Gene 843 (2022) 146795

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Phylogenomic and expression analysis of *Colossoma macropomum cyp19a1a* and *cyp19a1b* and their non-classical role in tambaqui sex differentiation

R.V. Paixão [a], G.F. Silva [b], A.R. Caetano [c], L.C. Cintra [d], E.S. Varela [e], F.L.A. O'Sullivan [b,*]

[a] Universidade Federal do Amazonas (UFAM), Programa de Pós-graduação em Ciência Animal e Recursos Pesqueiros, Avenida Rodrigo Otávio, CEP: 69080-900, 6200 Manaus, AM, Brazil
[b] Embrapa Amazônia Ocidental, Rodovia AM-010, Km 29, Caixa Postal 319, CEP: 69010-790, Brazil
[c] Embrapa Recursos Genéticos e Biotecnologia, Final Av. W/5 Norte, C.P. 02372, CEP 70770-917, Brasília, DF, Brazil
[d] Embrapa Agricultura Digital, Avenida André Tosselo, 209, Cidade Universitária, CEP: 13083-886, Campinas, SP, Brazil
[e] Embrapa Pesca e Aquicultura, Av. NS 10, cruzamento com a Av. LO 18 Sentido Norte Loteamento - Água Fria, Palmas, TO 77008-900, Brazil

ARTICLE INFO

Edited by X. Carette

Keywords:
Cyp19a1
Synteny
Promoter Structure
Sex development
Neotropical fish
aromatase

ABSTRACT

The genes coding for Cytochrome P450 aromatase (*cyp19a1a* and *cyp19a1b)* and estrogen (E_2) receptors (*esr1*, *esr2a* and *esr2b*) play a conserved role in ovarian differentiation and development among teleosts. Classically, the "gonad form" of aromatase, coded by the *cyp19a1a*, is responsible for the ovarian differentiation in genetic females via ligation and activation of the Esr, which mediates the endocrine and exocrine signaling to allow or block the establishment of the feminine phenotype. However, in neotropical species, studies on the molecular and endocrine processes involved in gonad differentiation as well as on the effects of sex modulators are recent and scarce. In this study, we combined *in silico* analysis, real-time quantitative PCR (qPCR) assay and quantification of E_2 plasma levels of differentiating tambaqui (*Colossoma macropomum)* to unveil the roles of the paralogs *cypa19a1a* and *cyp19a1b* during sex differentiation. Although the synteny of each gene is very conserved among characids, the genomic environment displays striking differences in comparison to model teleost species, with many rearrangements in *cyp19a1a* and *cyp19a1b* adjacencies and transposable element traces in both regulatory regions. The high dissimilarity (DI) of SF-1 binding motifs in *cyp19a1a* (DI = 10.06 to 14.90 %) and *cyp19a1b* (DI = 8.41 to 13.50 %) regulatory region, respectively, may reflect in an alternative pathway in tambaqui. Indeed, while low transcription of *cyp19a1a* was detected prior to sex differentiation, the expression of *cyp19a1b* and *esr2a* presented a large variation at this phase, which could be associated with sex-specific differential expression. Histological analysis revealed that anti-estradiol treatments did not affect gonadal sex ratios, although Fadrozole (50 mg kg^{-1} of food) reduced E_2 plasma levels ($p < 0{,}005$) as well *cyp19a1a* transcription; and tamoxifen (200 mg kg^{-1} of food) down regulated both *cyp19a1a* and *cyp19a1b* but did not influence E_2 levels. Altogether, our results bring into light new insights about the evolutionary fate of *cyp19a1* paralogs in neotropical fish, which may have generated uncommon roles for the gonadal and brain forms of *cyp19a1* genes and the unexpected lack of effect of endocrine disruptors on tambaqui sexual differentiation.

Abbreviations: aa, amino acid(s); AhR, Aryl hydrocarbon receptor; AIs, Aromatase inhibitors; AR, Androgen receptor; bp, base pair(s); cDNA, DNA complementary to RNA; CDS, Coding DNA sequence; CRE, cAMP response element; CRX, Cone-rod homeobox protein; dNTP, deoxyribonucleoside triphosphat; DBP, D site-binding protein binding site; EDCs, Endocrine Disrupting Chemicals; ELISA, enzyme-linked immunosorbent assay; ERE, Estrogen-response element; ERs, Estrogen receptors; ETOH, ethyl alcohol; FRA-1, Fos related antigen-1; Foxl2, forkhead box L2; GATA-1, GATA binding factor 1; hERE, ERE half sites; LTR, long terminal repeat(s); IRF-2, Interferon regulatory factor 2; Myr, million years; NF-Y, Nuclear transcription factor Y; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; P53, Cellular tumor antigen p53; PR, Progesterone receptor; SERMs, Selective Estrogen Receptor Modulators; SF-1, Steroidogenic factor 1; SMAD3, Mothers against decapentaplegic homolog 3; SOX9, SRY-Box Transcription Factor 9; TBP, Transcription binding protein; TCF-4E, Transcription factor 4; TEs, Transposable Elements; TFBSs, transcription factor binding sites; TIR, Terminal inverted repeats; UTR, untranslated region(s); WT1-KS, Wilms tumor protein.

* Corresponding author at: Rodovia AM-010, Km 29, Caixa Postal 319, CEP: 69010-970, Manaus, Amazonas, Brazil.
E-mail address: fernanda.almeida@embrapa.br (F.L.A. O'Sullivan).

https://doi.org/10.1016/j.gene.2022.146795
Received 8 February 2022; Received in revised form 15 July 2022; Accepted 5 August 2022
Available online 9 August 2022

1. Introduction

In teleosts, steroids are well known for playing key roles in sex differentiation and reproduction. A general and classic feature of ovarian development in teleost fish is the sex-biased synthesis of estradiol during the critical period of ovary differentiation, as a reflection of the sex-specific differential expression of the *cyp19a1*. The *cyp19a1* encodes for Cytochrome P450 aromatase, the catalytic enzyme that aromatizes androgens into estrogens (Devlin & Nagahama, 2002; Guiguen et al., 2010; Rajakumar & Senthilkumaran, 2020). Furthermore, estrogens, mainly 17β-estradiol, act as primarily transcriptional regulators of target genes mediated by the nuclear estrogen receptors (ERs), which recognize and activate gene transcription through binding to the genomic element called the estrogen-response element (ERE) (Amenyogbe et al., 2020). The administration of endocrine disruptor chemicals (EDCs) such as aromatase inhibitors (AIs) and Selective Estrogen Receptor Modulators (SERMs), which block this E_2-ER path, induces sex reversal in fish, resulting in genetic females that are functional males (Dang & Kienzler, 2019).

Due to a second, fish-specific, whole genome duplication (3R) event that occurred during the evolution of teleosts, the *cyp19a1* gene was duplicated, giving rise to the paralogs *cyp19a1a*, mostly expressed in the gonads, and *cyp19a1b*, which is mostly expressed in the brain (Guiguen et al., 2010; Diotel et al., 2010). Additionally, although both Cyp19a1 proteins have similar enzymatic activity, some divergences in their gene promoter sequences have supported their differential tissue expression and regulation in different fish species (Lin et al., 2020). Recent studies on teleosts using different gene editing methodologies (TALEN and CRISPR) showed that the knockout of *cyp19a1a* results in all-male offspring in zebrafish (Lau et al. 2016) and tilapia (Zhang et al., 2017). However, *cyp19a1a* deficiency in XX medaka mutants did not affect the ovarian development, but caused a partial ovary degeneration and partial sexual inversion after puberty due to a failure in yolk accumulation (Nakamoto et al., 2018). This establishes that not in all cases the initial expression of *cyp19a1a* is important for the gonadal fate, but it is important for the normal ovary development. In contrast, *cyp19a1b*-null mutants showed a normal sex ratio in zebrafish and medaka, providing evidence that *cyp19a1b,* is not crucial for ovarian fate determination in these species (Yin et al., 2017, Nakamoto et al., 2018).

While the process of sex differentiation and the role of estrogen and AIs in the gonad final phenotype are well characterized in model and some economically important fish species, the elucidation of sex determination mechanisms and associated molecular pathways for sex differentiation in neotropical fishes remain key areas of research (Fernandino & Hattori, 2019). The Characiformes order is the most speciose and diverse group of fish in the neotropics, where they inhabit a large variety of freshwater habitats. Moreover, neotropical fish are the basis of native finfish production in Brazil, which recently increased steadily along with the continuous growth of the aquaculture industry in South America. Among different farmed neotropical species, the tambaqui (*Colossoma macropomum*), a member of the family Serrasalmidae, outstands as the major native species in fish farming in Brazil and in neighboring countries (Peixe, 2022). Juvenile female tambaqui apparently proceeds throughout ovarian cavity formation, which marks the ovary differentiation, in the absence of *cyp19a1a.* Moreover, RNA seq data of undifferentiated tambaqui showed that the expression of *cyp19a1b* and *esr2a* were higher in the trunks of putative males (Lobo et al., 2020). To better explore this unconventional pattern, we performed a phylogenomic analysis of both *cyp19a1* paralogs and of their coded aromatases, comparing with other teleost's and evaluated the expression of *cyp19a1a*, *cyp19a1b* and *esr2a* in different stages of tambaqui development, under normal growing conditions and treated with AI compounds before and during sex differentiation.

2. Materials and methods

2.1. In silico analysis

2.1.1. Identification of cyp19a1a and cyp19a1b sequences in tambaqui and other teleost

The identification of tambaqui *cyp19a1a* and *cyp19a1b* was done by blast searches, using *Pygocentrus nattereri* CDS sequences as query, against juvenile trunks, ovary and testis RNAseq data, and on a local assembly of the *C. macropomun* genome (*unpublished*) in addition to the *C. macropomum* genome assembly at National Center for Biotechnology Information (NCBI; Bioproject PRJEB40318). The introns prediction, splicing junctions and protein deduction from the genomic databases were performed with Augustus or manually by BLASTX tool with identification of nucleotides consensus for introns 5′(GT) and 3′(AG). The retrieved sequences were manually curated using the Unipro UGENE v. 35.1 software to construct the predicted CDS, which were subsequently used to design species-specific primers for real-time quantitative PCR (qPCR) and to deduce the amino acid sequences.

