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

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

***Correspondence:**
Beau A. Alward
balward@uh.edu

[†]These authors share first authorship.

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

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, as well as 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; 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 Lombés, 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, followed by an ancestral progesterone receptor (Thornton, 2001; Thornton et al., 2003). Specifically, extant steroid receptors evolved 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; Baker, 2019; Dube

This is a provisional file, not the final typeset article

et al., 2023; Eick and Thornton, 2011, Hochberg et al. 2020). Taken together, 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 fundamental roles of steroid hormones 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 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. Despite previous efforts to unify the nomenclature of this nuclear receptor superfamily (Nuclear Receptors Nomenclature Committee, 1999), researchers that study steroid receptors often still use inconsistent and conflicting names and symbols when referring to steroid receptors of interest. This challenge at first may seem inconsequential, but a 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), 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 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 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 for that 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 fish research during a time of rapid technological advancements in genetic technology. We hope that our approach may build upon similar efforts (e.g., Tsai et al. 2018) to establish consistent naming systems for other 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 Cyprinodontimformes (toothcarps), Order Characiformes (tetras and piranhas), Order Siluriformes (catfish), Order Salmoniformes (salmonids), Order Gadiformes (cods), Order Beloniformes (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 (damselfishes 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 the availability of protein sequences for steroid receptors, many of which were acquired via whole-genome sequencing projects, and 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; *Callorhinchus milli* (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. milli*) 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 Figure 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 analysis: ARα, ARβ, ARβ1, ARβ2, ARa, ARb, ER, ERα, ERβ, ERβ1, ERβ2, ER2a, ER2b, PR, PR1, PR2, GR, GR1, GR2, and MR. Sequences that were described as alternative splice variants or as partial sequences were excluded from analysis. Protein sequences were downloaded from NCBI Protein and UniProt in FASTA format for inclusion in phylogenetic analysis (see ***Section 2.3***). 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, available or inferred full-length amino acid sequences were used to generate a molecular phylogeny. Construction of multiple alignments of receptor protein sequences with the neighbor-joining method were performed using the MUSCLE Alignment version 3.8.425 (Edgar, 2004; Saitou and Nei, 1987). To estimate molecular phylogenies, trees were constructed using 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 we had an available sequence [*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).

This is a provisional file, not the final typeset article

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 discovered 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 as well as the model of steroid receptor gene nomenclature in mice (*M. musculus*).

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 for the receptors 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 ARα and ARβ paralogs (Fig. 2). Overall, protein sequences for AR, ARα, and ARβ 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 ARα sequences ("Androgen Receptor Alpha" group), whereas the second clade mostly consisted of ARβ 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 genes for an AR protein with an *ar* base. Thus, genes encoding ARα will be called *ar1*, and genes encoding ARβ 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 ARα sequences clustered into the "Androgen Receptor Alpha" group (100% bootstrap value), while the majority of ARβ 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 ARα 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 ARβ sequences in the "Androgen Receptor Beta" group (96% bootstrap values). Interestingly, the ARα and ARβ sequences for *O. mykiss* (rainbow trout) and *A. japonica* (Japanese eel) clustered in the "Androgen Receptor Beta" group, indicating that those sequences being referred to as ARα in these species are more similar to ARβ sequences in other teleost

species than to other ARα sequences. These 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 ERβ (Fig. 3). Protein sequences for ER, ERα, and ERβ clustered into two well-supported clades (100% bootstrap values). The first clade largely consisted of ERα sequences ("Estrogen Receptor Alpha" group), whereas the second clade primarily contained ERβ 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 ERβ2 sequences (100% bootstrap value), and 2) the "Estrogen Receptor Beta 1" group, which mainly included ERβ1 sequences (69% bootstrap value). These results are consistent with the expression of two distinct ERβ 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 ERα, ERβ1, and ERβ2 *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 ERα sequences in the "Estrogen Receptor Alpha" group (≥ 90% bootstrap values; Fig. 3). Conversely, we found some discrepancies in nomenclature based on protein sequence alignment for the "Estrogen Receptor Beta" clade. The tetrapod outgroup sequences for ERβ ["estrogen receptor beta 1" and "estrogen receptor beta 2" for *X. laevis* (African clawed frog)] clustered with *L. erinacea* (little skate), *C. milli* (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 ERβ 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 ERβ2 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 ERβ2 sequences in the "Estrogen Receptor β2" group (100% bootstrap values). Similarly, most ERβ1 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 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 ERβ1 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 relationships in this lineage (Fig. 1). Importantly, our search

