



OPEN ACCESS

EDITED BY

Xu Wang,
Auburn University, United States

REVIEWED BY

Qing-Ping Xie,
Zhejiang Academy of Agricultural Sciences,
China
Xiaoli Ma,
Jiangsu Normal University, China
Zhiyuan Ruan,
Merck, United States

*CORRESPONDENCE

Chao Bian,
✉ bianchao@szu.edu.cn
Qiong Shi,
✉ shiqiong@szu.edu.cn,
✉ shiqiong@genomics.cn

†These authors have contributed equally to this work

RECEIVED 05 February 2024

ACCEPTED 03 April 2024

PUBLISHED 11 April 2024

CITATION

Zhao C, Bian C, Mu X, Zhang X and Shi Q (2024),
Gonadal transcriptome sequencing reveals
sexual dimorphism in expression profiling of
sex-related genes in Asian arowana
(*Scleropages formosus*).
Front. Genet. 15:1381832.
doi: 10.3389/fgene.2024.1381832

COPYRIGHT

© 2024 Zhao, Bian, Mu, Zhang and Shi. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Gonadal transcriptome sequencing reveals sexual dimorphism in expression profiling of sex-related genes in Asian arowana (*Scleropages formosus*)

Chenxi Zhao^{1,2†}, Chao Bian^{2,3*†}, Xidong Mu⁴, Xinhui Zhang^{1,2} and Qiong Shi^{1,2,3*}

¹College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, ²Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, Shenzhen, China, ³Laboratory of Aquatic Genomics, College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China, ⁴Key Laboratory of Prevention and Control for Aquatic Invasive Alien Species, Ministry of Agriculture and Rural Affairs, Guangdong Modern Recreational Fisheries Engineering Technology Center, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, China

Asia arowana (*Scleropages formosus*) is an ornamental fish with high economic value, while its sex determination mechanism is still poorly understood. By far, no morphological evidence or molecular marker has been developed for effective distinguishment of genders, which poses a critical challenge to our captive breeding efforts. In this study, we sequenced gonadal transcriptomes of adult Asian arowanas and revealed differential expression profiling of sex-related genes. Based on the comparative transcriptomics analysis of testes (n = 3) and ovaries (n = 3), we identified a total of 8,872 differentially expressed genes (DEGs) and 18,490 differentially expressed transposable elements (TEs) between male and female individuals. Interestingly, the expression of TEs usually has been more significantly testis-biased than related coding genes. As expected, several genes related to females (such as *foxl2* and *cyp19a1a*) are significantly transcribed in the ovary, and some genes related to male gonad development (such as *dmrt1*, *gsdf* and *amh*) are highly expressed in the testis. This sexual dimorphism is valuable for ascertaining the differential expression patterns of sex-related genes and enriching the genetic resources of this economically important species. These valuable genetic materials thereby provide instructive references for gender identification and one-to-one breeding practices so as to expand fish numbers for a rapid elevation of economic value.

KEYWORDS

Asia arowana, gonad transcriptome, sex-biased, *foxl2*, *dmrt* gene family

Introduction

Teleost exhibit remarkable diversity of sex determination mechanisms, including genetic sex determination (GSD; Myosho et al., 2012), environmental sex determination (ESD; Li et al., 2023), and various combinations of these two modes. In the GSD system (such as XX/XY and ZW/ZZ), the expression of sex-determining genes initiates a series of

cascadic signaling pathways of sex determination and differentiation, thereby inducing the primordial gonads to develop into ovaries or testes (Li and Gui, 2018). Many master sex-determining genes or sex differentiation-related genes have been reported to play a main role in regulation of sex development in various fishes, such as *dmrt1* (doublesex and mab-3 related transcription factor 1), *sox9* (SRY-box transcription factor 9), and *foxl2* (forkhead box protein L2). Some transcription factors, involved in regulating gene expression, play a crucial role in the development of male or female differentiation (Cocquet et al., 2002; Cui et al., 2017; Hu et al., 2021).