We further examined the *cyp19a1* sequences from several animal species with available sequenced genomes at NCBI and Ensembl databases, using the TBLASTN algorithm. Teleost representatives include the superorder Elopomorpha (*Megalops cyprinoides*, *Anguilla anguilla*), Osteoglossomorpha (*Sclerophages formosus*, *Arapaima gigas*, *Heterotis niloticus*, *Gymnarchus niloticus*, *Mormyrus iriodes*, *Paramormyrops kingsleyae*, *Brevimyrus niger*, *Mormyrus lacerda*), Otocephala (*Denticeps clupeoides*, *Coilia nasus*, *Limnothrissa miodon*, *Alosa alosa*, *Clupea harengus*, *Chanos chanos*, *Danio rerio*, *Carassius auratus*, *Cyprinus carpio*, *Electrophorus electricus*, *Ictalurus punctatus*, *Tachysurus fulvidraco*, *Pangasianodon hypophthalmus*, *Ameiurus melas*, *Hepsetus odoe*, *Pygocentrus nattereri*, *Astyanax mexicanus*) and Eutelostei (*Gasterosteus aculeatus*, *Takifugu rubripes*, *Dicentrarchus labrax*, *Monopterus albus*, *Esox lucius*, *Oreochromis niloticus*, *Oryzias latipes*, *Gadus morhua*, *Salmo salar*, *Onchorhynchus mykiss*). We also included non-teleost actinopterygian such as *Lepisosteus oculatus*, *Acipenser ruthenus* and *Erpetoichthys calabaricus*, some tetrapods (Sarcopterygii; *Latimeria chalumnae*, *Homo sapiens*, *Gallus gallus*, *Xenopus tropicalis*, *Anolis cariolinensis*), Chondrichthyes (*Callorhinchus milii*, *Amblyraja radiata*, *Rinchodon typus*) and one Cyclostomata representative (*Eptatretus burgeri*). Cyp19 sequences found in *Branchiostoma floridae* (Cephalochordata) were used as outgroup. Annotations were added to sequences from species (9) with non-annotated databases.

2.1.2. Promoter analysis of tambaqui cyp19a1a and cyp19a1b

The comparative intergenic analysis of both *cyp19a1* genes was carried out on the entire region from the 5′ flanking region to their respective adjacent gene neighbors. We used the fishTEDB and NCBI as databases to search for transposable elements through homology search by BLASTn and BLASTx tools, respectively. We further searched for putative transcription factor-binding sites in the promoter by applying the PROMO web tool (https://alggen.lsi. upc.es/home.html), using previous reports that have been validated TFBS from aromatase promoter region (Tanaka et al., 1995, Tchoudakova et al., 2001, Chang et al., 2005, Sridevi et al., 2012, Lin et al., 2020).

2.1.3. Phylogenetic analysis of Cyp19a1a and Cyp19a1b

Multiple sequence alignments were performed with MUSCLE, included in MEGA7. Phylogenetic relationships were estimated using maximum likelihood and Bayesian approaches. The best-fitting model for amino acid substitution matrix (JTT + I + G4) was selected based on the proposed model tool from IQ-Tree 2.0, which was also used to perform the Maximum likelihood analysis to obtain the best tree. Node support was assessed with 10,000 bootstrap pseudo replicates using the ultrafast routine. Bayesian searches were conducted in MrBayes v.3.1.248, with two independent runs of six simultaneous chains for 1,000,000 generations, and every 1000 generations were sampled using

default priors.

2.1.4. Synteny analysis of cyp19a1a and cyp19a1b loci among teleosts

The Genomicus vertebrate server (https://www.dyogen.ens.fr/genomicus), which is synchronized with Ensembl releases, was used for chromosomal localization of aromatase genes and neighboring genes that appear in conserved position between the phylogenetic representative fish species for synteny analysis. The genomic region (44 neighboring genes) of *cyp19a1a* and *cyp19a1b* from *C. macropomum* were identified and manually compared using the NCBI scaffold annotation from other representative species.

2.2. In vivo experiments

2.2.1. Ethics statement

The authors assert that all procedures of this work comply with the Ethical Principles of Animal Experimentation, adopted by the Brazilian College of Animal Experimentation (COBEA), and were approved by the local Ethics Committee on Animal Use (CEUA) – Embrapa Amazônia Ocidental (n° 05/2018, SEI 21158.003669/2018-77). The project has an Authorization of Access to Genetic Heritage under register A5784B5.

2.2.2. Animals and rearing conditions

Recently-hatched tambaqui larvae were obtained from a commercial farm (from Presidente Figueiredo, Brazil) and transported to the Embrapa Amazônia Ocidental facilities, in Manaus. Fish were maintained in acclimatization for 7 days, fed crumbled fish pellets with 45 % crude protein (CP), *ad libitum*, three times/day. After this period, juveniles were randomly distributed into nine polyethylene tanks (310 L), in a density of 200 juveniles/tank for receiving the treatments in triplicates. The temperature, oxygen, and pH of the water were measured daily. Treatments lasted 6 weeks during which the feed (experimental diets) was provided *ad libitum* four times a day.

2.2.3. Experimental diets

Fadrozole at dose of 50 mg kg^{-1} and tamoxifen at dose of 200 mg kg^{-1} were dissolved in 70 % ethanol and incorporated into fish pellets (according to Hines & Watts, 1995) for feeding sexually undifferentiated tambaqui (2 cm of length) for 40 days, which covered the period when the first morphological signs of sex differentiation appears (Lobo et al., 2020). The control diet was treated with ethanol 70 %, in the same volume: pellet weight as treatment diets. Hormone dosages were established based on different studies showing effects of tamoxifen and fadrozole on fish sex ratio (Afonso et al., 2001; Navarro-Martín et al., 2009; Singh et al., 2012). All feed required for the entire treatment was prepared in one batch. After the treatment, each fish group was transferred to and maintained in net cages up to 12 months of age, with the exception of one replicate of the tamoxifen treatment, which was lost due to an accident. All cages were placed in the same earth pond. During this phase, the animals were fed fish pellets of 32 % CP at a rate of 5 % of the tank biomass, twice a day.

2.2.4. Samplings

The first sampling was done just before the start of the treatment (only in the control tanks; T_0, n = 5/tank), 20 and 40 days after the beginning of treatment (at the middle and end of the treatment, respectively; n = 9/tank). Fish were desensitized with benzocaine and sacrificed by medullary sectioning. The trunks were snap frozen in liquid nitrogen, and stored at −80 °C. At the last sampling, blood was collected from the caudal vein of each fish prior to killing. The plasma was obtained by centrifugation, frozen and stored at −80 °C until analysis. The gonads were dissected, fixed in the Bouin solution for 24 hrs, and kept in 70 % ETOH for histological processing.

To characterize *cyp19a1a*, *cyp19a1b* and *esr* expression levels in non-treated tambaqui, 5 to 15 fish were collected at hatching, and in the sizes of 1.5, 2, 3, 4 and 6 cm of total length, covering the period of pre-differentiation, and morphological differentiation of the gonads (Lobo et al., 2020). Additionally, we collected gonads and brains of 6 sexually immature adult tambaqui (3 males and 3 females), and 8 gonads from sexually maturing tambaqui (5 females and 3 males) for RNA isolation and qPCR analysis (Supplementary Fig. 1C to F).

2.2.5. Gene expression assays

Total RNA from juveniles' trunks, brains and extracted gonads was extracted using mortar and pestle for tissue maceration with TRIzol reagent (Life Technologies; Carsbald, USA) in a proportion of 1 ml per 100 mg of tissue. Total RNA samples were treated with RNAse-free DNA (RQ1 RNAse-free; Promega, Madison, USA) to remove any possible genomic DNA residues. The concentration and integrity of the RNA were assessed by spectrophotometry (Nanodrop 1000; Thermo Scientific) and in agarose gel electrophoresis (1.5 %), respectively. Only samples with a 260/280 ratio between 1.8 and 2.1 were used for cDNA synthesis, which was performed using High-Capacity cDNA Reverse Transcription Kit (Thermo Scientific, Waltham, MA, USA), according to the manufacturer's protocol. The cDNAs were used as a template to amplify and quantify the transcript levels of *cyp19a1a, cyp19a1b* and *esr2a*. All qPCR primers (Supplementary Table 1) were designed using the Integrated DNA Technologies (IDT) tools (https://www.idtdna.com), based on tambaqui genomic nucleotide sequences. Amplification efficiency for each primer set was calculated from a 1:4 serial dilution of a pool of tambaqui ovaries and testes cDNA. Melt curves were inspected to ensure amplification of single amplicons, which were sequenced afterwards to confirm the primers specificity of each assay. The qPCR reactions were performed in the 7500 Fast Real-Time PCR System v2.3 (Applied Biosystems), in duplicates, using 2 μL of cDNA (120 ng), 1 μL of each primer (200 nM), 12,5 μL Fast SYBR Green PCR Master Mix (Applied Biosystems), and nuclease-free water to a final volume of 25 μL. qPCR data was analyzed using the $2^{-(\Delta\Delta Ct)}$ method where the expression of the target gene was first normalized to the *β-actin* (Nascimento et al., 2016) and then calibrated to the mean of the pool of larvae (first sampling) for the non-treated fish (i.e. differentiating juveniles and gonads of adult tambaqui) and of the control group in the treated fish (experimental groups).

2.2.6. Quantification of E_2 plasmatic concentration and histology

The circulating concentration of 17βestradiol was estimated by an enzyme-linked immunosorbent assay (ELISA) kit (Invitrogen, Camarillo, USA) and the samples were analyzed in duplicates. Absorbance was measured using a microplate reader (Multiskan FC Thermo Scientific®, Leicestershire, UK). Analyses were carried out following the manufacturer's instructions, and a standard curve was run for each ELISA plate. No dilution was needed in any assay.

For sex identification, the fixed gonads were dehydrated in increasing concentrations of ETOH, cleared in xylene and embedded in paraffin. The sections were cut in 3 μm and stained with haematoxylin-eosin following conventional histological procedures. Histological slides were used to identify the phenotypic sex of each fish, based on gonadal structures, such as ovarian lamellae and/or the presence of oogonial nests in females and the presence of spermatogonia and/or spermatogenic cysts in males (Supplementary Fig. 1 A and B).

2.2.7. Statistical analysis of the data

To calculate the variation in sex ratios between treatment groups vs control, Chi-square (X^2) tests were performed. Based on ΔΔCt values, the differences of gene expression among different time points of tambaqui early development as well as the differences between control and treatments in some time points were tested by Kruskal-wallis test followed by post hoc Dunn's multiple comparison test, since that did not confirm one or all of the assumptions for parametric test (normal distribution, homogeneity of variances and homoscedasticity). One-way ANOVA was used to identify significant differences of gene expression and E_2 levels between control and treated fish. All statistical analyses

were performed in Graph Pad Prism v. 8.4.3.

