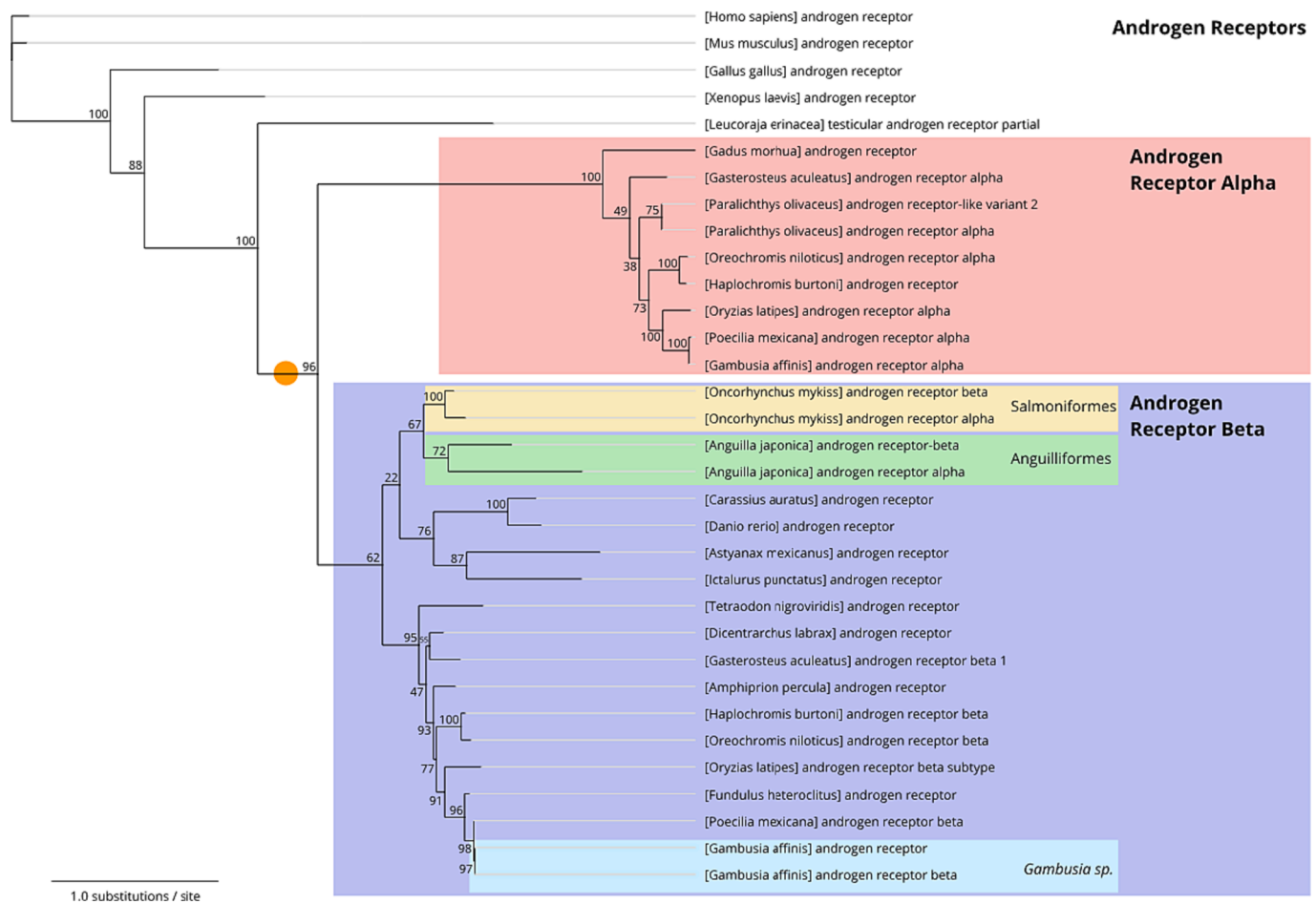


Fig. 2. Consensus phylogenetic tree of androgen receptors. Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the approximate position of the teleost-specific whole-genome duplication event (TS-WGD). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

results suggest that these sequences would be more appropriately called “Androgen Receptor Beta 1” and “Androgen Receptor Beta 2” for each species.

3.2. Estrogen receptors

Our phylogenetic analysis for ERs revealed several inconsistencies in the nomenclature of estrogen receptor paralogs in teleost fishes based on protein sequence similarities, particularly for ER β (Fig. 3). Protein sequences for ER, ER α , and ER β clustered into two well-supported clades (100 % bootstrap values). The first clade largely consisted of ER α sequences (“Estrogen Receptor Alpha” group), whereas the second clade primarily contained ER β sequences (“Estrogen Receptor Beta” group). Furthermore, there were two additional clusters within the “Estrogen Receptor Beta” group that were supported by our analysis (86 % bootstrap value): 1) the “Estrogen Receptor Beta 2” group, which was mostly comprised of ER β 2 sequences (100 % bootstrap value), and 2) the “Estrogen Receptor Beta 1” group, which mainly included ER β 1 sequences (69 % bootstrap value). These results are consistent with the expression of two distinct ER β genes, *esr2a* and *esr2b*, in medaka (*O. latipes*; Nishiike et al., 2021) and Burton’s mouthbrooder (*A. burtoni*; Maruska et al., 2020). We propose calling the genes for ER α , ER β 1, and ER β 2 *esr1*, *esr2a*, and *esr2b*, respectively (Table 1).