This is a provisional file, not the final typeset article

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 discovered in closely related teleost species (e.g., zebrafish; Tang et al., 2016), suggesting that these sequences might represent a gene duplication that is specific to the goldfish lineage. 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 glucocorticoid receptors persisted in the teleost lineage following the TS-WGD. To remain consistent with the naming systems for the other steroid receptor proteins, we propose calling "glucocorticoid receptor 1" GRα and "glucocorticoid receptor 2" GRβ, with gene names of *nr3c1a* and *nr3c1b* (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 GRα. Because duplicate GR sequences were discovered 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 receptors

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 specious 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 in making comparative insights in the field of behavioral neuroendocrinology. 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 scheme 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 extra 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 ARβ sequences in the "Androgen Receptor Beta" group (in contrast to Douard et al., 2008), suggesting this species has two ARβ 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 ARα and ARβ known in the literature (Ikeuchi et al., 1999; Todo et al., 1999) are either splice variants or paralogs of the teleost ARβ. It is intriguing to speculate whether these sequences represent evidence of early loss of function (e.g., non-functionalization) of the ARα receptor in this lineage. The AR sequences in rainbow trout *(O. mykiss)* also clustered with ARβ 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 (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 ARβ sequences. Thus, this sequence should be referred to as ARβ (*ar2*) based on our proposed nomenclature system in teleosts, which takes into account that the vast majority of teleost species possess two ARs (ARα and ARβ). We propose that for such cases, including widely-studied models like zebrafish and goldfish, a phylogenetic approach should be used to unambiguously identify that the gene in question belongs to the ARβ (*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 ARβ 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

This is a provisional file, not the final typeset article

functions are extensive and broadly studied, 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, 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.

6. Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7. 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.

8. Funding

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.).

9. 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.

10. Supplementary Material

Table S1. Protein sequences of androgen receptors used for phylogenetic analysis.
Table S2. Protein sequences of estrogen receptors used for phylogenetic analysis.
Table S3. Protein sequences of progesterone receptors used for phylogenetic analysis.
Table S4. Protein sequences of glucocorticoid receptors used for phylogenetic analysis.
Table S5. Protein sequences of mineralocorticoid receptors used for phylogenetic analysis.

11. Data Availability Statement

422 Data from this study are available on GitHub (https://github.com/AlwardLab) and are publicly
423 accessible. In addition, a complete list of sources and accession numbers for the protein sequences
424 used in our phylogenetic analyses are available in the Supplementary Material (Tables S1-S5).
425

426 **12. References**

427 Allendorf, F.W., Thorgaard, G.H., 1984. Tetraploidy and the Evolution of Salmonid Fishes, in: Turner,
428 B.J. (Ed.), Evolutionary Genetics of Fishes, Monographs in Evolutionary Biology. Springer
429 US, Boston, MA, pp. 1–53. https://doi.org/10.1007/978-1-4684-4652-4_1

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

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

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

437 Baker, M.E., 2019. Steroid receptors and vertebrate evolution. Mol Cell Endocrinol 496, 110526.
438 https://doi.org/10.1016/j.mce.2019.110526

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

442 Betancur-R, R., Wiley, E.O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., Ortí, G.,
443 2017. Phylogenetic classification of bony fishes. BMC Evol Biol 17, 162.
444 https://doi.org/10.1186/s12862-017-0958-3.

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

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

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

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

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

This is a provisional file, not the final typeset article

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

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

463 Douard, V., Brunet, F., Boussau, B., Ahrens-Fath, I., Vlaeminck-Guillem, V., Haendler, B., Laudet,
464 V., Guiguen, Y., 2008. The fate of the duplicated androgen receptor in fishes: a late
465 neofunctionalization event? BMC Evolutionary Biology 8, 336. https://doi.org/10.1186/1471-
466 2148-8-336

467 Dube, N., Khan, S.H., Sasse, R., Okafor, C.D., 2023. Identification of an Evolutionarily Conserved
468 Allosteric Network in Steroid Receptors. J. Chem. Inf. Model. 63, 571–582.
469 https://doi.org/10.1021/acs.jcim.2c01096

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

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

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

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

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

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

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

486 Gronemeyer, H., Gustafsson, J.-A., Laudet, V., 2004. Principles for modulation of the nuclear receptor
487 superfamily. Nat Rev Drug Discov 3, 950–964. https://doi.org/10.1038/nrd1551

488 Guindon, Stéphane, et al. "New Algorithms and Methods to Estimate Maximum-Likelihood
489 Phylogenies: Assessing the Performance of PhyML 3.0." Systematic Biology, vol. 59, no. 3,
490 Mar. 2010, pp. 307–21, https://doi.org/10.1093/sysbio/syq010.