In fish, *dmrt1* is the master sex-determining gene in half-smooth tongue sole (Cui et al., 2017), and its homologous gene *dmy* (the *dmrt1* paralogous located in Y-chromosome) was verified to be the sex-determining gene in medaka (Nanda et al., 2002; Kobayashi et al., 2004). Nearly all polypeptide chains encoded by *dmrt* genes contain a highly conserved zinc finger DNA-binding and transcription regulating motif (known as DM domain), which consists of six conserved cysteines and two histidines for binding into the minor groove of any target DNA. High expression of *dmrt1* in males activates genetic programs to promote testicular differentiation. There is only one copy of the *dmrt2* gene in mammals, but two copies (*dmrt2a* and *dmrt2b*) are commonly identified in teleost fishes. It has been reported that the *dmrt2* (a/b) genes are involved in determination of the body axis and segment differentiation of vertebrates and are also related to the gonadal sex differentiation process (Han et al., 2021). In bony fishes, *dmrt3* is usually highly expressed in the testis and nervous system, which presumably plays an important role in the differentiation and development of nerve and germ cells as well as testes (Li et al., 2008; Dong et al., 2010). In contrast, in females, high expression of *foxl2* inhibits the expression of *dmrt1* while upregulates *cyp19a1*, thereby promoting the development of ovarian tissues (Zhang et al., 2017). The transforming growth factor- β (TGF- β) signaling pathway mainly participates in mediating the formation of tissues and organs as well as reproductive development by regulating cell growth, proliferation, differentiation and other processes (Li and Gui, 2018; Pan et al., 2021). Many previous studies have proved that this signaling pathway is related to sex determination and differentiation in various fishes. Meanwhile, some members of the TGF- β superfamily have been considered as initiators or key regulators of sexual differentiation, such as *amh* (anti-Müllerian hormone), *amhr2* (anti-Müllerian hormone receptor type 2; (Kamiya et al., 2012)), and *gsdf* (gonadal soma-derived factor; (Myosho et al., 2012)).

Transposable elements (TEs) have been integral parts of vertebrate genomes throughout the evolution process and play important roles in promoting genetic innovation, sex determination and reproduction (Biémont and Vieira, 2006). Recent bursts of TEs and simple repeat accumulations were observed around young sex determination loci, indicating the strong association between TE amplification and the formation of sex determination regions on sex chromosomes (Chalopin et al., 2015). TEs could carry regulatory elements and modify the expression of neighboring genes, thereby participating in the evolution of germ cells and gonadal regulatory networks. The expression and transposition of TEs are critical for vertical transmission to progeny and persistence in lineages (Dechaud

et al., 2019). Interestingly, TEs are particularly prone to being recruited in sexual development, since they are typically expressed in the gonads (Brunet et al., 2018). For instance, in *Oryzias latipes* approximately 1.2% and 3.5% of ovarian and testicular transcripts were reported to be attributed to TE expression, respectively (Dechaud et al., 2019).

Asian arowana (*Scleropages formosus*, also known as dragonfish) belongs to the ancient family Osteoglossidae, and it is a freshwater species with a native distribution in Southeast Asia. It has become a well-known ornamental fish for its vibrant colors and cultural significance. Mainly due to habitat destruction and overfishing, arowanas have been classified as endangered species in Appendix I by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). In recent years, significant progress has been made in the genetics research and breeding techniques of arowanas (Chang, 2009; Mu et al., 2012; Shen et al., 2014; Austin et al., 2015; Bian et al., 2016; Mu et al., 2022).

Asian arowana often reaches sexual maturity at 2–4 years of age (Chang, 2009), and several previous studies proposed that the genetic sex determination system of this species may be ZW/ZZ (Shen et al., 2014; Bian et al., 2016). Because there are no obvious morphological differences during all life stages of the Asian arowana, even after sexual maturity, it is very difficult to phenotypically differentiate the sex of individuals. One recent report showed that its putative ZW system is more likely a polymorphic pattern that occurs in the 18th chromosome pair, including the accumulation of constitutive heterochromatin and 18S rDNA (Toma et al., 2023). Meanwhile, although its genomes of both sexes have been published by us (Mu et al., 2022), our in-depth comparative genomics between males and females has failed to identify any significant region with obvious genomic differences. Therefore, we performed this study to identify coding and non-coding regions that are differentially expressed between ovaries and testes of adult Asian arowanas through transcriptome sequencing, which may become differential signatures between both sexes for practical gender identification and one-to-one captive breeding.

Materials and methods

Sample collection

Three female and three male individuals of Asian arowana, around 5-year-old, were obtained from Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences (Guangzhou, Guangdong, China). After dissection, the sex of these fishes was determined based on morphological inspection of the gonads. A total of six gonad tissues, including three ovaries and three testes, were frozen in liquid nitrogen immediately and stored at -80°C until use.

RNA isolation, library construction, and transcriptome sequencing

Total RNAs were extracted from the ovary and testis tissues by using a TRIZOL Kit (Invitrogen, Carlsbad, CA, United States of

America) following the manufacturer's instructions. The extracted RNAs were then digested by DNase I to eliminate genomic DNA's contamination. Purified RNA integrity and quality were assessed with an Agilent 2,100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA, United States of America). Only those RNA samples with RIN (RNA integrity number) > 7.0 were utilized for library construction. A total of six cDNA libraries with insert sizes of 300–400 bp were generated in DNA nanoballs (DNBs) according to the manufacturer's protocol of DNBSEQ sequencing platform, and then sequenced on a MGISEQ-2000 platform (MGI, BGI Shenzhen, China) to obtain 150-bp paired-end reads.