All sequences labeled “estrogen receptor” in NCBI Protein and UniProt, including those of *D. rerio* (zebrafish), *C. auratus* (goldfish), *A. mexicanus* (Mexican tetra), *O. mykiss* (rainbow trout), *O. latipes* (medaka), and *A. percula* (orange clownfish), as well as “estrogen

receptor 1” for *A. japonica* (Japanese eel) and the “type I estrogen receptor” for *O. niloticus* (Nile tilapia), clustered with the ER α sequences in the “Estrogen Receptor Alpha” group (≥ 90 % bootstrap values; Fig. 3). Conversely, we found some discrepancies in nomenclature for the “Estrogen Receptor Beta” clade based on protein sequence alignment. The tetrapod outgroup sequences for ER β [“estrogen receptor beta 1” and “estrogen receptor beta 2” for *X. laevis* (African clawed frog)] clustered with *L. erinacea* (little skate), *C. milii* (Australian ghostshark), and *L. oculatus* (spotted gar), species that diverged from teleosts prior to the TS-WGD, and these sequences were basal to all teleost ER β sequences that were included in our analysis (100 % bootstrap value). These data support the hypothesis that the estrogen receptor alpha and beta divergence is more ancient than the TS-WGD (Kelley and Thackray, 1999). The majority of ER β 2 sequences, which are labeled “estrogen receptor beta 2” and “estrogen receptor 2b,” clustered into a single well-supported clade (“Estrogen Receptor Beta 2,” 86 % bootstrap value) that was consistent with known phylogenetic relationships in the teleost lineage (Fig. 1). Importantly, there was one sequence within the “Estrogen Receptor Beta 2” group that was unexpected: the “estrogen receptor beta 1” sequence for *D. rerio* (zebrafish, 100 % bootstrap value). The sequences “estrogen receptor type beta” for *I. punctatus* (channel catfish) and “estrogen receptor BetaB” for *F. heteroclitus* (Atlantic killifish) also clustered with the ER β 2 sequences in the “Estrogen Receptor Beta 2” group (100 % bootstrap values). Similarly, most ER β 1 sequences, which are labeled “estrogen receptor beta 1” or “estrogen receptor 2a,” clustered into a single clade (69 % bootstrap value) that typically followed known phylogenetic relationships in teleost fishes. Moreover, all