3. Results

3.1. Tambaqui cyp19a genes

In the tambaqui genome, the *cyp19a1a* (LOC118804009) is located in the scaffold NW_023496020.1, while *cyp19a1b* (LOC118811907) is

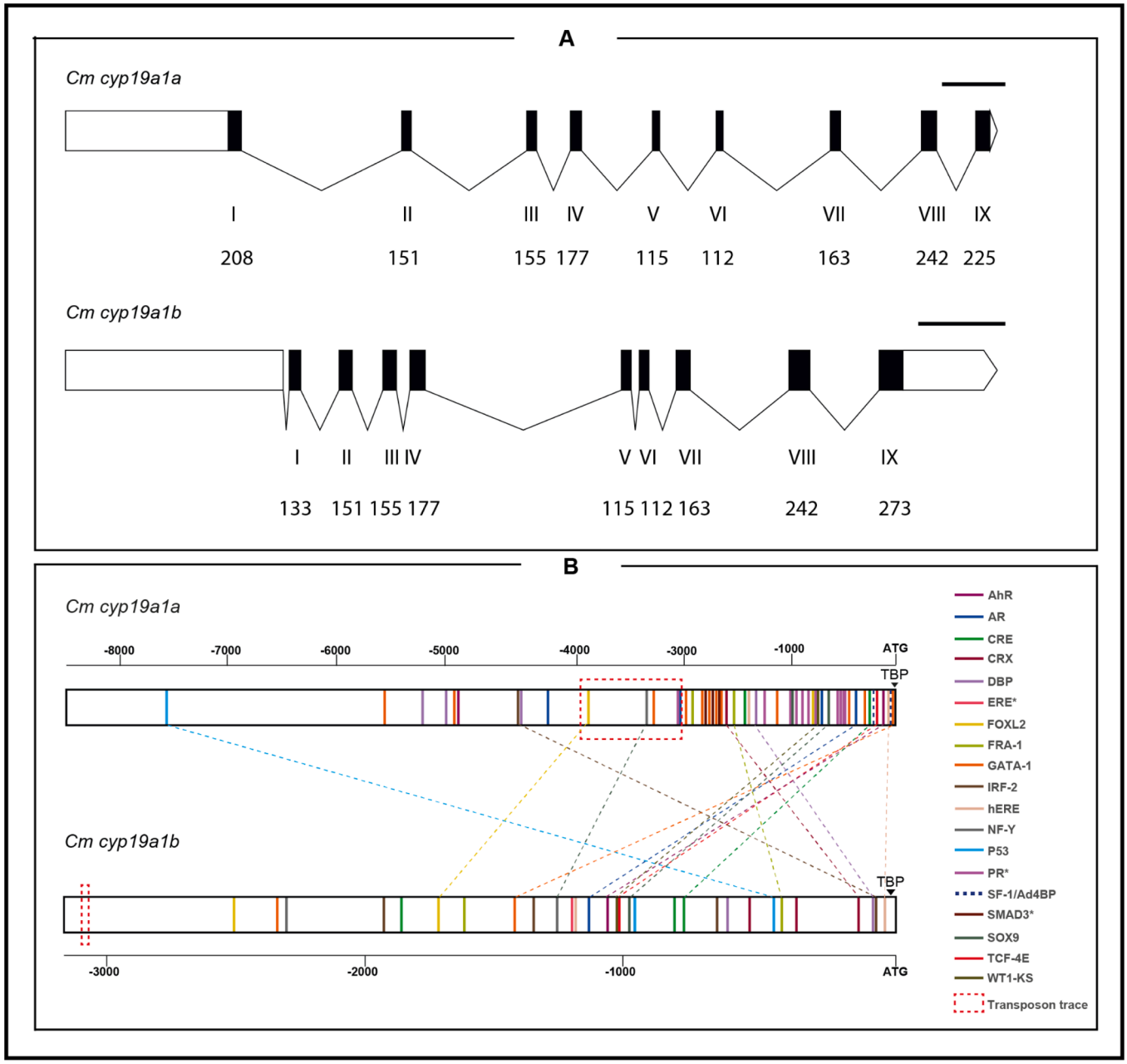

Fig. 1. Characterization of the genomic sequences of *Colossoma macropomum cyp19a1* paralogues. A) Illustrative scheme of relative positions of introns and exons of *C. macropomum cyp19a1a* and *cyp19a1b*. The exons are boxed and the introns are depicted by lines. Position and length of exons are indicated by roman numerals and cardinal numbers, respectively. The thin line shown above the graphic is a scale bar with 1000 base pairs; B) cis-regulatory elements distribution in the promoter region of *C. macropomum cyp19a1a* and *cyp19a1b*. TATA-box are indicated by arrowheads. Binding sites that were predicted to be<5 % dissimilar to the matrix (DI < 5 %) are represented by solid colored lines, with the exception of the SF-1 site which is represented by dotted line (DI > 10 %). Elements that are shared by both *cyp19a1* copies are linked by dotted lines with its respective colour, connecting the closest elements from the ATG (start codon). Asterisks indicate elements that are exclusively present in one of the duplicates. Transposon traces are indicated by red dotted squares. AhR: Aryl hydrocarbon receptor responsive element; AR, androgen receptor binding site; CRE, cAMP response element; CRX, Cone-rod homeobox protein binding site; DBP, D site-binding protein binding site; ERE, estrogen response element; Foxl2, forkhead box L2 binding site; FRA-1, Fos related antigen-1 binding site; GATA-1, GATA binding factor 1 responsive element; IRF-2, Interferon regulatory factor 2 binding site; NF-Y, Nuclear transcription factor Y binding site; SF-1/Ad4BP, steroidogenic factor 1/Ad4-binding protein; hERE, ERE half sites; P53, Cellular tumor antigen p53 binding site; PR, Progesterone receptor binding site; SMAD3, Mothers against decapentaplegic homolog 3 binding site; SOX9, SRY-Box Transcription Factor 9 binding site; TCF-4E, Transcription factor 4 binding site; WT1-KS, Wilms tumor protein binding site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

located in the scaffold NW_023494792.1. Both genes are organized into nine exons, with 1548 and 1521 bp of CDS sequence, respectively (Id shared = 72 %), resulting in a predicted protein of 515 and 506 amino acids, respectively (Id shared = 66,7%) (Fig. 1a). Cyp19a1a predicted aa sequence shares, respectively, 94, 80 and 75 % of identity with *Pygocentrus nattereri*, *Astyanax mexicanus* and *Danio rerio*, whereas Cyp19a1b have the highest amino acid identity with *P. nattereri* (Id = 95 %), followed by *D. rerio* (Id = 77,5%) and then by *A. mexicanus* (Id = 72,8%).

3.2. Intergenic and promoter analysis of tambaqui cyp19a1 genes

Tambaqui *cyp19a1a* and *cyp19a1b* display striking differences in 5′intergenic regions, with 8358 bp and 3248 bp, respectively. The tambaqui *cyp19a1a* promoter region from −3004 to −3993 bp had high similarity (79.33 %) to the PM_Contig_11 sequence of the *Periophthalmus magnuspinnatus,* classified as class II transposon, terminal inverted repeats (TIR) order and TC1-Mariner superfamily in FishTEDB database, and 41.08 % of identity with TCB1 transposase from *Polypterus senegalus* in the NCBI database. Similarly, tambaqui *cyp19a1b* 5′ intergenic region also contains a small trace from −3155 to −3194, which is 99 % similar to the TC1-Mariner from *Anguilla japonica* (Supplementary Fig. 2).

Sequence analysis upstream of the ATG start codon of tambaqui *cyp19a1a* and *cyp19a1b* revealed a TBP motif as well as several potential regulatory elements (Fig. 1b; Supplementary Fig. 3). We found putative binding sites for nuclear transcription factor Y (NF-Y), GATA-1, Nkx2-1, WT1-KS, cAMP response element-binding protein (CRE), SOX9, Forkhead box L2 (FOXL2), aryl hydrocarbon receptor/nuclear transfer responsive element (AhR/Arnt), androgen receptor (AR) and estrogen receptor recognition half sites (hERE), in both *cyp19a1a* and *cyp19a1b* 5′ flanking region. Most of these predicted sites displayed next to 0 % of dissimilarity index (DI) to the reference matrix, except for AhR and NF-Y binding sites which displayed up to 1 % and 2 % of DI, respectively. Moreover, these binding sites have been described in the promoter regions of both *cyp19a1a* and *cyp19a1b* from other teleost species (Tanaka et al., 1995, Tchoudakova et al., 2001, Chang et al., 2005, Sridevi et al., 2012, Lin et al., 2020).

Tambaqui *cyp19a1a* and *cyp19a1b* promoter regions share many putative transcription factor binding sites, but there are differences in position and number of copies that may reflect the particularities in their individual regulation. Among several GATA-1 putative binding sites predicted in tambaqui *cyp19a1a* 5′ flanking region (n = 29), four are located proximal to the transcription start site (nt −54, −285, −500, −1112) in contrast *cyp19a1b* that exhibited only five putative GATA-1 binding sites (nt −1428, −2103, −2355, −2363, −2918). Additionally, we detected four putative AR binding sites in the tambaqui *cyp19a1a* promoter region (nt −458, −747, −2062, −3308) and only one in *cyp19a1b* (nt −1138).

A full ERE binding site (GGGTCAGCCTGACCT, nt −1204) was manually identified in the tambaqui *cyp19a1b* but not in the *cyp19a1a* promoter region, which is consistent with previous studies in fish (Tanaka et al., 1995, Tchoudakova et al., 2001, Chang et al., 2005, Sridevi et al., 2012, Lin et al., 2020). Moreover, no SF-1 response element was predicted in the tambaqui *cyp19a1a* promoter region with less than 10 % of the DI. Otherwise, two SF-1 responsive elements were predicted in the tambaqui *cyp19a1b* promoter region with 10.36 % (nt −777) and 8.41 % (nt −3209) of DI.

Among the less discussed transcription factors in fish, we predicted 8 putative binding sites for progesterone receptor A (PR-A), 2 for progesterone receptor B (PR-B) and 3 for SMAD3, all exclusively present in tambaqui *cyp19a1a* promoter region with 0 % of DI. Moreover, the tambaqui *cyp19a1b* promoter region contains putative binding sites for IRF-2 (nt −79, −700, −1355, −1940, −2262), DBP (nt −91, −653), CRX (nt −147,-389, −572), FRA-1 (nt −447, −1626) and P53 (nt −477, −959), while same binding sites were predicted in tambaqui *cyp19a1a* promoter region, but in a position more distant from transcription start site (up to 1 kb).

3.3. Phylogenetic analysis of teleost's Cyp19a1a and Cyp19a1b

A total of 86 aromatase protein sequences from 52 genomes (Supplementary Fig. 4) were retrieved to assess the phylogenetic relationship and to predict ancestral gene duplications and losses among vertebrates. The reconstructed aromatase phylogenetic tree clustered the sequences in three major groups of vertebrates, namely chondrichthyan, sarcopterygian and actinopterygian (the last comprehending non-teleosts, basal teleosts and clupeocephala) (Fig. 2a). In both Cyp19a1a and Cyp19a1b subclades, *C. macropomum*, together with *P. nattereri, Hepsetus odoe* and *A. mexicanus* sequences represented the Characiformes clade which, together with the Siluriformes and Gymnotiformes, represents the unexplored Characiphysi group.

The single *CYP19A1* of non-teleost actinopterygians, reedfish (*Erpetoichthys calabaricus*), spotted gar (*Lepisosteus oculatus*) and the sterlet sturgeon (*Acipenser ruthenus*), branched basis to the well supported actinopterygian teleost clade (bootstrap probabilities in ML/posterior probability in Bayesian tree inference: 100/1.0), in agreement with the phylogenetic relationships within this major group. The single Cyp19a1 of basal teleosts, including European eel (*Anguila anguila*), Asian arowana (*Scleropages formosus*) and the Amazonian pirarucu (*Arapaima gigas*) grouped in a well-supported cluster at the base of the clade composed by duplicated Cyp19a1 sequences from the remaining teleosts (Clupeocephala).