Screening and functional analysis of sex-biased differentially expressed genes (DEGs)

Raw reads with low quality, adapter sequences, and/or highly unknown N bases were filtered by using SOAPnuke v1.5.6 with optimized parameters “filter -n 0.01 -L 15 -q 0.4 -G -Q 2” (Chen et al., 2017). The clean RNA reads from six samples were subsequently aligned onto the previously published female genome assembly (Mu et al., 2022) by using STAR v2.5.3 (Dobin et al., 2013), and then RSEM v1.2.8 (Li and Dewey, 2011) was applied to quantify transcription levels of genes and transcripts. The Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) algorithms were used to normalize the mRNA expression levels. Based on the quantitative data, we utilized the DESeq2 R package (Love et al., 2014) to characterize DEGs between ovaries and testes. Adjusted *p*-value (false discovery rate, FDR) < 0.05 and absolute value of \log_2 (fold change) > 1.5 were assigned as the stringent threshold for significant DEGs (Stachowiak et al., 2024).

Functional annotations of the DEGs were performed via the NCBI NR database, Gene Ontology (GO) (Ashburner et al., 2000) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) pathway enrichment analysis. In addition, we compared all genes to the AnimalTFDB2.0 (Zhang et al., 2015) database to obtain transcription factor (TF) families. We subsequently applied DIAMOND (Buchfink et al., 2015) to align the DEGs with the sequences of Asia arowana in the STRING v12.0 database (Szklarczyk et al., 2023), constructed a potential protein-protein interaction (PPI) network among potential sex-related genes, and visualized the results using Cytoscape v3.10.1 (Su et al., 2014).

Quantification of TE expression

To analyze the expression of TEs, we applied Tetranscripts v2.2.3 (Jin et al., 2015) to estimate TE expression at a copy-level resolution by using sorted bam files from STAR. TE quantification was performed using RepeatMasker (Tarailo-Graovac and Chen, 2009) TE annotation. TE transcripts were used for a differential expression analysis between ovary and testis tissues with DESeq2 (Love et al., 2014) to generate a normalized count matrix, variance-stable count matrix, and matrix of differential gene expression. We further classified TE categories into LTRs, LINES, SINES, DNA TEs, and unknowns as provided in the published TE annotation of arowana (Mu et al., 2022). The numbers of differentially expressed TEs by sex within each category were reported for ovaries and testes.

To characterize the distributions of sex-biased genes and TEs across the arowana genome, we examined whether sex-biased genes clustered on the genome according to a published pipeline (https://gitlab.com/Coreend/gene_clusters_pyth) (Dechaud et al., 2021; Toubiana et al., 2021), which calculated the local average log₂FC of transcripts in a sliding window of 1 Mb with a 50-kb step size and applied a bootstrap method to detect regions of significant deviation (*p* < 0.05). We then counted the TE copies expressed in the testes and ovaries according to 500-kb windows and visualized the distributions of co-expression clusters and sex-biased TEs by RectChr (v1.36; <https://github.com/BGI-shenzhen/RectChr>).

Identification of *dmrt* family genes for phylogenetic analysis

We applied two strategies to obtain the protein sequences of *dmrt* family genes in one mammal (*Homo sapiens*, Hs), one cartilaginous fish (*Amblyraja radiata*, Ara), 18 representative teleost fishes, including *Danio rerio* (Dre), *O. latipes* (Ola), *Gasterosteus aculeatus* (Gac), *Lepisosteus oculatus* (Loc), *Albula glossodonta* (Agl), *Anguilla japonica* (Aja), *Arapaima gigas* (Agi), *Clarias batrachus* (Cba), *Clupea harengus* (Cha), *Gambusia affinis* (Gaf), *Heterotis niloticus* (Hni), *Ictalurus punctatus* (Ipu), *Mastacembelus armatus* (Mar), *Megalops atlanticus* (Mat), *Megalops cyprinoides* (Mcy), *Oreochromis aureus* (Oau), *Scophthalmus maximus* (Sma) and arowana (Sfo) (Supplementary Table S1).

For those species with public annotations, we downloaded related gene sequences from the NCBI (Supplementary Table S2) or extracted the sequences through BLASTP (Ye et al., 2006) for using as the reference sequences. Then we employed GeneWise (Birney et al., 2004) to predict related protein-coding sequences in the arowana genome using the reference protein sequences, and obtained neighboring genes from the genome annotation or using BLAST with an E-value of 1e-5 against the arowana genome in order to validate the synteny of *dmrt* genes. We converted coding sequences (CDS) to protein sequences and used MUSCLE v3.8 (Edgar, 2004) to perform global alignments. RaxML (Stamatakis, 2006) was employed to construct a gene-family phylogenetic tree with the PROTGAMMAAUTO model. We also searched the domains of these protein sequences by using NCBI Batch CD-Search and generated visualizations via iTOL (Letunic and Bork, 2021) and IBS 2.0 (Xie et al., 2022).