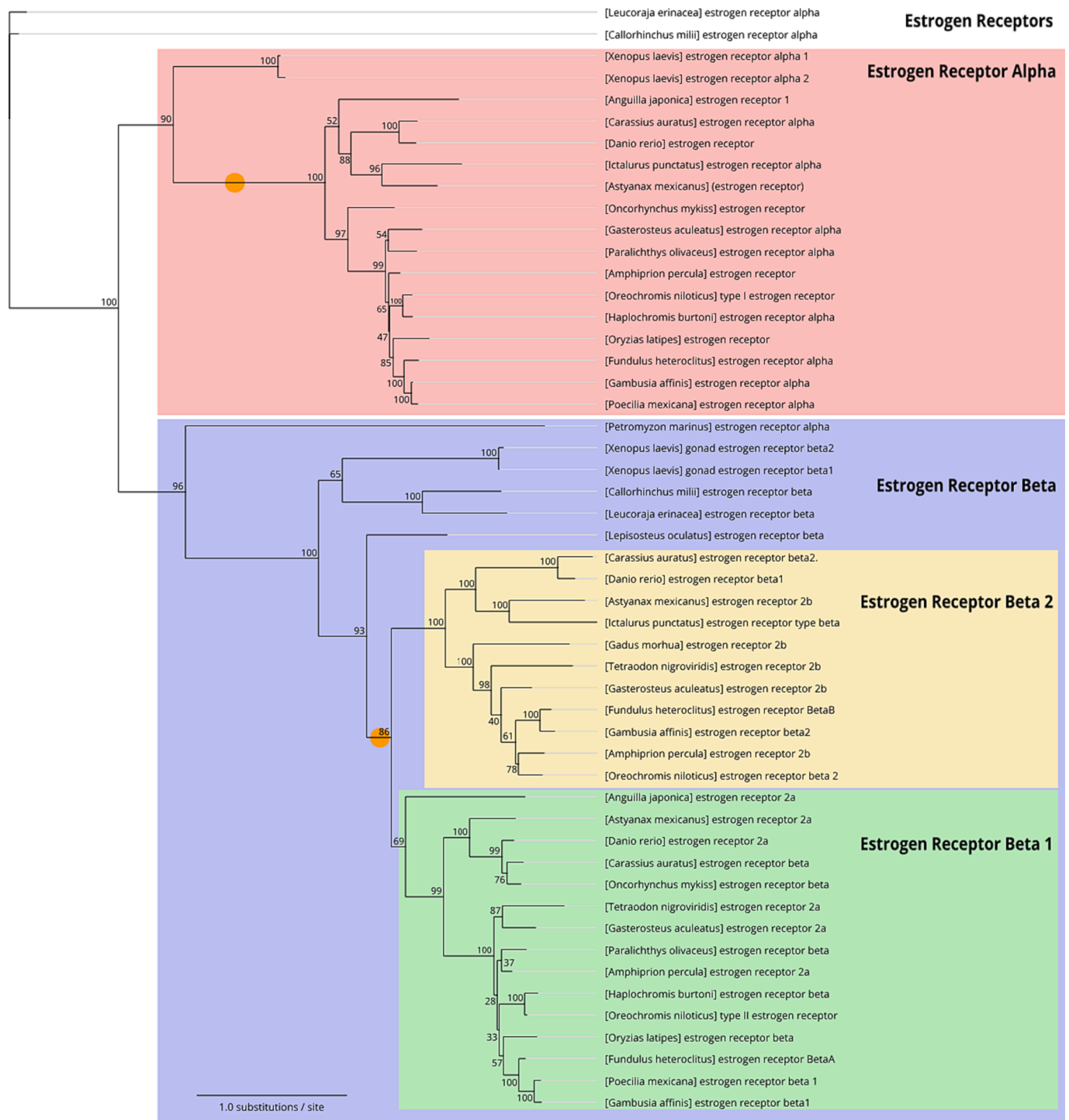


Fig. 3. Consensus phylogenetic tree of estrogen receptors alpha (α) and beta (β). Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). Orange circles represent the approximate position of the teleost-specific whole-genome duplication event (TS-WGD), which impacted the independent duplication of estrogen receptor β , but not estrogen receptor α . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sequences labeled “estrogen receptor beta” in NCBI Protein and UniProt, including those of *C. auratus* (goldfish), *O. mykiss* (rainbow trout), *O. latipes* (medaka), and *P. olivaceus* (olive flounder), as well as “estrogen receptor BetaA” for *F. heteroclitus* (Atlantic killifish) and “type II estrogen receptor” for *O. niloticus* (Nile tilapia), clustered with the ER β 1 sequences in the “Estrogen Receptor Beta 1” group (≥ 80 % bootstrap values; Fig. 3).

3.3. Progesterone receptors

Our analysis of PR suggests that there was no widespread retention of orthologous PR genes following the TS-WGD, consistent with previous studies (Ren et al., 2019; Fig. 4). The tetrapod sequences for PR clustered as a sister group to all teleost sequences (100 % bootstrap value), and all teleost PR sequences clustered into a single well-supported clade (100 % bootstrap value) that mostly agreed with established phylogenetic

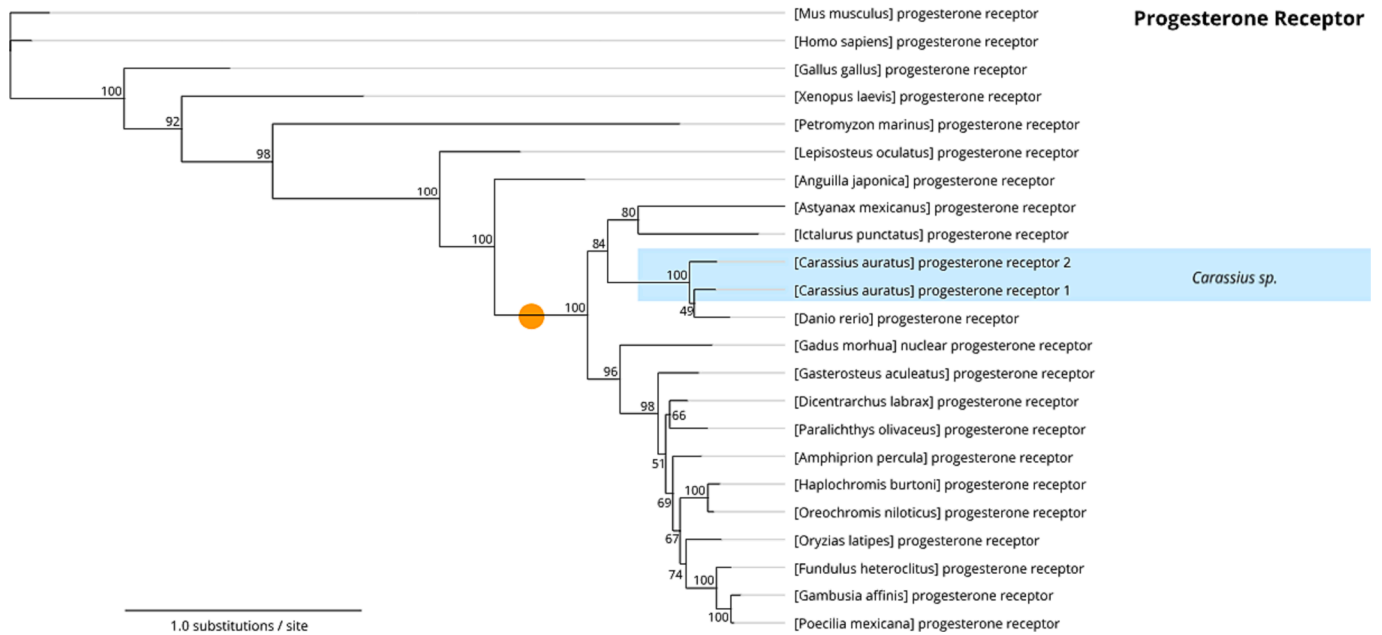


Fig. 4. Consensus phylogenetic tree of progesterone receptor. Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the approximate position of the teleost-specific whole-genome duplication event (TS-WGD). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

relationships in this lineage (Fig. 1). Importantly, our search yielded two PR sequences in *C. auratus* (goldfish; PR1 and PR2), which clustered together in a clade with *D. rerio* (zebrafish, PR) (100 % bootstrap value; Fig. 4). No additional PR sequences have been identified in closely related teleost species (e.g., zebrafish; Tang et al., 2016), suggesting that these sequences represent a gene duplication that is specific to the goldfish lineage (see Chen et al., 2019). The current name for the PR gene is *pgr*, and this nomenclature was consistent across all teleost species examined in our study.