In teleosts, the duplicated aromatase clade most likely resulted from a specific third round of whole genome duplication event (TWGD) that occurred in the lineage. Additional paralogs that were found in teleosts, such as salmonids and carps, that independently underwent another round (4R) of WGD, and in the non-teleost *A. ruthenus,* which lineage experienced the independently 3R after their divergence from the other ray-finned fish, are indicated. The absence of *cyp19a1a* sequence in the black bullhead catfish (*Amerius melas*) and in The Lake Tanganyika sardine (*Limnothrissa miodon*; Fig. 2b), as well as *cyp19a1b* additional paralogs in the common carp (*Cyprinus carpio)* and Atlantic salmon (*Salmo salar*) databases suggest gene loss in these lineages.

3.4. Synteny analysis

The order of 38 (out of 44) genes is perfectly conserved between tambaqui's scaffold NW_023496020.1 carrying *cyp19a1a* and the red-bellied piranha's chromosome 7 (Fig. 3a). We also found that *cyp19a1a* from Mexican tetra surface fish (*A. mexicanus*) is located at an unplaced scaffold (NW_019171364.1), even though the assembly is at chromosome level, sharing 22 genes with the congeneric representative species (11 in complete synteny). However, small rearrangements are evident as part of tambaqui and red-bellied piranha's *cyp19a1a* 3′ neighboring genes are in synteny with Mexican tetra chromosome 18.

Compared to distant related species, the degree of rearrangements in *cyp19a1a* adjacencies are more evident. In zebrafish (*D. rerio*) chromosome 8, sixteen genes out of 44 were found to be conserved around the *cyp19a1a*, whereas thirteen neighboring genes are in 3′ adjacencies and only three were found in 5′. In contrast, from 25 *cyp19a1a* conservative neighboring genes shared between japanese medaka (*O. latipes*) chromosome 3 and red-bellied piranha chromosome 7, eighteen are in 5′ and only seven are in 3′ adjacencies. Moreover, the reversed order of *usp8* from the twenty-second gene upstream to *cyp19a1a* in red-bellied piranha to the eighth gene downstream japanese medaka *cyp19a1a*, indicates local rearrangement in the *cyp19a1a* gene environment. Despite these particularities, the comparison of all ortholog regions showed that the syntenic organization of the gene group *gldn-cyp19a1a-tnfaip8l3* is highly conserved among the species analyzed.

In regard to *cyp19a1b* neighboring genes, there is a strong synteny conservation among red-bellied piranha chromosome 11, tambaqui scaffold NW_023494792.1 and mexican tetra chromosome 8, despite the duplication of some genes such as *vw5a*-like that varied in the number of copies in each species (Fig. 3b). In the zebrafish chromosome 25,

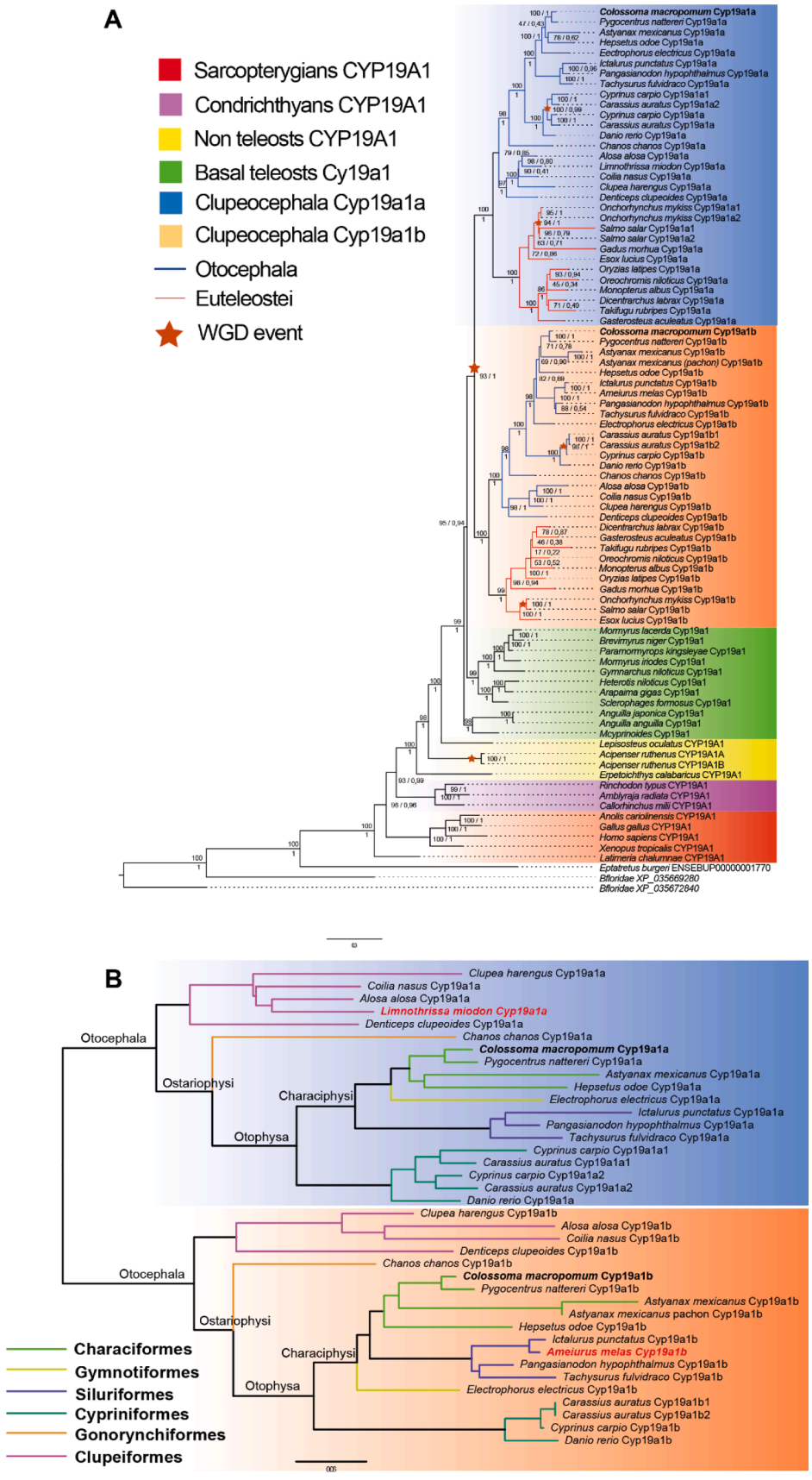

Fig. 2. Phylogenetic tree of vertebrate aromatase amino acid sequences. A) Phylogenetic tree built using Maximum Likelihood and Bayesian analysis on the basis of 86 aromatase sequences from 52 species representative of Sarcopterygians, Chondrichthyans and Actinopterygians. The Cephalochordata *Branchiostoma floridae* CYP19 tandem duplicated protein sequences and the Cyclostomata *Eptatretus burgeri* CYP19 were used as the outgroup. Values on the nodes are relative to the Bootstrap of ML (%; upper of branch or left of slash) and posterior probability of BI (lower of branch or right of slash) respectively. The red stars show the position of the teleost-specific (Ts3R), the salmonid-specific (Ss4R), the carp-specific (Cs4R) and sterlet (Ars3R) whole-genome duplications (WGD) after the 1R/2R events. The scale bar indicates the average substitutions per site; the dotted lines associate the taxon names with the branch ends. B) Phylogenetic tree of aromatase sequences from representative species of Otocephala group evidencing single aromatase Cyp19a1a in *Limnothrissa miodon* and single aromatase Cyp19a1b in *Ameiurus melas*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

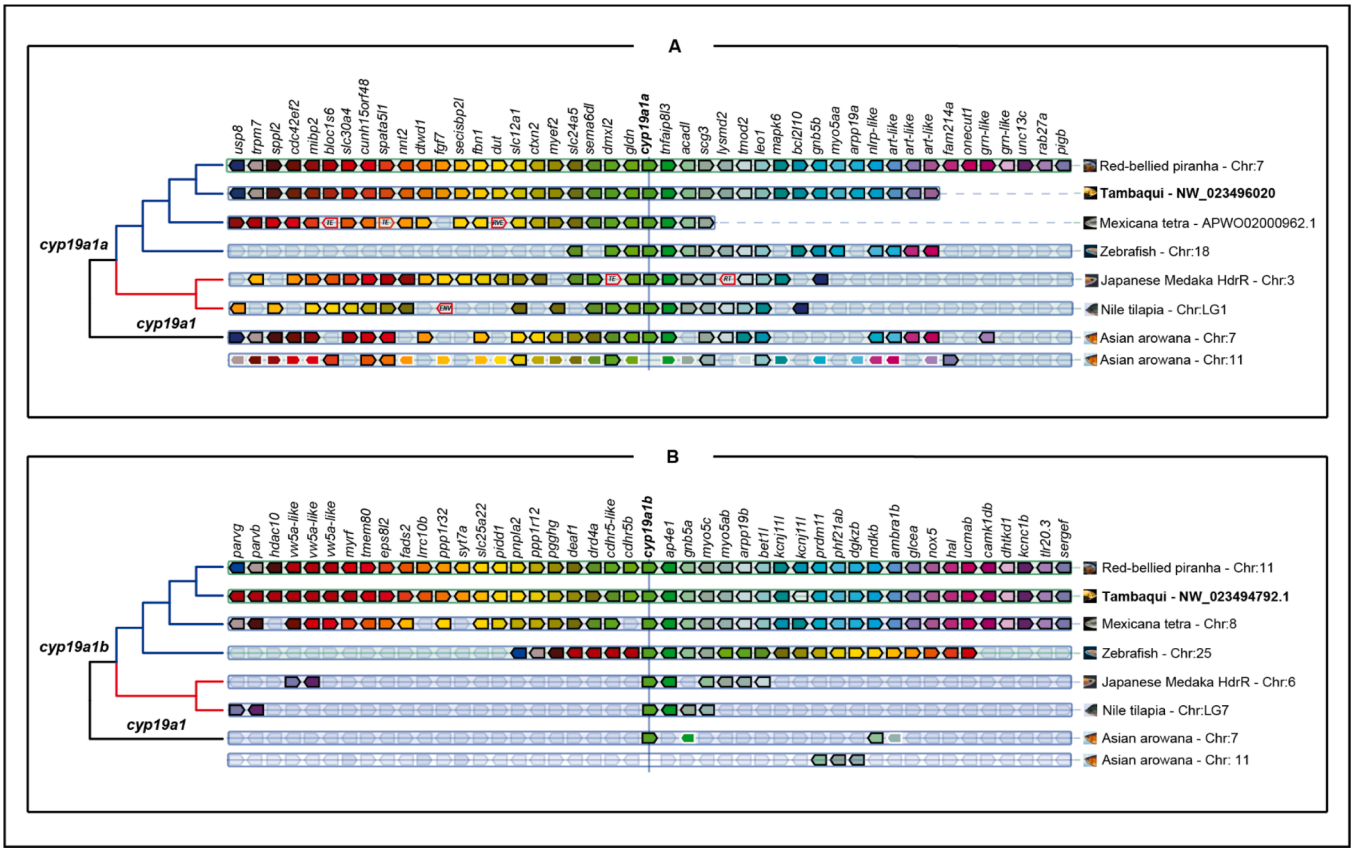

Fig. 3. Synteny analysis of A) *cyp19a1a* and B) *cyp19a1b* neighbouring genes in *C. macropomum* and in different species that share the same ancestral species. Orthologs of these genes in other species are shown in matching colours. Genes outlined in white are paralogs of the genes in the same colour but outlined in black. Shaded genes correspond to genes that are not orthologous to any genes from the species used in the query (reference). The tree on the left of the display is the phylogenetic tree of the gene shown in the middle that intersects the vertical line.

however, the *vw5a*-like tandem duplications lie upstream and adjacent to *cyp19a1b*, while these genes are distant located upstream of *cyp19a1b* in the characiform representatives. Additionally, the zebrafish region features the reversed order and orientation of a block containing fourteen genes, including the *cdhr5b* which is the conserved *cyp19a1b* upstream adjacent gene among red-bellied piranha, tambaqui and mexican tetra.