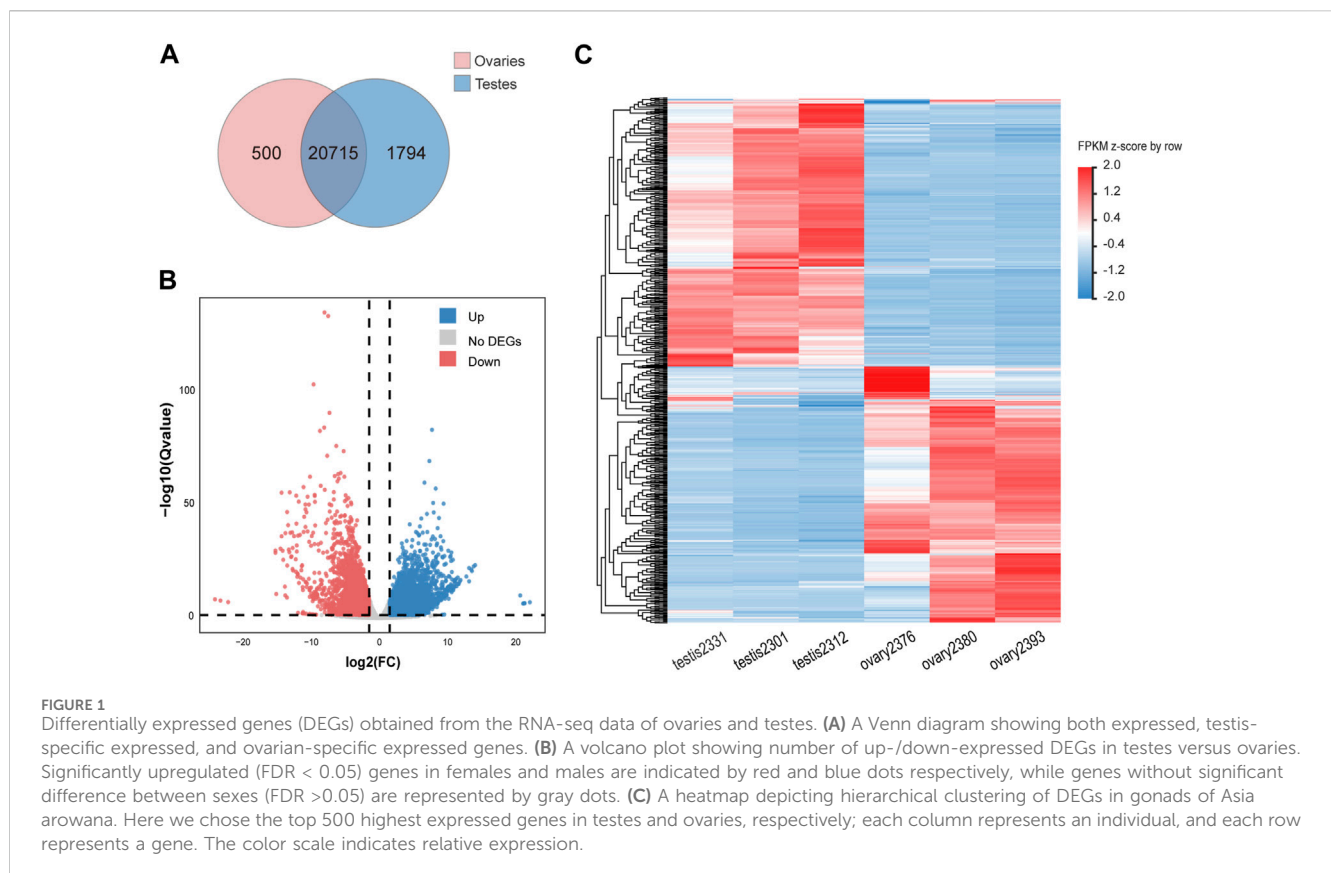
Results

Summary of the sequencing data and quality analysis

Transcriptome sequencing (RNA-Seq) of the six libraries, constructed in triplicates from ovaries and testes, yielded 32.16 Gb and 29.76 Gb of raw reads, respectively. After data filtering, we obtained 29.43 Gb of clean reads for three ovarian samples and 27.72 Gb of clean reads for three testis samples. In addition, the average percentages of bases with quality values greater than 20 (Q20) and 30 (Q30) in the six samples accounted for 97.51% and 90.45% of the total bases, respectively. The average mapping

TABLE 1 Summary of the sequencing reads of gonad transcriptomes in Asia arowana.