3.4. Glucocorticoid receptors

Our analysis revealed several inconsistencies in the nomenclature of GR (Fig. 5). Protein sequences for GR, GR1, and GR2 clustered into two clades (100 % bootstrap value) that were consistent with known phylogenetic relationships in teleosts (Fig. 1). The first clade mostly consisted of GR2 sequences (“Glucocorticoid Receptor Beta” group), while the second clade primarily contained GR1 sequences (“Glucocorticoid Receptor Alpha” group; Fig. 5), suggesting that two

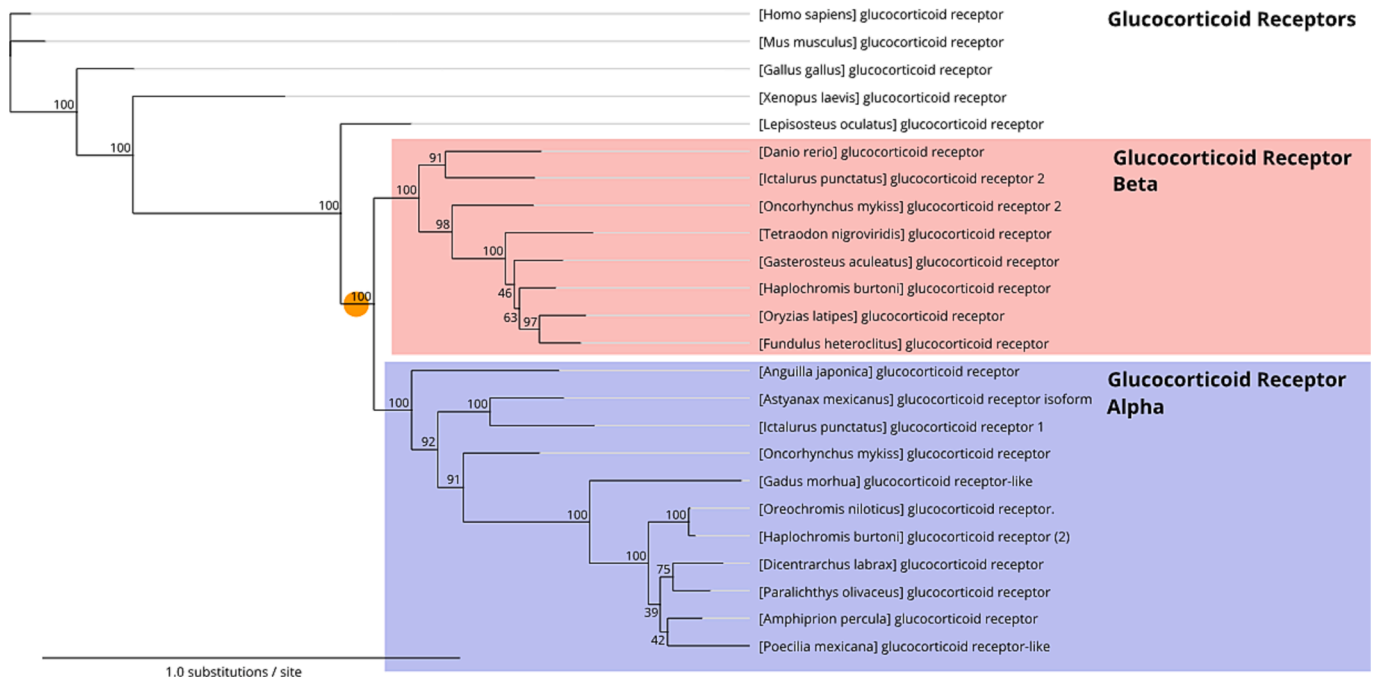


Fig. 5. Consensus phylogenetic tree of glucocorticoid receptors. Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the approximate position of the teleost-specific whole-genome duplication event (TS-WGD). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

glucocorticoid receptors persisted in the teleost lineage following the TS-WGD. To remain consistent with the naming systems for other steroid receptor proteins, we propose calling “glucocorticoid receptor 1” GR α and “glucocorticoid receptor 2” GR β , with gene names of *nr3c1a* and *nr3c1b*, respectively (Table 1).