491 Hochberg, G.K.A., Liu, Y., Marklund, E.G., Metzger, B.P.H., Laganowsky, A., Thornton, J.W., 2020.
492 A hydrophobic ratchet entrenches molecular complexes. Nature 588, 503–508.
493 https://doi.org/10.1038/s41586-020-3021-2

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

497 Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E.,
498 Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C.,
499 Scott, C., Barrett, J.C., Koch, R., Rauch, G.-J., White, S., Chow, W., Kilian, B., Quintais, L.T.,
500 Guerra-Assunção, J.A., Zhou, Y., Gu, Y., Yen, J., Vogel, J.-H., Eyre, T., Redmond, S.,
501 Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S.F., Laird, G.K., Lloyd, D., Kenyon, E.,
502 Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M.,
503 Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver,
504 K., Clark, R., Riddle, C., Elliott, D., Threadgold, G., Harden, G., Ware, D., Begum, S.,
505 Mortimore, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson,
506 C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker,
507 N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H.,
508 Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K.,
509 Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C.,
510 Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray,
511 E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond,
512 S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N.,
513 Ellwood, M., Woodmansey, R., Clark, G., Cooper, J.D., Tromans, A., Grafham, D., Skuce, C.,
514 Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner,
515 P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C.M., Ersan-Ürün, Z., Eser,
516 C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberländer, M.,
517 Rudolph-Geiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A.,
518 Widaa, S., Langford, C., Yang, F., Schuster, S.C., Carter, N.P., Harrow, J., Ning, Z., Herrero,
519 J., Searle, S.M.J., Enright, A., Geisler, R., Plasterk, R.H.A., Lee, C., Westerfield, M., de Jong,
520 P.J., Zon, L.I., Postlethwait, J.H., Nüsslein-Volhard, C., Hubbard, T.J.P., Crollius, H.R.,
521 Rogers, J., Stemple, D.L., 2013. The zebrafish reference genome sequence and its relationship
522 to the human genome. Nature 496(7446), 498-503.

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

528 Ikeuchi, T., Todo, T., Kobayashi, T., Nagahama, Y., 1999. cDNA cloning of a novel androgen receptor
529 subtype. Journal of Biological Chemistry 274, 25205–25209.
530 https://doi.org/10.1074/jbc.274.36.25205

531 Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer,
532 C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C.,
533 Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N., Castellano, S., Anthouard, V.,
534 Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Biémont, C., Skalli, Z., Cattolico, L., Poulain,
535 J., De Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.-P., Gouzy, J., Parra,
536 G., Lardier, G., Chapple, C., McKernan, K.J., McEwan, P., Bosak, S., Kellis, M., Volff, J.-N.,
537 Guigó, R., Zody, M.C., Mesirov, J., Lindblad-Toh, K., Birren, B., Nusbaum, C., Kahn, D.,
538 Robinson-Rechavi, M., Laudet, V., Schachter, V., Quétier, F., Saurin, W., Scarpelli, C.,

This is a provisional file, not the final typeset article

539 Wincker, P., Lander, E.S., Weissenbach, J., Crollius, H.R., 2004. Genome duplication in the
540 teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature
541 431(7011), 946-957.

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

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

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

550 Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi,
551 K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A.,
552 Sasaki, T., Shimizu, A., Asakawa, S., N., S., Hashimoto, S., Yang, J., Lee, Y., Matsushima, K.,
553 Sugano, S., Sakaizumi, M., Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K.,
554 Morishita, T., Endo, T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y., 2007. The medaka
555 draft genome and insights into vertebrate genome evolution. Nature 447(7145), 714-719.