On the other hand, the japanese medaka chromosome 6 segment shares only seven genes with red-bellied piranha's *cyp19a1b* neighboring genes. The *kcnc1b* and *sergef* genes are in the reverse order and the remaining five genes form the syntenic gene group *cyp19a1b, ap4e1, gnb5a-myo5c, myo5ab and arpp19b* are positionally conserved in all species analyzed. Moreover, a second *gldn* paralog is present and is adjacent to japanese medaka *cyp19a1b* upstream region, which is a conserved feature previously described by Lin et al (2020) in the single *cyp19a1* from basal and non-teleost fishes, as well as in *cyp19a1a* from Clupeocephala. Such results indicates that critical rearangements occured in the *cyp19a1b* cis regulatory region across teleost lineages.

3.5. Transcript level profiling of estrogen signaling cyp19a1a, cyp19a1b and esr2a in tambaqui juvenile, gonads and brain.

Transcripts of *cyp19a1a*, *cyp19a1b* and *esr2a* could be detected in the pool of larvae (12 h post-hatching) samples and it was used as the calibrator for the expression in pre- and differentiating fish (decapitated juveniles); ovaries, testes and brain of differentiated tambaqui (Fig. 4). The expression of *cyp19a1a* was marked by a slight, but significant, increase in 2 cm juvenile trunks. The values remained then unchanged until they reached 6 cm total length. However, in pre-differentiating juveniles of 3 and 4 cm long, the expression of *cyp19a1a* tended to split into two groups. After the end of ovary differentiation (6 cm long) the values gathered again. Maturing ovaries and testes presented the highest expression of *cyp19a1a*.

On the contrary, the expression of *cyp19a1b* increased 10 fold from hatched larvae to juveniles with 1.5 cm and presented a large variation of values in fish of 3 cm, i. e. prior to morphological differentiation. After this phase, the values decreased, but two groups could be identified at 4 and 6 cm length. The expression of *cyp19a1b* in trunks of pre- and differentiating tambaqui was very similar to those found in adult gonads and male brains, and lower than in female brains.

Esr2a was overexpressed in 2 cm (pre-differentiating) fish ($p < 0.005$). At this stage, the expression average increased over 1200 fold, and the individual values could be distinguished in two groups, one whose expression was around 300 times over the hatched larvaes and other with values above thousand times.

In the differentiated gonads, *cyp19a1a* transcription was highest in the testis. Among females, *cyp19a1a* transcripts were more abundant in maturing ovaries while *esr2a* was more expressed in immature ovaries. Furthermore, in sexually immature males the expression of *cyp19a1b* and *esr2a* transcripts were higher in testis than in the brain, while in immature females *cyp19a1b* was more expressed in the brain than in the ovaries. However, both *cyp19a1a* and *cyp19a1b* transcripts were more abundant in the brain of immature females compared to immature males.

3.6. Effect of EDCs fadrozole and tamoxifen on tambaqui sex development

There was no mortality during the experiment, nor any visual effects on the body or behavior of the fish during or after the treatments with

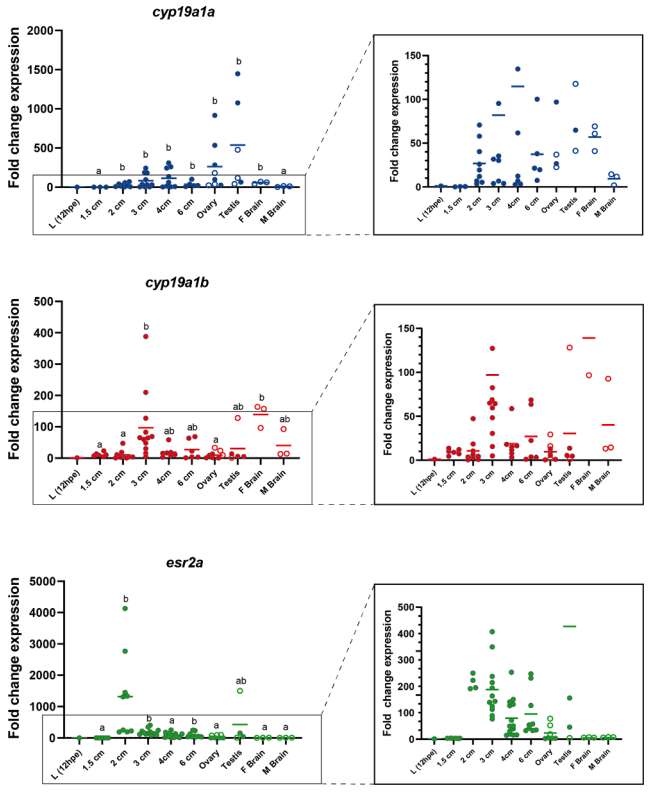

Fig. 4. Expression of *cyp19a1a*, *cyp19a1b* and *esr2a* from different stages of juvenile development and gonads of *C. macropomum*. Results are presented as scatterplots of individual expression values of relative expression to the values of the pool of 12 h post-hatched larvae (L) and the averages as lines. Gonad and brain samples from immature fish are represented by unfilled circles and from maturing fish by filled circles. Different letters mean different averages between the groups. $P < 0.005$.

fadrozole (Fz) or tamoxifen (Tx). There was no masculinization effect in neither treatment (n = 74 to 97 fish evaluated/treatment; Supplementary Table 2). Exposure to Fz reduced the E2 plasma concentration at the end of the treatment (fortieth day), but after 20 days the values were similar to the controls (Fig. 5). Exposure to Tx did not affect E2 plasma values. In both treatments, twenty days after the beginning of EDCs

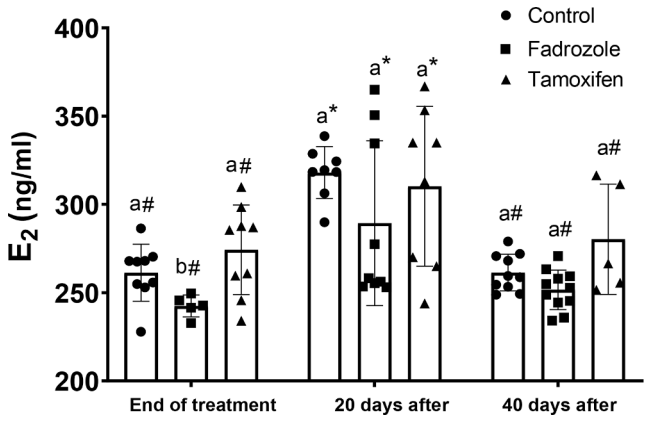

Fig. 5. Concentration of 17β estradiol (E_2) in *C. macropomum* treated with fadrozole (50 mg kg^{-1} of food) or tamoxifen (200 mg kg^{-1}) during 40 days. Samples were collected at the end of treatment, 20 and 40 days after the end of the treatment. Each bar represents the mean ± SD. Different letters denote statistical differences between treatments at the same point and different symbols show statistical differences between the collection points in the same treatment ($p < 0.05$).

administration, *cyp19a1a* transcription was down-regulated in comparison to controls. Otherwise, *cyp19a1b* transcription at day 20 was only down-regulated by Tx and neither Tx nor Fz altered the expression of *esr2a*. No further differences in gene expression were found in EDC-treated tambaqui compared to the control group at 40 days (end) of treatment (Fig. 6). Since no sex changes were observed, it is expected that after some time the transcriptional levels return to normal.

4. Discussion

The role of aromatase on the sex differentiation of vertebrates has been frequently addressed at the gonad level where estradiol, acting via estradiol receptors, is a pivotal factor. There is a single aromatase gene (*CYP19A1*) in Chondrichthyes, Sarcopterygii and non-teleost Actinopterygii, but some mammals as pigs, peccaries and other suiformes have two or more genes resulting from the duplication of a common ancestor (Conley et al., 2009). However, many teleosts exhibit two functionally conserved aromatase genes, namely *cyp19a1a* and *cyp19a1b*, as a result of the third round of whole-genome duplication (TWGD) specific to the teleost lineage. The most recent phylogenomic study revealed a single 3R-aromatase paralog *cyp19a1* present in basal teleosts including eels and Asian arowana, that most likely exhibits conserved properties of the ancestral single *CYP19A1*, such as the conservation of the brain-and gonad-type responsive elements in the *cyp19a1* promoter, in

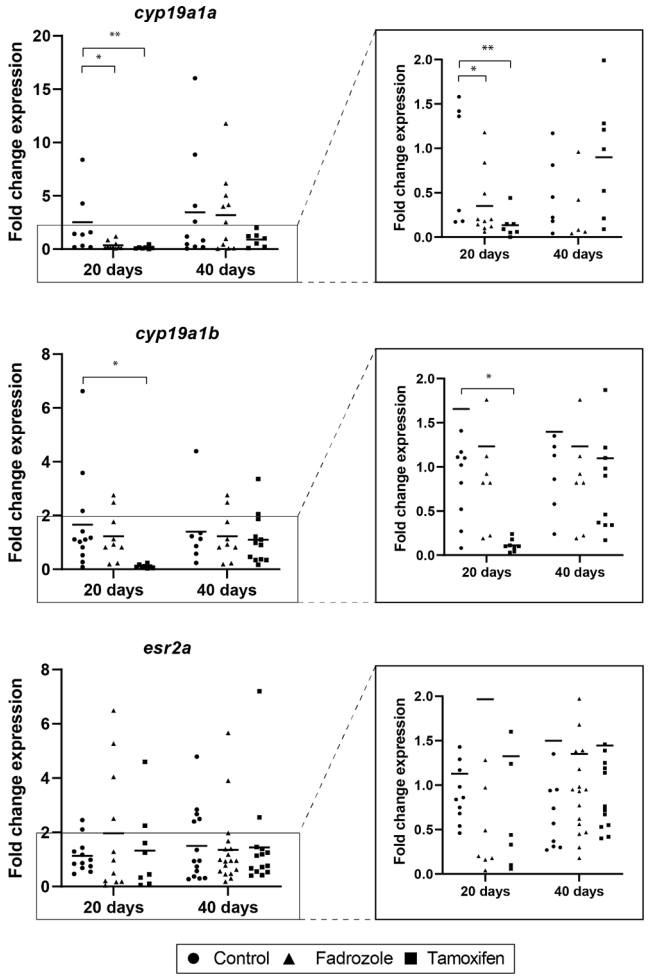

Fig. 6. Transcript abundance of *cyp19a1a*, *cyp19a1b* and *esr2a* in undifferentiated tambaqui trunks during treatment with fadrozole (50 mg kg^{-1} of food) or tamoxifen (200 mg kg^{-1}). Results are presented as scatterplots with individual expression values. Asterisks denote significant differences between control and treatment group ($p < 0.05$).

agreement with the main expression and role of the single *cyp19a1* in both brain and gonad (Lin et al., 2020). In the present study, more aromatase sequences from the Otocephala representatives were included in the phylogenetic analysis which displayed a tree topology similar to previous studies. We also discovered the existence of a single aromatase in the black bullhead fish (*Ameiurus melas*; Siluriformes) and *Limnothrissa miodon* (Clupeiformes) genomes, which correspond to the Cyp19a1b and Cyp19a1a clade, respectively.