The tetrapod outgroup sequences for GR clustered with that of *L. oculatus* (spotted gar) at the base of the GR phylogenetic tree (100 % bootstrap value; Fig. 5). Interestingly, most GR sequences used in our analysis were labeled “glucocorticoid receptor,” yet these sequences still grouped into two separate clades. Specifically, sequences labeled “glucocorticoid receptor” for *D. rerio* (zebrafish), *T. nigroviridis* (spotted green pufferfish), *F. heteroclitus* (Atlantic killifish), *O. latipes* (medaka), *G. aculeatus* (three-spined stickleback), and *A. burtoni* (Burton’s mouthbrooder) clustered in the “Glucocorticoid Receptor Beta” group (100 % bootstrap value), whereas sequences labeled “glucocorticoid receptor” for *A. japonica* (Japanese eel), *A. mexicanus* (Mexican tetra), *O. mykiss* (rainbow trout), *A. percula* (orange clownfish), *O. niloticus* (Nile tilapia), *D. labrax* (European seabass), and *P. olivaceus* (olive flounder), in addition to “glucocorticoid receptor-like” for *G. morhua* (Atlantic cod) and *P. mexicana* (shortfin molly), clustered in the “Glucocorticoid Receptor Alpha” group (100 % bootstrap value). A few species in our analysis had two different GR sequences, including *I. punctatus* (channel catfish; “glucocorticoid receptor 1” and “glucocorticoid receptor 2”), *A. burtoni* [Burton’s mouthbrooder; “glucocorticoid receptor” and “glucocorticoid receptor (2)”] and *O. mykiss* (rainbow trout; “glucocorticoid receptor” and “glucocorticoid receptor 2”). Although these sequences generally clustered with other similarly named sequences, the “glucocorticoid receptor (2)” sequence for *A. burtoni* clustered with sequences in the “Glucocorticoid Receptor Alpha” group, suggesting that this sequence should be called GR α . Because duplicate GR sequences were identified in highly derived clades (e.g., *A. burtoni*), our results suggest that GR paralogs may exist, but remain undescribed in some teleost species.

3.5. Mineralocorticoid receptor

Our analysis of MR indicates no widespread retention of orthologous MR genes following the TS-WGD (Fig. 6). The tetrapod outgroup sequences for MR clustered with *P. marinus* (sea lamprey) and were basal to all teleost MR sequences (100 % bootstrap value). MR protein sequences from teleost fishes clustered into a single well-supported clade (100 % bootstrap value) that was consistent with known phylogenetic relationships in this lineage (Fig. 1). While one of the sequences included in our analysis was labeled “mineralocorticoid receptor-like” [from

P. mexicana (shortfin molly)], this protein sequence confidently clustered with other MRs in our analysis (100 % bootstrap value), suggesting that this sequence encodes MR (Fig. 6). The current name for the MR gene is *nr3c2*, and this nomenclature was consistent in all teleost species included in our study.

4. Discussion

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

In our analysis, we identified several instances of supporting evidence for novel steroid receptors, including PRs in goldfish, which possess two different PRs encoded by distinct genes. Future work investigating whether other species possess additional PR genes is warranted. Several intriguing observations were also made for the ARs. For example, in the Japanese eel (*A. japonica*), both identified AR sequences clustered with AR β sequences in the “Androgen Receptor Beta” group (in contrast to Douard et al., 2008), suggesting that this species has two AR β receptors. It is unclear whether these two receptors are encoded by distinct genes, but unlike in other teleosts, no significant differences in the androgen-binding capacity of the two eel receptors has been detected (Peñaranda et al., 2014). These results suggest that the eel AR α and AR β described in the literature (Ikeuchi et al., 1999; Todo et al., 1999) are either splice variants or paralogs of the teleost AR β . It is intriguing to speculate whether these sequences represent evidence of early loss of function (e.g., non-functionalization) of the AR α receptor in this lineage. The AR sequences in rainbow trout (*O. mykiss*) also clustered with AR β sequences in the “Androgen Receptor Beta” group, which is consistent with prior work suggesting that these paralogs resulted from a lineage specific WGD that occurred in the Salmoniformes

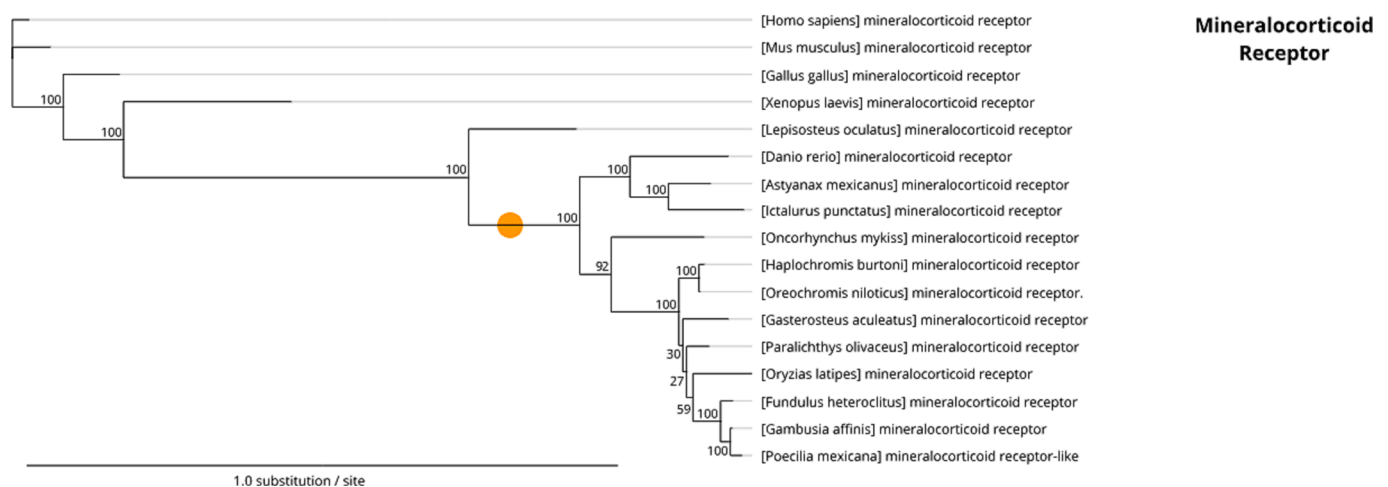


Fig. 6. Consensus phylogenetic tree of mineralocorticoid receptor. Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the approximate position of the teleost-specific whole-genome duplication event (TS-WGD). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Allendorf and Thorgaard, 1984; Takeo and Yamashita, 1999). The framework we propose can provide researchers with novel hypotheses to test, enabling them to disentangle AR evolution and selection in teleost fishes and vertebrates more broadly.