556 Kelley, S. T., and V. G. Thackray. "Phylogenetic Analyses Reveal Ancient Duplication of Estrogen
557 Receptor Isoforms." Journal of Molecular Evolution, vol. 49, no. 5, Nov. 1999, pp. 609–14,
558 https://doi.org/10.1007/pl00006582.

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

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

575 Lorin, T., Salzburger, W., Böhne, A., 2015. Evolutionary fate of the androgen receptor-Signaling
576 pathway in ray-finned fishes with a special focus on cichlids. G3: Genes, Genomes, Genetics
577 5, 2275–2283. https://doi.org/10.1534/g3.115.020685

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

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

582 Maruska, K.P., Butler, J.M., Anselmo, C., Tandukar, G., 2020. Distribution of aromatase in the brain
583 of the African cichlid fish Astatotilapia burtoni: Aromatase expression, but not estrogen
584 receptors, varies with female reproductive-state. Journal of Comparative Neurology 528, 2499–
585 2522. https://doi.org/10.1002/cne.24908

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

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

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

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

597 Nuclear Receptors Nomenclature Committee,. "A Unified Nomenclature System for the Nuclear
598 Receptor Superfamily." Cell, vol. 97, no. 2, 1999, pp. 161–63, https://doi.org/10.1016/S0092-
599 8674(00)80726-6.

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

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

607 Peñaranda, D., Mazzeo, I., Gallego, V., Hildahl, J., Nourizadeh-Lillabadi, R., Pérez, L., Weltzien, F.-
608 A., Asturiano, J., 2014. The Regulation of Aromatase and Androgen Receptor Expression
609 During Gonad Development in Male and Female European Eel. Reprod Dom Anim 49, 512–
610 521. https://doi.org/10.1111/rda.12321

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

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

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

This is a provisional file, not the final typeset article

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

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

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

626 Takeo, J., Yamashita, S., 1999. Two distinct isoforms of cDNA encoding rainbow trout androgen
627 receptors. J Biol Chem 274, 5674–5680. https://doi.org/10.1074/jbc.274.9.5674

628 Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by
629 ligand exploitation and serial genome expansions. Proc. Natl. Acad. Sci. USA 98(10), 5671-
630 5676.

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

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

635 Todo, T., Ikeuchi, T., Kobayashi, T., Nagahama, Y., 1999. Fish androgen receptor: cDNA cloning,
636 steroid activation of transcription in transfected mammalian cells, and tissue mRNA levels.
637 Biochem Biophys Res Commun 254, 378–383. https://doi.org/10.1006/bbrc.1998.9919

638 Tsai, P.-S., 2018. Gonadotropin-releasing hormone by any other name would smell as sweet. Gen
639 Comp Endocrinol 264, 58–63. https://doi.org/10.1016/j.ygcen.2017.09.010

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

641 Wada, H., 2008. Glucocorticoids: mediators of vertebrate ontogenetic transitions. Gen Comp
642 Endocrinol 156, 441–453. https://doi.org/10.1016/j.ygcen.2008.02.004

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

646 Yin, Y., Tang, H., Liu, Y., Chen, Y., Li, G., Liu, X., Lin, H., 2017. Targeted disruption of aromatase
647 reveals dual functions of cyp19a1a during sex differentiation in zebrafish. Endocrinology 158,
648 3030–3041. https://doi.org/10.1210/en.2016-1865

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

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

655 Zhu, B., Ge, W., 2018. Genome editing in fishes and their applications. Gen. Comp. Endocrinol. 257,
656 3-12.

657

This is a provisional file, not the final typeset article

658 **TABLES**

659

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

661

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
PR	*pgr*	
PRα	*pgr1*	Fig. 4
PRβ	*pgr2*	Fig. 4
GR	*nr3c1*	
GRα	*nr3c1a*	Fig. 5
GRβ	*nr3c1b*	Fig. 5
MR	*nr3c2*	Fig. 6