A deep phylogenomic view into the Otocephala clade clearly shows additional possibilities regarding the function of duplicated aromatases. Despite the great conservation of *cyp19a1a* and *cyp19a1b* neighboring genes across Characiformes, the variations in their organization compared to the commonly studied zebrafish (*Danio rerio*) and japanese medaka (*Oryzias latipes*) draw our attention. The zebrafish genome notably seems to have experienced many interchromosomal rearrangements during evolution by extensive translocations, while the medaka genome has preserved its Clupeocephala ancestral genomic structure without undergoing major interchromosomal rearrangements (Kasahara et al., 2007). Thus, the variations detected right upstream the *cyp19a1b* genomic location among these species suggest that lineage-specific rearrangement events happened in this locus after the split of Clupeocephala into Otocephala and Euteleostei which may have somehow impacted the regulatory area in different levels.

Changes in the genomic neighborhood after the split from the common ancestor are linked to divergences in gene expression levels in different tissues, potentially leading to phenotypic divergences (De et al., 2009). Although the adjacent genes of *cyp19a1a* among all species compared are conserved, the presence of transposable element (TE) traces such as transposase at this location served as a variation source in the gonadal aromatase genomic environment. Additionally, TEs can be the source of developmental innovations through their recruitment as new coding sequences and new ncRNAs, and by acting as regulatory sequences (Etchegaray et al., 2021). Indeed, an alternate 5′ regulatory region of the human and mouse *CYP19A1* includes a promoter derived from LTR retrotransposon that drives expression in the placenta, as well as an associated upstream enhancer (Kamat et al., 2005). Therefore, TE traces found within *gldn-cyp19a1a* intergenic region in tambaqui, covering the promoter region of both genes, can be associated with notable differences in bp size compared to *cdhr5b-cyp19a1b* intergenic region and with the alterations in *cis*-regulatory elements detected among the groups analyzed.

In this context, the tambaqui *cyp19a1a* promoter sequence provided the most controversial insight due to the absence of steroidogenic factor 1 (SF-1) responsive element, conversely found in the 5′-flanking regions of all Clupeocephala *cyp19a1a*, basal teleost *cyp19a1* and non-teleost actinopterygian *CYP19A1* sequences (Lin et al., 2020). The modulation of *cyp19a1a* expression by SF-1 nuclear receptor, also referred to as AD4BP, NR5A1 or FTZ-F1, has already been demonstrated in Nile tilapia (Yoshiura et al., 2003), medaka (Watanabe et al., 1999), zebrafish (von Hofsten and Olsson, 2005), human ovary (Michael et al., 1995) and rat granulosa cells (Shapiro et al., 1996). Moreover, even though SF-1 binding sites were found in other Characiformes, such as red-bellied piranha and mexican tetra *cyp19a1a* promoter sequences, the highest dissimilarity (DI) values compared to the consensus sequence predicted in all other teleost investigated suggest that remarkably differences in *cyp19a1a* regulation may occur in this clade, even between phylogenetically close species.

Regarding the *cyp19a1b* promoter region, no SF-1 response elements were predicted in tambaqui, similar to other Clupeocephala species such as zebrafish and tilapia, but with the exception of medaka (Lin et al., 2020). This particularity can be associated with the retention of *gldn* paralog linked to *cyp19a1b* in Japanese medaka. Such diversity in aromatase cis regulatory region may result in disparities in gene expression, affecting the plasticity of the gene regulation underlying gonadal morphogenesis (Nakamoto et al., 2018; Li & Ge, 2020; Imarazene et al., 2021). Overall, although both tambaqui *cyp19a1a* and *cyp19a1b* promoter regions share many putative transcription factor binding sites, the differences in position and number of copies may reflect the particularities in their individual regulation and expression.

Although the main molecular pathway towards tambaqui ovary and testis differentiation could be characterized by RNA seq of the trunks of undifferentiated tambaqui, only the *cyp19a1b* (significantly more expressed in the future males than females), but not *cyp19a1a,* could be detected by this approach (Lobo et al., 2020). In the present study, the expression of both genes were identified by real-time quantitative PCR (qPCR) during the early stages of sex development. As expected, based on the previous RNA seq data, *cyp19a1b* transcripts in tambaqui larvae trunks increased after hatching (while *cyp19a1a* decreased) and peaked at very high values just prior to differentiation. On the other hand, the increase in *cyp19a1a* expression occurs later (fish from 2 cm) and the pattern of two distinct groups could be observed in fish of 3 and 4 cm length. Although we couldn't link the transcription profile to the corresponding sex of the individuals, in *A. mexicanus,* a non-dimorphic expression profiles of *cyp19a1a* were detected in males and females during early development (Imarazene et al., 2021).

Given the presence of a well-conserved estrogen-responsive element (ERE) in the proximal promoter region of *cyp19a1b* (Fig. 1b), and the positive auto-regulatory loop of estradiol via ERs in fish brains (Diotel et al., 2010), we traced a parallel between *esr2a* and *cyp19a1b* transcription profiles during sex development of tambaqui. From hatching, the expression of estrogen receptor beta 1 (*esr2a)* increased in pre-differentiating tambaqui up to 1000-fold, which are values far above the *esr2a* expression in differentiated gonads (in immature and mature juveniles). Moreover, at this undifferentiated phase, *esr2a* expression in fish trunks split into two distant groups and in fish bigger are back to similar (and low) transcription values. Soon after the split of *esr2a*, the expression of *cyp19a1b* rises, reaching levels above the recorded in brain and gonads of adult fish. *cyp19a1b* transcripts then decrease during morphological differentiation (4–6 cm), at the same time as *cyp19a1a* expression splits into two groups. Therefore, we suspect that *cyp19a1b* and *esr2a* are involved in the gene pathway that drives tambaqui sex differentiation, while *cyp19a1a* seems to be a consequence of the already triggered E_2-ER system activation and might be responsible for the further maintenance of the developing ovary. The involvement of *cyp19a1b* in sex differentiation has been suggested in pejerrey *Odontesthes bonariensis,* however with high expression in brain, and not in the trunk, preceding the expression of *cyp19a1a* in the trunk before the first signs of gonadal differentiation (Karube et al., 2007; Strobl-Mazzulla et al., 2008).

Fradozole reduced the E_2 plasma levels at the end of treatment, and down regulated *cyp19a1a* (significantly) and *cyp19a1b* at the beginning of the treatment, and tamoxifen reduced the expression of both aromatase coding genes, but without effect on the levels of estradiol. These observations are in agreement with the physiological effects of EDCs in fish in which SERMs blocks estrogen function by binding to ERs and suppressing *cyp19a1a* or *cyp19a1b* expression, while AIs decrease E_2 levels by acting directly on aromatase production (Cheshenko et al., 2008; Scholz and Klüver, 2009). However, none of the treatments changed the expression of *esr2a* and nor were able to distress the sex ratio in tambaqui. Such data is contradictory to the masculinization effect of aromatase inhibitors widely reported in a variety of fish species (Babiak et al., 2012, Shen et al., 2015, Evliyaoğlu et al., 2019, Qin et al., 2020, Ayobahan et al., 2020), even at very low concentrations (Muth-Köhne et al., 2016, Joshi et al., 2019). And also incompatible to the tamoxifen effect that, acting as either agonist or antagonist, stimulating or inhibiting gonadal differentiation, leads to abnormal sexual development and biased sex ratio (Singh et al, 2012, Navarro-Martín et al., 2009, Singh et al., 2015, García-Hernández et al., 2016, Pandit et al., 2017). Such complexity in EDC responses has been reported in *Astyanax altiparane* as a result of a mechanism of resilience, in which males resisted the E_2-induced feminization by enhancing the expression of genes related to testicular differentiation (Martinez-Bengochea et al.,

2020). Likewise, zebrafish females are able to resist heat-induced masculinization, maintaining ovarian morphology despite changing their gonadal transcriptome to a testis-like one (Ribas et al., 2017) and *A. burtoni* ovaries cannot be transformed into functional testis by AI, but rapid changes towards a male-like phenotype happens, such as body coloration, hormone levels and behavior (Göppert et al., 2016).

The higher expression of *cyp19a1a* in testis than in ovaries and the upregulation of *esr2a* in all pre-differentiation fish suggest that the estrogen pathway is necessary for tambaqui gonad development and maintenance in both sexes. Nonetheless, the reproductive state may influence the level of *cyp19a1a*, and more reproductive states should be analyzed in the future. Estrogen production via aromatase in fish testis has been described to be involved in spermatogenesis, where estrogen is necessary for the renewal of spermatogonial stem cells and possibly differentiation (Schulz et al., 2010; Kobayashi et al., 2011). Similar *cyp19a1a* transcription sex bias towards males were observed in *A. mexicanus*, with high values in testes and low basal levels in ovaries during gametogenesis (Imarazene et al., 2021). In another Characid, *A. altiparanae, cyp19a1a* are expressed equally in adult testes and ovaries (Martinez-Bengochea et al., 2020).

The expression pattern of *cyp19a1a* and *cyp19a1b* was more abundant in female brains when compared to even some individual gonads. This observation also confronts the classical assumption of the tissue-specificity of aromatase expression (Guiguen et al., 2010). In most fishes studied so far, the *cyp19a1b* is more expressed in the brain of males than of the females and it is responsible for the male sexual behaviour such as aggression and court strategies (Goto-kazeto et al., 2004; Black et al., 2005; Strobl-Mazzulla et al., 2008; Gonçalves et al., 2008; Renn et al., 2008; Colman et al., 2009; Diotel et al., 2010; Huffman et al., 2013; Göppert et al., 2016, Silva de Assis et al., 2018). However, it was recently demonstrated through the *cyp19a1b* expression on the female brain of the African cichlid fish, *Astatotilapia burtoni*, that the estrogenic signaling in female brain might have important roles in the regulation of reproductive cycling and social behaviors (Maruska et al., 2020). These results on the *cyp19a* expression in gonads and brains of male and female tambaqui are uncommon in fish, and therefore require proper and specific investigation.