Most of the name changes we propose are straightforward, aligning inconsistencies between species with what is otherwise a consensus within the field. Proposing any universal naming system for certain proteins and genes, however, remains a challenge. For example, some teleosts, such as *D. rerio*, only have one described AR protein and one AR gene. Traditionally, animals with a single AR use the name AR (*ar*). Our analysis shows, however, that the AR protein sequence in *D. rerio* clusters with AR β sequences. Thus, this sequence should be referred to as AR β (*ar2*) based on our proposed nomenclature system in teleosts, which takes into account that the vast majority of teleost species possess two ARs (AR α and AR β). We propose that for such cases, including widely-studied models like zebrafish and goldfish, a phylogenetic approach using full-length amino acid sequences can be used to unambiguously identify that the gene in question belongs to the AR β (*ar2*) lineage of receptors. When, as is the case of this example, an extensive body of literature exists that retains the ambiguous naming convention, we suggest that explicit reference is made to the protein or gene of interest's evolutionary context. By rendering synonymous the original name AR or *ar* and the more nuanced AR β or *ar2*, we believe future research will facilitative collaboration as well as yield comparative insights within the diverse teleost clade.

5. Conclusions

Within the most popular genetic model systems, such as house mice, there is broad consensus on names of some of the most widely-studied genes that are relevant to numerous subfields of biology. This is clearly the case for steroid receptors, in which the names of proteins and genes encoding steroid receptors are consistent across the literature. For teleost fishes, where studies of steroid signaling functions are extensive, the names used for these proteins and genes have been inconsistently assigned. Our study addresses this challenge by providing a cohesive nomenclature system for steroid receptors in teleosts based on straightforward phylogenetic analyses of easily-accessed full-length amino acid sequences, an essential step for enhancing communication and promoting collaborative efforts. More broadly, we hope our analysis and recommendations will propagate a shift towards consistency in the names used for these proteins and genes, enabling us to better understand the evolution and function of steroid receptors across taxa.

Author Contributions

A.P.H. conceived of the original idea for the manuscript. B.A.A., A.P.H., and K.M.M. conceived of the approach to analyses and final framing of the ideas presented in the manuscript. A.P.H. and K.M.M. chose the species to include in the phylogenetic analyses. A.P.H. conducted the phylogenetic analyses, generated the phylogenies, and made the figures. K.M.M. collected protein sequences, made the table in the main text, and made the [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 and references and edited significant portions of all sections of subsequent drafts. All authors reviewed, edited, and approved the final versions of the manuscript and [supplementary material](#).

CRedit authorship contribution statement

Kathleen M. Munley: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing, Visualization. **Andrew P. Hoadley:** Conceptualization, Data curation, Formal

analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Beau A. Alward:** Conceptualization, Funding acquisition, Software, Supervision, Writing – original draft, Writing – review & editing.

Data availability

Data from this study are available on GitHub (<https://github.com/AlwardLab>) and are publicly accessible.

Acknowledgments

We thank members of the Alward Lab for their intellectual contributions to the ideas discussed in this paper. We are also grateful to Dr. Vincent Laudet and two anonymous reviewers for their valuable feedback on a previous version of this manuscript. This research was funded by NIGMS grant R35GM142799 (to B.A.A.), University of Houston – National Research University Fund grant R0503962 (to B.A.A.), and a Beckman Young Investigator Award (to B.A.A.).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ygcen.2023.114436>.

References

- Allendorf, F.W., Thorgaard, G.H., 1984. Tetraploidy and the Evolution of Salmonid Fishes. In: Turner, B.J. (Ed.), *Evolutionary Genetics of Fishes*, Monographs in Evolutionary Biology. Springer, US, Boston, MA, pp. 1–53. https://doi.org/10.1007/978-1-4684-4652-4_1.
- Alward, B.A., Laud, V.A., Skalnik, C.J., York, R.A., Juntii, S.A., Fernald, R.D., 2020. Modular Genetic Control of Social Status in a cichlid fish. *PNAS* 117 (45), 28167–28174.
- 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>.
- Baker, M.E., 2019. Steroid receptors and vertebrate evolution. *Mol. Cell. Endocrinol.* 496, 110526 <https://doi.org/10.1016/j.mce.2019.110526>.
- 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., 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 Crolius, 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, Z., Omori, Y., Koren, S., Shirokiya, T., Kuroda, T., Miyamoto, A., Wada, H., Fujiyama, A., Toyoda, A., Zhang, S., Wolfsberg, T.G., Kawakami, K., Phillippy, A.M., NISC Comparative Sequencing Program, Mullikin, J.C., Burgess, S.M., 2019. De novo assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication. *Sci. Adv.* 5, eaav0547. <https://doi.org/10.1126/sciadv.aav0547>.
- Chen, P., Li, B., Ou-Yang, L., 2022. Role of estrogen receptors in health and disease. *Front. Endocrinol.* 13, 839005.
- Nuclear Receptors Nomenclature Committee, "A Unified Nomenclature System for the Nuclear Receptor Superfamily." *Cell*, vol. 97, no. 2, 1999, pp. 161–63, [https://doi.org/10.1016/S0092-8674\(00\)80726-6](https://doi.org/10.1016/S0092-8674(00)80726-6).
- Conte, M.A., Gammerding, 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.
- Douard, V., Brunet, F., Boussau, B., Ahrens-Fath, I., Vlaeminck-Guillem, V., Haendler, B., Laudet, V., Guiguen, Y., 2008. The fate of the duplicated androgen receptor in fishes:

- a late neofunctionalization event? *BMC Evol. Biol.* 8, 336. <https://doi.org/10.1186/1471-2148-8-336>.
- Dube, N., Khan, S.H., Sasse, R., Okafor, C.D., 2023. Identification of an Evolutionarily Conserved Allosteric Network in Steroid Receptors. *J. Chem. Inf. Model.* 63, 571–582. <https://doi.org/10.1021/acs.jcim.2c01096>.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
- Eick, G.N., Thornton, J.W., 2011. Evolution of steroid receptors from an estrogen-sensitive ancestral receptor. *Mol. Cell. Endocrinol.* 334 (1–2), 31–38.
- Faught, E., Vijayan, M.M., 2018. The mineralocorticoid receptor is essential for stress axis regulation in zebrafish larvae. *Sci. Rep.* 8, 1–11. <https://doi.org/10.1038/s41598-018-36681-w>.
- Faught, E., Vijayan, M.M., 2019. Loss of the glucocorticoid receptor in zebrafish improves muscle glucose availability and increases growth. *Am. J. Phys. Endocrinol. Metab.* 316, E1093–E1104. <https://doi.org/10.1152/ajpendo.00045.2019>.
- Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y., Postlethwait, J., 1999. Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics* 151, 1531–1545. <https://doi.org/10.1093/genetics/151.4.1531>.
- Glasauer, S.M.K., Neuhauss, S.C.F., 2014. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol. Genet. Genomics* 289 (6), 1045–1060.
- Godwin, J., 2010. Neuroendocrinology of sexual plasticity in teleost fishes. *Frontiers in Neuroendocrinol.* 31, 203–216.
- Gronemeyer, H., Gustafsson, J.-A., Laudet, V., 2004. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discov.* 3, 950–964. <https://doi.org/10.1038/nrd1551>.
- Guindon, Stéphane, et al. “New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.” *Systematic Biology*, vol. 59, no. 3, Mar. 2010, pp. 307–21. <https://doi.org/10.1093/sysbio/syq010>.
- Hochberg, G.K.A., Liu, Y., Marklund, E.G., Metzger, B.P.H., Laganowsky, A., Thornton, J. W., 2020. A hydrophobic ratchet entrenches molecular complexes. *Nature* 588, 503–508. <https://doi.org/10.1038/s41586-020-3021-2>.
- Hoegg, S., Brinkmann, H., Taylor, J.S., Meyer, A., 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J. Mol. Evol.* 59, 190–203. <https://doi.org/10.1007/s00239-004-2613-z>.
- Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Bertelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J.C., Koch, R., Rauch, G.-J., White, S., Chow, W., Kilian, B., Quintais, L.T., Guerra-Assunção, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.-H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle, C., Elliott, D., Threadgold, G., Harden, G., Ware, D., Begum, S., Mortimore, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gildershorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., 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-Urün, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirm, A., Konantz, J., Konantz, M., Oberländer, M., Rudolph-Geiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A., Widada, 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., 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. *PNAS* 115, 6249–6254. <https://doi.org/10.1073/pnas.1719358115>.
- Ikeuchi, T., Todo, T., Kobayashi, T., Nagahama, Y., 1999. cDNA cloning of a novel androgen receptor subtype. *J. Biol. Chem.* 274, 25205–25209. <https://doi.org/10.1074/jbc.274.36.25205>.
- Jackson, L.J., Lopez, M.S., Alward, B.A., 2023. Breaking through the bottleneck: Krogh’s principle in behavioral neuroendocrinology and the potential of gene editing. *Integr. Comp. Biol.* 63 (2), 428–443.
- 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., Anthonard, 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., Quétiér, F., Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissbach, J., Crollius, H.R., 2004.
- Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431 (7011), 946–957.
- 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, 260–272. <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., 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.
- Kelley, S.T., Thackray, V.G., 1999. Phylogenetic Analyses Reveal Ancient Duplication of Estrogen Receptor Isoforms. *J. Mol. Evol.* 49 (5), 609–614. <https://doi.org/10.1007/pl00006582>.
- Lehmann, R., Lightfoot, D.J., Schunter, C., Michell, C.T., Ohyanagi, H., Mineta, K., Foret, S., Berumen, M.L., Miller, D.J., Aranda, M., Gojbori, T., Munday, P.L., Ravasi, T., 2019. Finding Nemo’s Genes: A chromosome-scale reference assembly of the genome of the orange clownfish *Amphiprion percula*. *Mol. Ecol. Resour.* 19, 570–585. <https://doi.org/10.1111/1755-0998.12939>.
- 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. *J. Comp. Neurol.* 528, 2499–2522. <https://doi.org/10.1002/cne.24908>.
- Munley, K.M., Alward, B.A., 2023. Control of social status by sex steroids: insights from teleost fishes. *Mol. Psychol. Brain, Behavior, Soc.* 2, 21.
- 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. *Curr. Biol.* 31, 1699–1710.e6. <https://doi.org/10.1016/j.cub.2021.01.089>.
- Ocampo Daza, D., Bergqvist, C.A., Larhammar, D., 2022. The evolution of oxytocin and vasotocin receptor genes in jawed vertebrates: A clear case for gene duplications through ancestral whole-genome duplications. *Front. Endocrinol.* 12, 792644.
- 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. *Trans. Am. Fish. Soc.* 99, 120–130. [https://doi.org/10.1577/1548-8659\(1970\)99<120:TEDIGS>2.0.CO;2](https://doi.org/10.1577/1548-8659(1970)99<120:TEDIGS>2.0.CO;2).
- Olender, T., Jones, T.E.M., Bruford, E., Lancet, D., 2020. A unified nomenclature for vertebrate olfactory receptors. *BMC Evol. Biol.* 20, 42.
- 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., 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.
- Peichel, C.L., McCann, S.R., Ross, J.A., Nafaty, 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.