662

663 Proposed protein and gene names for androgen receptors, estrogen receptors, progesterone receptors, glucocorticoid
664 receptors, and mineralocorticoid receptors in teleost fishes based on phylogenetic analysis. The location of figures
665 containing phylogenies associated with each steroid receptor subtype are also shown. *Abbreviations: AR, androgen*
666 *receptor; ar1, androgen receptor 1 gene; ar2, androgen receptor 2 gene; ARα, androgen receptor alpha protein; ARβ,*
667 *androgen receptor beta protein; ER, estrogen receptor; ERα, estrogen receptor alpha protein; ERβ, estrogen receptor*
668 *beta protein; ERβ1, estrogen receptor beta 1 protein; ERβ2, estrogen receptor beta 2 protein; esr1, estrogen receptor 1*
669 *gene; esr2a, estrogen receptor 2a gene; esr2b, estrogen receptor 2b gene; GR, glucocorticoid receptor protein; GRα,*
670 *glucocorticoid receptor alpha protein; GRβ, glucocorticoid receptor beta protein; MR, mineralocorticoid receptor;*
671 *nr3c1, nuclear receptor subfamily 3 group C member 1 gene; nr3c1a, nuclear receptor subfamily 3 group C member 1a*
672 *gene; nr3c1b, nuclear receptor subfamily 3 group C member 1b gene; nr3c2, nuclear receptor subfamily 3 group C*
673 *member 2 gene; pgr, progesterone receptor gene; pgr1, progesterone receptor 1 gene; pgr2, progesterone receptor 2*
674 *gene; PR, progesterone receptor; PRα, progesterone receptor alpha protein; PRβ, progesterone receptor beta protein.*

675

676

677

678

679

680

FIGURE LEGENDS

Figure 1: Phylogeny identifying the evolutionary relationships between taxa considered in this paper [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.*

Figure 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).

Figure 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 ERβ, but not ERα.

Figure 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).

Figure 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).

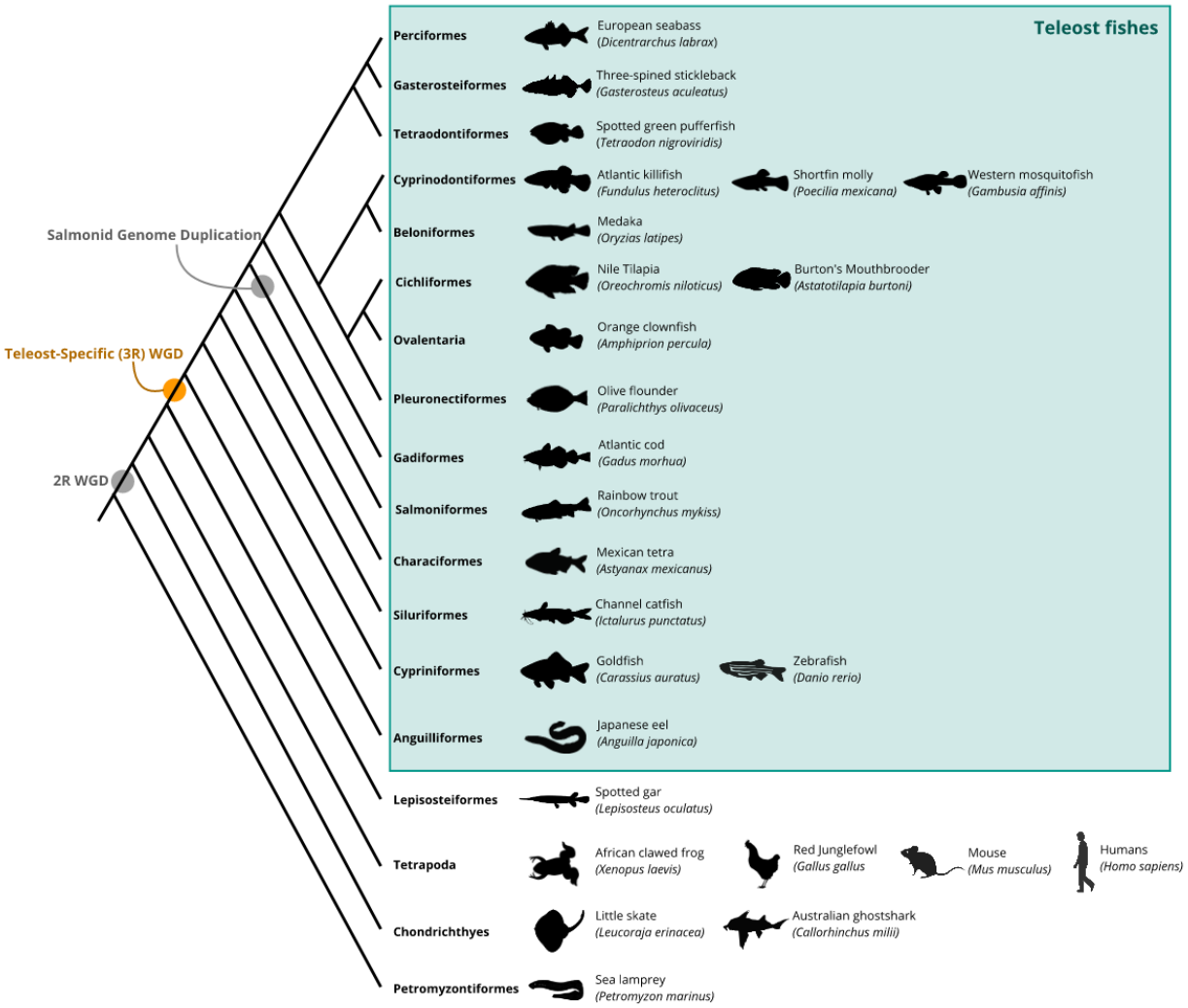
Figure 6: Consensus phylogenetic tree of mineralocorticoid receptor. Numbers shown at each branch node indicate bootstrap values (%), and the scale bar indicates phylogenetic distance (0.1 substitutions/site). The orange circle represents the approximate position of the teleost-specific whole-genome duplication event (TS-WGD).