Taken together, the present study reveals another level of flexibility on the sex-differentiation pathway of teleosts, contributing to the recent discussion on the unexpected expression of genes related to ovarian differentiation, especially in characids. Such an unconventional mechanism can be explained by the modifications on the *cyp19a1a* promoter region as a consequence of transposable elements' activity. Indeed, the chromosomal evolution of Serrasalmidae involved different rearrangements (fissions, translocations and pericentric inversions) due to TEs, which have crucial roles in the differentiation and evolution of sex chromosomes. During tambaqui sex differentiation, besides the bimodal expression of *cyp19a1b* first, followed by *cyp19a1a*, the outstanding bimodal gene expression of *esr2a* from hatching to differentiation, suggests that estrogen receptor is crucial for the sex identity in *C. macropomum*. Likewise, although the endocrine disruptors caused some changes in *cyp19* genes transcription, the *esr2a* expression and sex ratio were not altered by those compounds. In conclusion, our data on the complete structure, synteny, phylogenetic and expression characterization of tambaqui *cyp19a1a* and *cyp19a1b* provided new information concerning the evolution and evolutionary impact of TEs in the Neotropical freshwater fish genomes, which can be considered groundbreaking in the classical knowledge regarding estrogens and estrogen receptors duties in fish sex differentiation.

CRediT authorship contribution statement

R.V. Paixão: Methodology, Investigation, Validation, Formal analysis, Writing – original draft. **G.F. Silva:** Conceptualization, Supervision, Investigation, Writing – review & editing. **A.R. Caetano:** Resources, Data curation, Writing – review & editing. **L.C. Cintra:** Resources, Data curation. **E.S. Varela:** Resources, Data curation. **F.L.A. O'Sullivan:** Conceptualization, Supervision, Funding acquisition, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This study was financialy supported by Embrapa, Empresa Brasileira de Pesquisa Agropecuária (SEG 12.16.05.018.00.00) and by CAPES, Coordenação de Aperfeiçoamento Pessoal de Nível Superior, and FAPEAM, Fundação de Amparo à Pesquisa do Estado do Amazonas, through the cooperation program for the consolidation of graduate studies and human resources training in Amazonas – Brazil (CAPES/FAPEAM – Edital N° 24/2014).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2022.146795.

References

Afonso, L.O.B., Wassermann, G.J., Terezinha de Oliveira, R., 2001. Sex reversal in Nile tilapia (Oreochromis niloticus) using a nonsteroidal aromatase inhibitor. J. Exp. Zool. 290 (2), 177–181. https://doi.org/10.1002/jez.1047.

Amenyogbe, E., Chen, G., Wang, Z., Lu, X., Lin, M., Lin, A.Y., 2020. A Review on Sex Steroid Hormone Estrogen Receptors in Mammals and Fish. Int. J. Endocrinol. 1–9. https://doi.org/10.1155/2020/5386193.

Ayobahan, S.U., Eilebrecht, S., Baumann, L., Teigeler, M., Hollert, H., Kalkhof, S., Eilebrecht, E., Schäfers, C., 2020. Detection of biomarkers to differentiate endocrine disruption from hepatotoxicity in zebrafish (*Danio rerio*) using proteomics. Chemosphere 240, 124970. https://doi.org/10.1016/j.chemosphere.2019.124970.

Babiak, J., Babiak, I., van Nes, S., Harboe, T., Haugen, T., Norberg, B., 2012. Induced sex reversal using an aromatase inhibitor, Fadrozole, in Atlantic halibut (*Hippoglossus hippoglossus* L.). Aquaculture 324–325, 276–280. https://doi.org/10.1016/j.aquaculture.2011.11.001.

Black, M.P., Balthazart, J., Baillien, M., Grober, M.S., 2005. Socially induced and rapid increases in aggression are inversely related to brain aromatase activity in a sex-changing fish, Lythrypnus dalli. Proc. Roy. Soc. B: Biol. Sci. 272 (1579), 2435–2440. https://doi.org/10.1098/rspb.2005.3210.

Chang, X., Kobayashi, T., Senthilkumaran, B., Kobayashi-Kajura, H., Sudhakumari, C.C., Nagahama, Y., 2005. Two types of aromatase with different encoding genes, tissue distribution and developmental expression in Nile tilapia (*Oreochromis niloticus*). Gen. Comp. Endocrinol. 141 (2), 101–115. https://doi.org/10.1016/j.ygcen.2004.11.020.

Cheshenko, K., Pakdel, F., Segner, H., Kah, O., Eggen, R.I.L., 2008. Interference of endocrine disrupting chemicals with aromatase CYP19 expression or activity, and consequences for reproduction of teleost fish. Gen. Comp. Endocrinol. 155 (1), 31–62. https://doi.org/10.1016/j.ygcen.2007.03.005.

Colman, J.R., Baldwin, D., Johnson, L.L., Scholz, N.L., 2009. Effects of the synthetic estrogen, 17α-ethinylestradiol, on aggression and courtship behavior in male zebrafish (*Danio rerio*). Aquat. Toxicol. 91 (4), 346–354. https://doi.org/10.1016/j.aquatox.2008.12.001.

Conley, A.J., Corbin, C.J., Hughes, A.L., 2009. Adaptive evolution of mammalian aromatases: lessons from Suiformes. J. Exp. Zool. Part A: Ecol. Genet. Physiol. 311A (5), 346–357. https://doi.org/10.1002/jez.451.

Dang, Z., Kienzler, A., 2019. Changes in fish sex ratio as a basis for regulating endocrine disruptors. Environ. Int. 130, 104928 https://doi.org/10.1016/j.envint.2019.104928.

De, S., Teichmann, S.A., Babu, M.M., 2009. The impact of genomic neighborhood on the evolution of human and chimpanzee transcriptome. Genome Res. 19 (5), 785–794. https://doi.org/10.1101/gr.086165.108].

Devlin, R.H., Nagahama, Y., 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208 (3–4), 191–364. https://doi.org/10.1016/S0044-8486(02)00057-1.

Diotel, N., Page, Y.L., Mouriec, K., Tong, S.-K., Pellegrini, E., Vaillant, C., Anglade, I., Brion, F., Pakdel, F., Chung, B., Kah, O., 2010. Aromatase in the brain of teleost fish:

Expression, regulation and putative functions. Front. Neuroendocrinol. 31 (2), 172–192. https://doi.org/10.1016/j.yfrne.2010.01.003.
Etchegaray, E., Naville, M., Volff, J.-N., Haftek-Terreau, Z., 2021. Transposable element-derived sequences in vertebrate development. Mobile DNA 12 (1), 1. https://doi.org/10.1186/s13100-020-00229-5.
Evliyaoğlu, E., Eroldoğan, O.T., Yılmaz, H.A., Genç, M.A., Genç, E., Duncan, N., Aktaş, M., Güroy, D., 2019. Artificial sex reversal of white grouper (*Epinephelus aeneus*) utilizing aromatase inhibitor (Fadrozole). Aquac. Res. 50 (5), 1539–1546. https://doi.org/10.1111/are.14030.
Fernandino, J.I., Hattori, R.S., 2019. Sex determination in Neotropical fish: Implications ranging from aquaculture technology to ecological assessment. Gen. Comp. Endocrinol. 273, 172–183. https://doi.org/10.1016/j.ygcen.2018.07.002.
García-Hernández, M.P., Rodenas, M.C., Cabas, I., García-Alcázar, A., Chaves-Pozo, E., García-Ayala, A., 2016. Tamoxifen disrupts the reproductive process in gilthead seabream males and modulates the effects promoted by 17α-ethynylestradiol. Comp. Biochem. Physiol. C: Toxicol. Pharmacol. 179, 94–106. https://doi.org/10.1016/j.cbpc.2015.09.005.
Gonçalves, D., Teles, M., Alpedrinha, J., Oliveira, R., 2008. Brain and gonadal aromatase activity and steroid hormone levels in female and polymorphic males of the peacock blenny *Salaria pavo*. Horm. Behav. 54 (5), 717–725. https://doi.org/10.1016/j.yhbeh.2008.07.014.
Göppert, C., Harris, R.M., Theis, A., Boila, A., Hohl, S., Rüegg, A., Hofmann, H.A., Salzburger, W., Böhne, A., 2016. Inhibition of Aromatase Induces Partial Sex Change in a Cichlid Fish: Distinct Functions for Sex Steroids in Brains and Gonads. Sexual Develop. 10 (2), 97–110. https://doi.org/10.1159/000445463.
Goto-Kazeto, R., Kight, K.E., Zohar, Y., Place, A.R., Trant, J.M., 2004. Localization and expression of aromatase mRNA in adult zebrafish. Gen. Comp. Endocrinol. 139 (1), 72–84. https://doi.org/10.1016/j.ygcen.2004.07.003.
Guiguen, Y., Fostier, A., Piferrer, F., Chang, C.-F., 2010. Ovarian aromatase and estrogens: A pivotal role for gonadal sex differentiation and sex change in fish. Gen. Comp. Endocrinol. 165 (3), 352–366. https://doi.org/10.1016/j.ygcen.2009.03.002.
Huffman, L.S., O'Connell, L.A., Hofmann, H.A., 2013. Aromatase regulates aggression in the African cichlid fish *Astatotilapia burtoni*. Physiol. Behav. 112–113, 77–83. https://doi.org/10.1016/j.physbeh.2013.02.004.
Imarazene, B., Beille, S., Jouanno, E., Branthonne, A., Thermes, V., Thomas, M., Herpin, A., Rétaux, S., Guiguen, Y., 2021. Primordial Germ Cell Migration and Histological and Molecular Characterization of Gonadal Differentiation in Pachón Cavefish *Astyanax mexicanus*. Sexual Develop. 1–18 https://doi.org/10.1159/000513378.
Joshi, H.D., Tiwari, V.K., Gupta, S., Sharma, R., Lakra, W.S., Sahoo, U., 2019. Application of nanotechnology for the production of masculinized Tilapia, *Oreochromis niloticus* (Linnaeus, 1758). Aquaculture 511, 734206. https://doi.org/10.1016/j.aquaculture.2019.734206.
Kamat, A., Smith, M.E., Shelton, J.M., Richardson, J.A., Mendelson, C.R., 2005. Genomic Regions that Mediate Placental Cell-Specific and Developmental Regulation of Human Cyp19 (Aromatase) Gene Expression in Transgenic Mice. Endocrinology 146 (5), 2481–2488. https://doi.org/10.1210/en.2004-1606.
Karube, M., Fernandino, J.I., Strobl-Mazzulla, P., Strüssmann, C.A., Yoshizaki, G., Somoza, G.M., Patiño, R., 2007. Characterization and expression profile of the ovarian cytochrome P-450 aromatase (cyp19A1) gene during thermolabile sex determination in pejerrey, Odontesthes bonariensis. J. Exp. Zool. Part A: Ecol. Genet. Physiol. 307 (11), 625–636.
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., Kohara, Y., 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447 (7145), 714–719. https://doi.org/10.1038/nature05846.
Kobayashi, Y., Nozu, R., Nakamura, M., 2011. Role of estrogen in spermatogenesis in initial phase males of the three-spot wrasse (Halichoeres trimaculatus): Effect of aromatase inhibitor on the testis. Dev. Dyn. 240 https://doi.org/10.1002/dvdy.22507.
Lau, E.-S.-W., Zhang, Z., Qin, M., Ge, W., 2016. Knockout of Zebrafish Ovarian Aromatase Gene (cyp19a1a) by TALEN and CRISPR/Cas9 Leads to All-male Offspring Due to Failed Ovarian Differentiation. Sci. Rep. 6 (1), 37357. https://doi.org/10.1038/srep37357.
Li, J., Ge, W., 2020. Zebrafish as a model for studying ovarian development: Recent advances from targeted gene knockout studies. Mol. Cell. Endocrinol. 507, 110778 https://doi.org/10.1016/j.mce.2020.110778.
Lin, C.-J., Maugars, G., Lafont, A.-G., Jeng, S.-R., Wu, G.-C., Dufour, S., Chang, C.-F., 2020. Basal teleosts provide new insights into the evolutionary history of teleost-duplicated aromatase. Gen. Comp. Endocrinol. 291, 113395 https://doi.org/10.1016/j.ygcen.2020.113395.
Lobo, I. K. C., Nascimento, Á. R. do, Yamagishi, M. E. B., Guiguen, Y., Silva, G. F. da, Severac, D., Amaral, A. da C., Reis, V. R., Almeida, F. L. de. 2020. Transcriptome of tambaqui *Colossoma macropomum* during gonad differentiation: Different molecular signals leading to sex identity. Genomics 112(3), 2478–2488. 10.1016/j.ygeno.2020.01.022.
Martinez-Bengochea, A., Doretto, L., Rosa, I.F., Oliveira, M.A., Silva, C., Silva, D.M.Z.A., Santos, G.R., Santos, J.S.F., Avelar, M.M., Silva, L.V., Lucianelli-Junior, D., Souza, E. R.B., Silva, R.C., Stewart, A.B., Nakaghi, L.S.O., Valentin, F.N., Nóbrega, R.H., 2020. Effects of 17β-estradiol on early gonadal development and expression of genes implicated in sexual differentiation of a South American teleost, *Astyanax altiparanae*. Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 248–249, 110467 https://doi.org/10.1016/j.cbpb.2020.110467.
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. Comparative Neurol. 528 (15), 2499–2522. https://doi.org/10.1002/cne.24908.
Michael, M.D., Kilgore, M.W., Morohashi, K., Simpson, E.R., 1995. Ad4BP/SF-1 Regulates Cyclic AMP-induced Transcription from the Proximal Promoter (PII) of the Human Aromatase P450 (CYP19) Gene in the Ovary. J. Biol. Chem. 270 (22), 13561–13566. https://doi.org/10.1074/jbc.270.22.13561.
Muth-Köhne, E., Westphal-Settele, K., Brückner, J., Konradi, S., Schiller, V., Schäfers, C., Teigeler, M., Fenske, M., 2016. Linking the response of endocrine regulated genes to adverse effects on sex differentiation improves comprehension of aromatase inhibition in a Fish Sexual Development Test. Aquat. Toxicol. 176, 116–127. https://doi.org/10.1016/j.aquatox.2016.04.018.
Nakamoto, M., Shibata, Y., Ohno, K., Usami, T., Kamei, Y., Taniguchi, Y., Todo, T., Sakamoto, T., Young, G., Swanson, P., Naruse, K., Nagahama, Y., 2018. Ovarian aromatase loss-of-function mutant medaka undergo ovary degeneration and partial female-to-male sex reversal after puberty. Mol. Cell. Endocrinol. 460, 104–122. https://doi.org/10.1016/j.mce.2017.07.013.
Nascimento, A.R., Silva, G.F., Gualberto, G.F., Almeida, F.L., 2016. SHORT-COMMUNICATION Validation of reference genes for real-time quantitative PCR in tambaqui (*Colossoma macropomum*). Genet. Mol. Res. 15 (4) https://doi.org/10.4238/gmr15049228.
Navarro-Martín, L., Blázquez, M., Piferrer, F., 2009. Masculinization of the European sea bass (*Dicentrarchus labrax*) by treatment with an androgen or aromatase inhibitor involves different gene expression and has distinct lasting effects on maturation. Gen. Comp. Endocrinol. 160 (1), 3–11. https://doi.org/10.1016/j.ygcen.2008.10.012.
Pandit, N.P., Ranjan, R., Wagle, R., Yadav, A.K., Jaishi, N., Mahato, I.S., 2017. Use of tamoxifen, a non-steroidal aromatase inhibitor, on sex reversal of Nile tilapia (*Oreochromis niloticus*). Nepalese J. Aquacult. Fisheries 4, 7.
Peixe, B.R., 2022. Associação brasileira da piscicultura. Anuário PEIXE-BR da piscicultura 2021. PEIXE-BR, São Paulo, 136 p.
Qin, Z., Yang, F., Tian, L., Chen, R., Xu, D., Takeuchi, Y., 2020. Induction of sex reversal in blue drum (*Nibea mitsukurii*) and gynogenetic yellow drum (*Nibea albiflora*) by oral administration of letrozole. Aquac. Res. 51 (3), 882–889. https://doi.org/10.1111/are.14434.
Rajakumar, A., Senthilkumaran, B., 2020. Steroidogenesis and its regulation in teleost-a review. Fish Physiol. Biochem. 46 (3), 803–818. https://doi.org/10.1007/s10695-019-00752-0.
Renn, S.C.P., Aubin-Horth, N., Hofmann, H.A., 2008. Fish and chips: functional genomics of social plasticity in an African cichlid fish. J. Exp. Biol. 211 (18), 3041–3056. https://doi.org/10.1242/jeb.018242.
Ribas, L., Liew, W.C., Díaz, N., Sreenivasan, R., Orbán, L., Piferrer, F., 2017. Heat-induced masculinization in domesticated zebrafish is family-specific and yields a set of different gonadal transcriptomes. Proc. Natl. Acad. Sci. 114 (6), E941–E950. https://doi.org/10.1073/pnas.1609411114.
Scholz, S., Klüver, N., 2009. Effects of Endocrine Disrupters on Sexual, Gonadal Development in Fish. Sexual Develop. 3 (2–3), 136–151. https://doi.org/10.1159/000223078.
Schulz, R.W., de França, L.R., Lareyre, J.-J., LeGac, F., Chiarini-Garcia, H., Nobrega, R. H., Miura, T., 2010. Spermatogenesis in fish. Gen. Comp. Endocrinol. 165 (3), 390–411. https://doi.org/10.1016/j.ygcen.2009.02.013.
Shapiro, D.B., Pappalardo, A., White, B.A., Peluso, J.J., 1996. Steroidogenic factor-1 as a positive regulator of rat granulosa cell differentiation and a negative regulator of mitosis. Endocrinology 137 (4), 1187–1195. https://doi.org/10.1210/endo.137.4.8625888.
Shen, Z.-G., Fan, Q.-X., Yang, W., Zhang, Y.-L., Wang, H.-P., 2015. Effects of 17α-Methyltestosterone and Aromatase Inhibitor Letrozole on Sex Reversal, Gonadal Structure, and Growth in Yellow Catfish *Pelteobagrus fulvidraco*. Biol. Bull. 228 (2), 108–117. https://doi.org/10.1086/BBLv228n2p108.
Silva de Assis, H.C., Navarro-Martín, L., Fernandes, L.S.P., Cardoso, C.C., Pavoni, D.P., Trudeau, V.L., 2018. Cloning, partial sequencing and expression analysis of the neural form of P450 aromatase (cyp19a1b) in the South America catfish *Rhamdia quelen*. Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 221–222, 11–17. https://doi.org/10.1016/j.cbpb.2018.04.001.
Singh, R., Singh, A.K., Tripathi, M., 2012. Effect of a non-steroidal tamoxifen on the gonad and sex differentiation in Nile tilapia, *Oreochromis niloticus*. J. Environ. Biol. 33(4), 799–803. http://www.ncbi.nlm.nih.gov/pubmed/23360010.
Singh, A.K., Srivastava, P.P., Verma, R., Srivastava, S.C., Kumar, D., Ansari, A., 2015. Effect of dietary administration of letrozole and tamoxifen on gonadal development, sex differentiation and biochemical changes in common carp (*Cyprinus carpio* L.). Reprod. Fertil. Dev. 27 (3), 449. https://doi.org/10.1071/RD13234.
Sridevi, P., Chaitanya, R.K., Dutta-Gupta, A., Senthilkumaran, B., 2012. FTZ-F1 and FOXL2 up-regulate catfish brain aromatase gene transcription by specific binding to the promoter motifs. Biochim. Biophys. Acta (BBA) - Gene Regulatory Mech. 1819 (1), 57–66. https://doi.org/10.1016/j.bbagrm.2011.10.003.
Strobl-Mazzulla, P.H., Lethimonier, C., Gueguen, M.M., Karube, M., Fernandino, J.I., Yoshizaki, G., Patiño, R., Strüssmann, C.A., Kah, O., Somoza, G.M., 2008. Brain aromatase (Cyp19A2) and estrogen receptors, in larvae and adult pejerrey fish *Odontesthes bonariensis*: Neuroanatomical and functional relations. Gen. Comp. Endocrinol. 158 (2), 191–201. https://doi.org/10.1016/j.ygcen.2008.07.006.
Tanaka, M., Fukada, S., Matsuyama, M., Nagahama, Y., 1995. Structure and Promoter Analysis of the Cytochrome P-450 Aromatase Gene of the Teleost Fish, Medaka (*Oryzias latipes*). J. Biochem. 117 (4), 719–725. https://doi.org/10.1093/oxfordjournals.jbchem.a124768.
Tchoudakova, A., Kishida, M., Wood, E., Callard, G.V., 2001. Promoter characteristics of two cyp19 genes differentially expressed in the brain and ovary of teleost fish.

J. Steroid Biochem. Mol. Biol. 78 (5), 427–439. https://doi.org/10.1016/S0960-0760(01)00120-0.

von Hofsten, J., Olsson, P.-E., 2005. Zebrafish sex determination and differentiation: involvement of FTZ-F1 genes. Reproduct. Biol. Endocrinol.: RB&E, 3, 63. https://www.google.com/search?client=firefox-b-d&q=10.1186%2F1477-7827-3-63.

Watanabe, M., Tanaka, M., Kobayashi, D., Yoshiura, Y., Oba, Y., Nagahama, Y., 1999. Medaka (*Oryzias latipes*) FTZ-F1 potentially regulates the transcription of P-450 aromatase in ovarian follicles: cDNA cloning and functional characterization. Mol. Cell. Endocrinol. 149 (1–2), 221–228. https://doi.org/10.1016/S0303-7207(99)00006-4.

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 (9), 3030–3041. https://doi.org/10.1210/en.2016-1865.

Yoshiura, Y., Senthilkumaran, B., Watanabe, M., Oba, Y., Kobayashi, T., Nagahama, Y., 2003. Synergistic Expression of Ad4BP/SF-1 and Cytochrome P-450 Aromatase (Ovarian Type) in the Ovary of Nile Tilapia, *Oreochromis niloticus*, During Vitellogenesis Suggests Transcriptional Interaction1. Biol. Reprod. 68 (5), 1545–1553. https://doi.org/10.1095/biolreprod.102.010843.

Zhang, X., Li, M., Ma, H., Liu, X., Shi, H., Li, M., Wang, D., 2017. Mutation of foxl2 or cyp19a1a results in female to male sex reversal in XX Nile tilapia. Endocrinology. https://doi.org/10.1210/en.2017-00127.