- Peñaranda, D., Mazzeo, I., Gallego, V., Hildahl, J., Nourizadeh-Lillabadi, R., Pérez, L., Weltzien, F.-A., Asturiano, J., 2014. The Regulation of Aromatase and Androgen Receptor Expression During Gonad Development in Male and Female European Eel. *Reprod Dom Anim* 49, 512–521. <https://doi.org/10.1111/rda.12321>.
- 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.
- Shao, C., Bao, B., Xie, Z., Chen, X., Li, B., Jia, X., Yao, Q., Orti, G., Li, W., Li, X., Hamre, K., Xu, J., Wang, L., Chen, F., Tian, Y., Schreiber, A.M., Wang, N., Wei, F., Zhang, J., Dong, Z., Gao, L., Gai, J., Sakamoto, T., Mo, S., Chen, W., Shi, Q., Li, H., Xiu, Y., Li, Y., Xu, W., Shi, Z., Zhang, G., Power, D.M., Wang, Q., Schartl, M., Chen, S., 2017. The genome and transcriptome of Japanese flounder provide insights into flatfish asymmetry. *Nat. Genet.* 49, 119–124. <https://doi.org/10.1038/ng.3732>.
- 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.
- Takeo, J., Yamashita, S., 1999. Two distinct isoforms of cDNA encoding rainbow trout androgen receptors. *J. Biol. Chem.* 274, 5674–5680. <https://doi.org/10.1074/jbc.274.9.5674>.
- Theofanopoulou, C., Gedman, G., Cahill, J.A., et al., 2021. Universal nomenclature for oxytocin–vasotocin ligand and receptor families. *Nature* 592, 747–755.
- Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *PNAS* 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.
- Todo, T., Ikeuchi, T., Kobayashi, T., Nagahama, Y., 1999. Fish androgen receptor: cDNA cloning, steroid activation of transcription in transfected mammalian cells, and tissue mRNA levels. *Biochem. Biophys. Res. Commun.* 254, 378–383. <https://doi.org/10.1006/bbrc.1998.9919>.
- Tsai, P.-S., 2018. Gonadotropin-releasing hormone by any other name would smell as sweet. *Gen. Comp. Endocrinol.* 264, 58–63. <https://doi.org/10.1016/j.ygcen.2017.09.010>.
- Vollf, J.-N., 2005. Genome evolution and biodiversity in teleost fish. *Heredity* 94 (3), 280–294.
- Wada, H., 2008. Glucocorticoids: mediators of vertebrate ontogenetic transitions. *Gen. Comp. Endocrinol.* 156, 441–453. <https://doi.org/10.1016/j.ygcen.2008.02.004>.
- Whirledge, S., Cidlowski, J.A., 2019. Chapter 5 - Steroid Hormone Action. In: Strauss III, J.F., Barbieri, R.L. (Eds.), *Yen & Jaffe's Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management*. Elsevier, pp. 115–131.e114.
- Yin, Y., Tang, H., Liu, Y., Chen, Y., Li, G., Liu, X., Lin, H., 2017. Targeted disruption of aromatase reveals dual functions of cyp19a1a during sex differentiation in zebrafish. *Endocrinology* 158, 3030–3041. <https://doi.org/10.1210/en.2016-1865>.
- Yong, L., Thet, Z., Zhu, Y., 2017. Genetic editing of the androgen receptor contributes to impaired male courtship behavior in zebrafish. *J. Exp. Biol.* 220, 3017–3021. <https://doi.org/10.1242/jeb.161596>.
- Zhu, B., Ge, W., 2018. Genome editing in fishes and their applications. *Gen. Comp. Endocrinol.* 257, 3–12.
- Ziv, L., Muto, A., Schoonheim, P.J., Meijsing, S.H., Strasser, D., Ingraham, H.A., Schaaf, M.J.M., Yamamoto, K.R., Baier, H., 2013. An affective disorder in zebrafish with mutation of the glucocorticoid receptor. *Mol. Psychiatry* 18, 681–691. <https://doi.org/10.1038/mp.2012.64>.