Sample	Total raw reads (Mb)	Total raw bases (Gb)	Total clean reads (Mb)	Total clean bases (Gb)	Clean reads Q20(%)	Clean reads Q30(%)	Total mapping (%)
ovary2376	42.65	5.86	39.08	5.86	97.76	91.32	93.01
ovary2380	86.31	11.94	79.59	11.94	97.32	90.05	97.19
ovary2393	84.79	11.63	77.51	11.63	97.04	89.16	97.53
testis2331	76.56	10.69	71.28	10.69	97.54	90.43	94.97
testis2301	65.57	9.18	61.17	9.18	97.55	90.41	94.48
testis2312	55.82	7.85	52.32	7.85	97.82	91.31	95.13



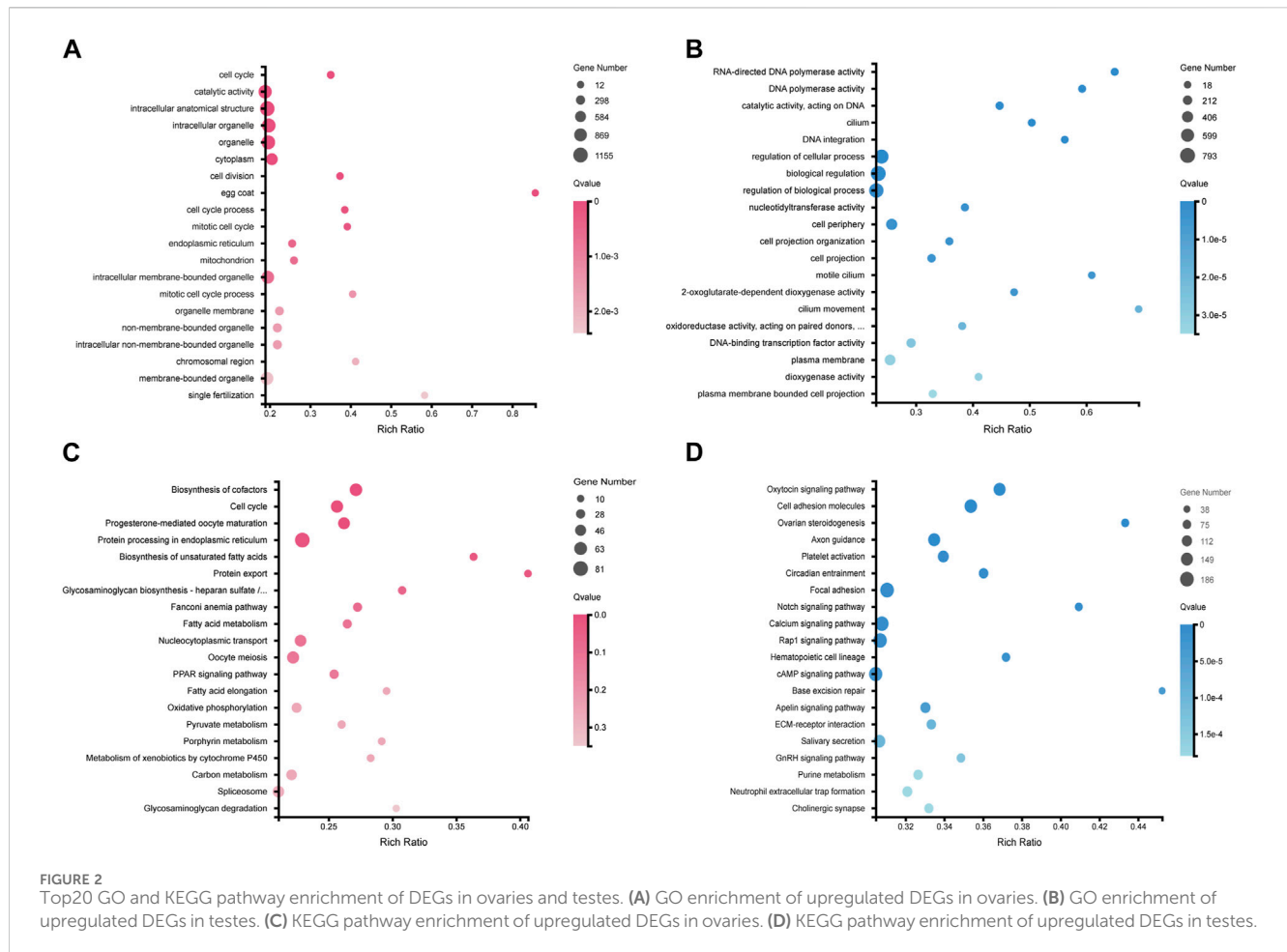
rates of transcriptome reads were 95.91% per ovary sample and 94.86% per testis sample, independently (Table 1).

Identification and enrichment of DEGs in ovaries and testes

A total of 23,009 expressed genes were detected in our research. Among the expressed genes, a total of 20,715 genes are expressed in both males and females, while 500 and 1,794 genes are specifically expressed in males and females respectively (Figure 1A). In total, 8,872 expressed genes (38.56%) were found to be sex-biased, including 5,153 (22.40%) upregulated DEGs and 3,719 (16.16%)

downregulated DEGs in testes compared with ovaries (Figures 1B, C). The remaining 14,137 genes were expressed without significant difference between genders.