This is a provisional file, not the final typeset article

714 **FIGURES:**

715

716 **Figure 1.**



717

718

719 **Figure 2.**

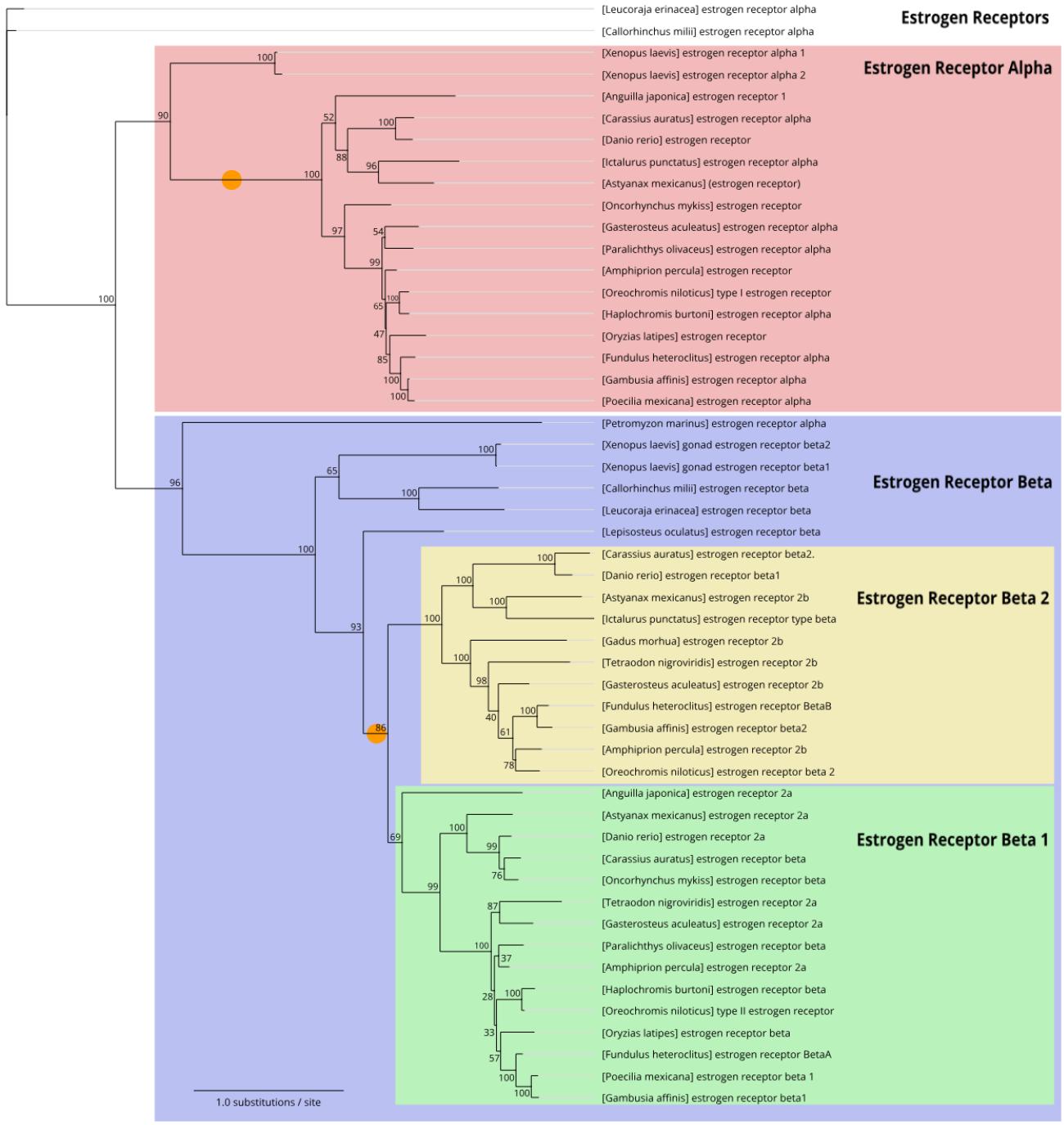
720

721

722

This is a provisional file, not the final typeset article