GO functional annotation and KEGG pathway annotation were further performed on the DEGs. A total of 5,651 DEGs (2,612 female-biased genes and 3,039 male-biased genes) were assigned to GO terms (Supplementary Figure S1), and 3,479 DEGs were annotated to KEGG pathways (Supplementary Figure S2). GO enrichment analysis showed that ovary-biased DEGs were enriched in GO terms such as cell cycle (GO:0007049), egg coat (GO:0035805) and single fertilization (GO:0007338) (Figure 2A; Supplementary Table S3), while testis-biased DEGs were enriched in GO terms such as RNA-directed DNA polymerase activity (GO:



0003964), motile cilium (GO:0031514) and regulation of cellular process (GO:0050794) (Figure 2B; Supplementary Table S4). Through KEGG pathway enrichment analysis, we observed that upregulated DEGs in ovaries were significantly enriched in cell cycle (ko04110), oocyte meiosis (ko04114), progesterone-mediated oocyte maturation pathways (ko04914) and TGF-β signaling pathway (ko04350) (Figure 2C; Supplementary Table S5), while upregulated DEGs in testes were enriched in some pathways including oxytocin signaling pathway (ko04921), focal adhesion (ko04510), cell adhesion molecules (ko04514), calcium signaling pathway (ko04020) (Figure 2D; Supplementary Table S6).

Expression patterns of sex determination and sexual differentiation related genes

Based on the functional annotations results of NCBI NR database as well as KEGG Orthology and enrichment analyses, we performed a further comparative analysis to detect that some well-known candidate sex determination or sexual differentiation-related genes and members of the TGF-β signaling pathway showed significant sexual dimorphism in gonadal expression. Those genes related to testis differentiation as well as sperm maturation and maintenance of teleost, such as *dmrt1*, *dmrtB1*, *sox9*, *cyp11a* (cholesterol side-chain cleavage enzyme), *cyp17a* (steroid 17-α-hydroxylase/17,20 lyase-like), *amh* and its receptor

amhr2 were detected with high transcription in testes; *foxl2*, *figla* (factor in the germline alpha), *cyp19a1* (aromatase), and *gdf9*, related to ovary differentiation, were significantly expressed in ovaries compared with testes (Figure 3A). Several members of *wnt* family also showed sex-biased in gonads. For example, *wnt4a* and *wnt11* showed a female bias, but *wnt5b*, *wnt6*, *wnt7b*, *wnt8b*, and *wnt10b* upregulated in testes; however, *wnt4b* was expressed without differential difference in these samples (Table 2; Figure 3A).

The STRING database was utilized to construct a PPI network for potential sex-related genes among the DEGs. It was revealed that 31 potential sex-related genes exhibited intricate interaction relationships. Notably, *dmrt1* and *amh* were associated with 15 sex-related genes, *cyp19a1* collaborated with 13 genes, and *Foxl2* interacted with five genes. *Dmrt1* not only interacted closely with testis-upregulated genes (such as *amh*, *sox9*, *dmrtB1*, and *rspo1*), but also with female sex-related genes (like *wnt4a*, *foxl2*, *figla*, and *cyp19a1*). Furthermore, *foxl2* interacted with its downstream functional gene *cyp19a1*, as well as with *dmrt1*, *amh*, and *amhr2* (Figures 3B, C).

Identification of TE copies and families with sex-biased expression

We characterized gonadal expression of TEs in order to analyze TE expression relative to gene expression. Unlike protein-coding

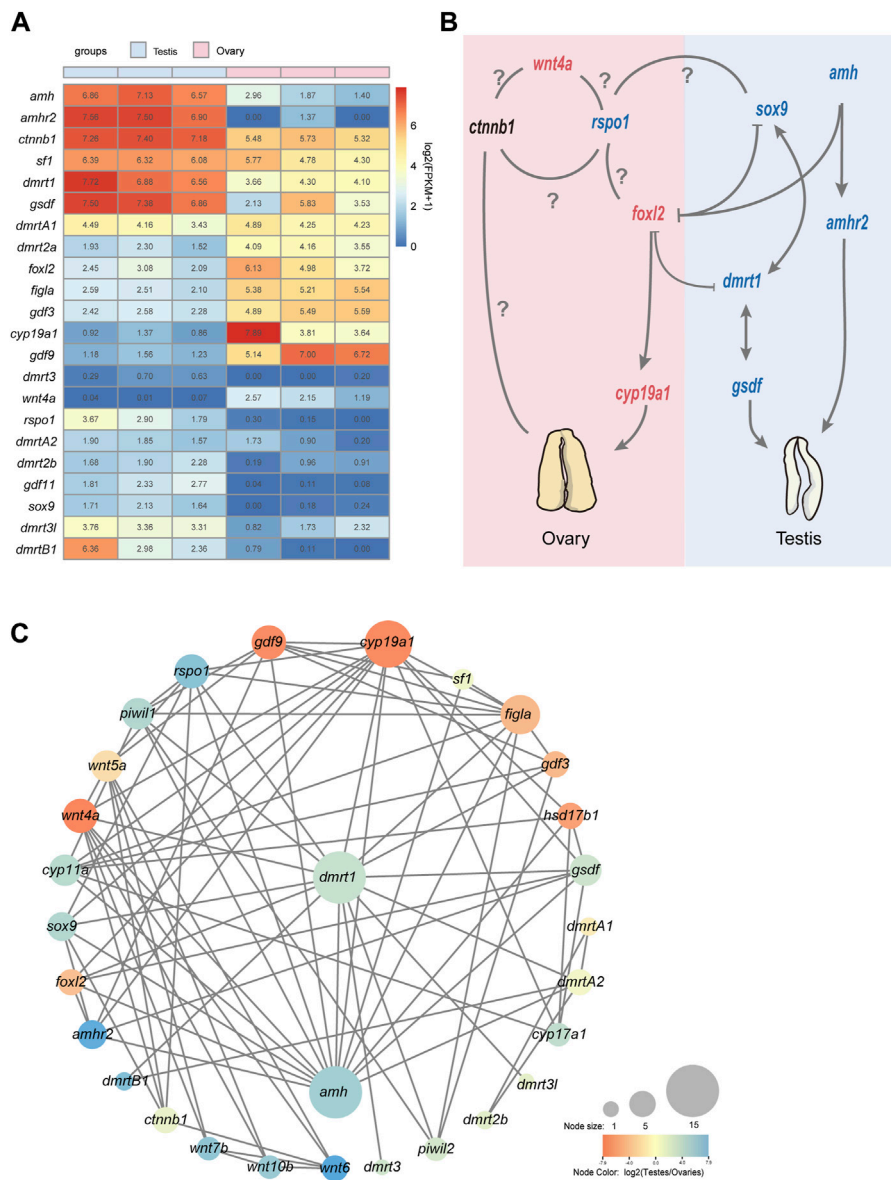


FIGURE 3 An overview of sex-related genes and a predicted regulatory network in Asia arowana. **(A)** The expression profiles of sex determination or differentiation-related genes in gonads of Asia arowana. **(B)** A predicted regulatory network for sex development in Asia arowana. Genes with upregulated expression in females and males are marked in red and blue, respectively. Note that *dmrt1* and *amh* are the key male sex regulators, while *foxl2* and *cyp19a1a* can induce female development. Symbols: ⊥: inhibition; →: stimulation; ?: unclearly. **(C)** PPI network analysis of sex determination or differentiation-related genes. Each node size reflects the number of interacting proteins, and various colors represent the differential expression profiles.

genes that have an approximately equal proportions of expression in both sexes, TEs appeared to have a significant testicular bias. In brief, 18,490 TE copies from the transcriptomes were detected with significantly differential expression in the gonads, of which 15,930 (86.15%) were testis-biased, while 2,560 (13.84%) were ovarian biased. We classified these testis- or ovary-biased TEs into five subfamilies by TE annotation subsequently (Figure 4A). Interestingly, we observed that sex-biased expressed genes often clustered in a genome-wide manner. And in some chromosomes, TE copies located near ovary-biased gene regions are still showing testis expression (such as the testis-biased TE is located between the two copies of the ovary-biased *rapunzel-like* on Chr23; Figure 4B, C-2),

which is consistent with the trend that the overall expression of TEs in the genome is more testis-biased (Figure 4A). Nevertheless, we also found some TE copies expressed near female related genes, such as an expressed *LTR retrotransposon* and an expressed DNA transposon at the upstream of the *foxl2* (Figure 4C-1).

Summary of *dmrt* family genes and related phylogenetic tree

Since multiple copies of *dmrt* family genes were detected by the transcriptome sequencing, we scanned the *dmrt* gene family

TABLE 2 Summary of sex determination and sex differentiation associated genes in Asia arowana.