723 **Figure 3.**



724

725

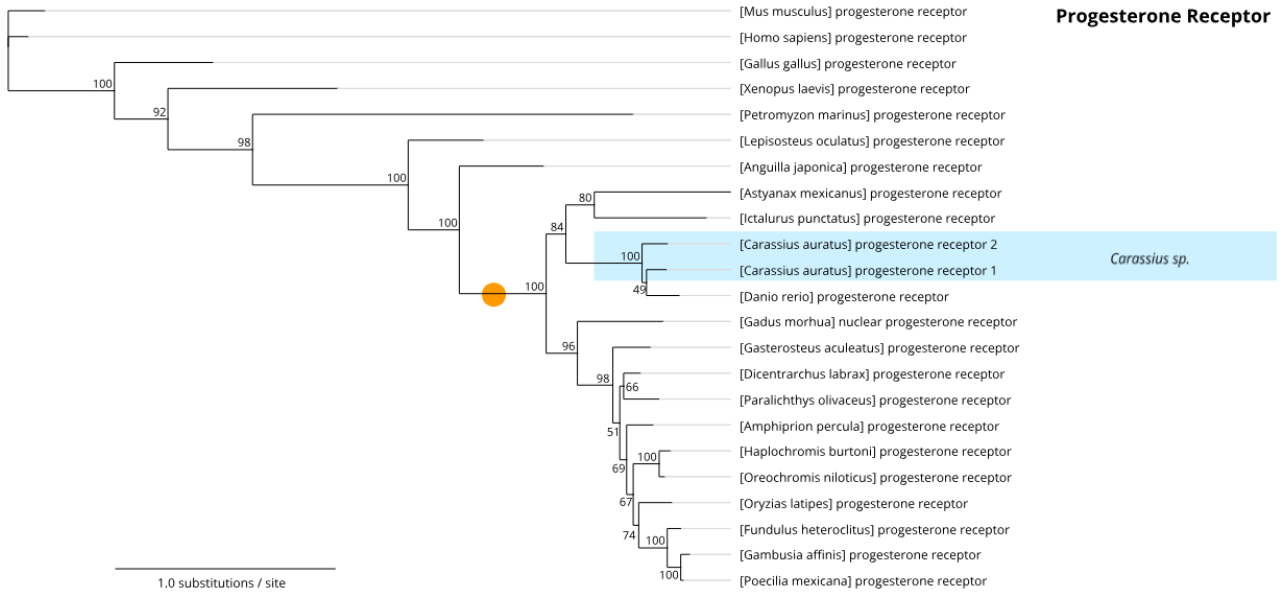
726

727

728

729

730 **Figure 4.**



731

732

733

734

735

736

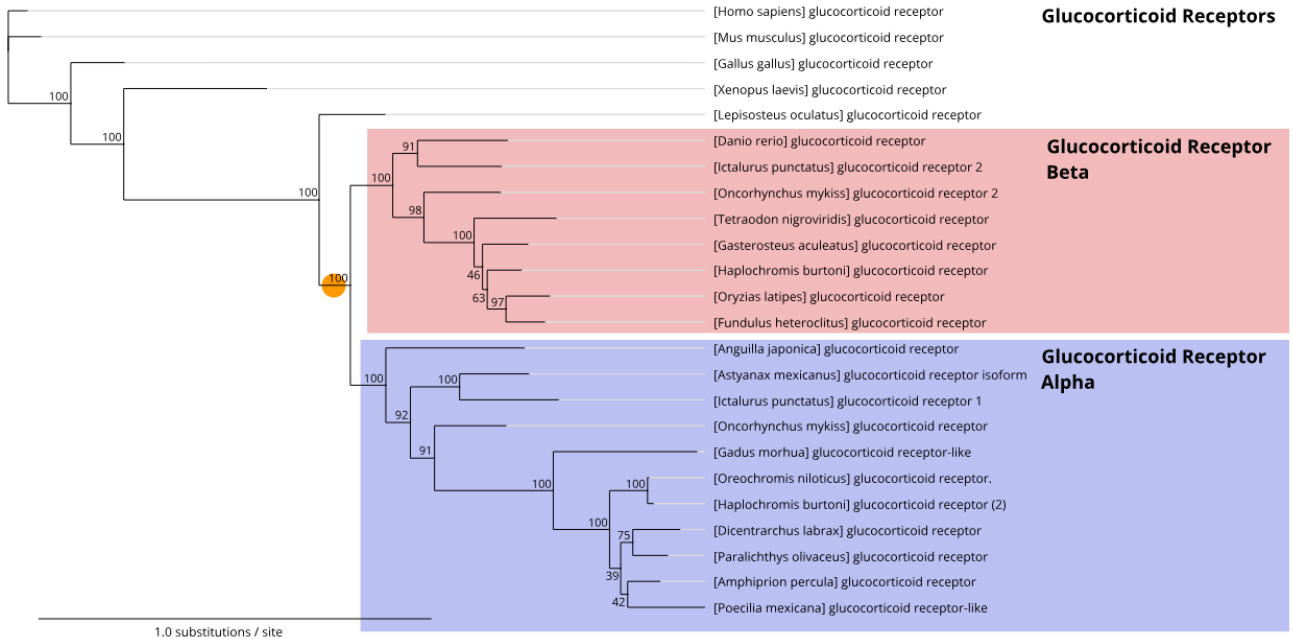
737

738

739

This is a provisional file, not the final typeset article

740 **Figure 5.**



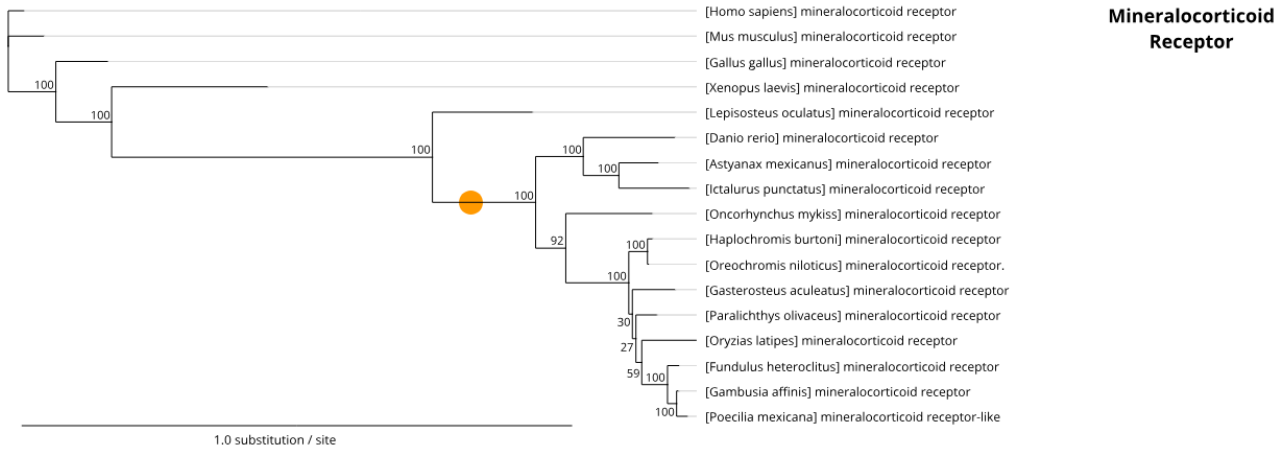
741

742

743

744

745 **Figure 6.**



746

747

This is a provisional file, not the final typeset article