Gene name	Gene description	log ₂ (Testes/Ovaries)	FDR	Sex-bias
<i>hsd17b7</i>	3-keto-steroid reductase/17-beta-hydroxysteroid dehydrogenase 7	1.64	4.66E-03	Male
<i>piwil2</i>	piwi-like protein 2	1.81	1.92E-04	Male
<i>gsdf</i>	growth differentiation factor 6	2.23	2.24E-03	Male
<i>dmrt1</i>	doublesex- and mab-3-related transcription factor 1	2.50	1.41E-09	Male
<i>wnt5b</i>	wingless-type MMTV integration site family, member 5b	2.76	3.25E-02	Male
<i>cyp17a1</i>	steroid 17-alpha-hydroxylase/17,20 lyase	2.77	6.04E-04	Male
<i>cyp11a</i>	cholesterol side-chain cleavage enzyme	3.14	2.97E-05	Male
<i>hsd17b3</i>	testosterone 17-beta-dehydrogenase 3	3.42	2.15E-02	Male
<i>piwil1</i>	piwi-like protein 1	3.46	1.01E-09	Male
<i>sox9</i>	transcription factor SOX-9a	3.46	5.38E-04	Male
<i>sox2</i>	transcription factor SOX-2	3.82	7.20E-02	Male
<i>amh</i>	muellerian-inhibiting factor	4.21	3.31E-17	Male
<i>wnt10b</i>	wingless-type MMTV integration site family, member 10b	4.64	7.22E-02	Male
<i>wnt2b</i>	wingless-type MMTV integration site family, member 2b	4.65	4.21E-06	Male
<i>wnt8b</i>	wingless-type MMTV integration site family, member 8b	4.87	3.95E-02	Male
<i>wnt7b</i>	wingless-type MMTV integration site family, member 7b	5.01	1.62E-02	Male
<i>rspo1</i>	R-spondin-1	5.41	6.82E-10	Male
<i>dmrtB1</i>	doublesex- and mab-3-related transcription factor 6	5.72	1.31E-04	Male
<i>gdf11</i>	growth/differentiation factor 11	5.90	8.65E-11	Male
<i>amhr2</i>	anti-Muellerian hormone type-2 receptor	7.42	9.60E-04	Male
<i>wnt6</i>	wingless-type MMTV integration site family, member 6	7.95	1.33E-08	Male
<i>wnt4a</i>	wingless-type MMTV integration site family, member 4	-7.26	8.33E-17	Female
<i>cyp19a1</i>	aromatase	-6.97	2.93E-11	Female
<i>gdf9</i>	growth/differentiation factor 9	-6.82	1.48E-24	Female
<i>hsd17b1</i>	estradiol 17-beta-dehydrogenase 1	-5.61	1.58E-07	Female
<i>wnt11</i>	wingless-type MMTV integration site family, member 11	-5.46	1.00E-24	Female
<i>hsd17b12</i>	17beta-estradiol 17-dehydrogenase/very-long-chain 3-oxoacyl-CoA reductase	-5.09	9.02E-24	Female
<i>sox32</i>	transcription factor SOX32	-4.76	2.62E-02	Female
<i>figla</i>	factor in the germline alpha isoform	-4.16	2.38E-23	Female
<i>gdf3</i>	growth/differentiation factor 3	-4.13	2.81E-18	Female
<i>foxl2</i>	forkhead box protein L2	-3.67	3.24E-08	Female
<i>dmrt2a</i>	doublesex- and mab-3-related transcription factor 2a	-3.39	3.90E-23	Female
<i>wnt5a</i>	wingless-type MMTV integration site family, member 5a	-1.80	7.97E-03	Female
<i>hsd17b4</i>	peroxisomal multifunctional enzyme type 2	-1.79	7.72E-04	Female
<i>hsd17b10</i>	3-hydroxyacyl-CoA dehydrogenase type-2	-1.63	3.75E-03	Female
<i>gdf7</i>	growth/differentiation factor 7	-1.65	4.54E-01	-
<i>hsd17b8</i>	estradiol 17-beta-dehydrogenase 8	-0.98	8.73E-02	-
<i>dmrtA1</i>	doublesex- and mab-3-related transcription factor A1	-0.91	1.17E-02	-

(Continued on following page)

TABLE 2 (Continued).

Gene name	Gene description	log2(Testes/Ovaries)	FDR	Sex-bias
<i>foxl3</i>	forkhead box protein L3	-0.87	8.19E-01	-
<i>sox17</i>	transcription factor Sox-17	-0.82	4.37E-01	-
<i>wnt9a</i>	wingless-type MMTV integration site family, member 9a	-0.79	8.39E-01	-
<i>wnt16</i>	wingless-type MMTV integration site family, member 16	-0.59	6.29E-01	-
<i>sox11</i>	transcription factor SOX-11	0.08	9.49E-01	-
<i>sox5</i>	transcription factor SOX-5	0.43	6.09E-01	-
<i>sox7</i>	transcription factor Sox-7	0.43	8.24E-01	-
<i>dmrtA2</i>	doublesex- and mab-3-related transcription factor A2	0.43	7.21E-01	-
<i>sox6</i>	transcription factor SOX-6	0.48	5.19E-01	-
<i>sf1</i>	splicing factor 1	0.54	2.92E-01	-
<i>ctnnb1</i>	catenin beta-1	0.83	1.06E-03	-
<i>dmrt3</i>	doublesex- and mab-3-related transcription factor 3	2.02	3.89E-01	-
<i>dmrt3l</i>	doublesex- and mab-3-related transcription factor 3a-like	1.06	2.03E-01	-
<i>sox8</i>	transcription factor Sox-8	1.06	8.33E-01	-
<i>sox21</i>	transcription factor SOX-21	1.11	1.59E-01	-
<i>dmrt2b</i>	doublesex- and mab-3-related transcription factor 2b	1.16	2.53E-01	-
<i>gdf5</i>	growth/differentiation factor 5	1.34	4.51E-01	-
<i>sox13</i>	transcription factor SOX-13	1.39	1.42E-01	-
<i>wnt7a</i>	wingless-type MMTV integration site family, member 7a	1.84	4.21E-01	-
<i>sox3</i>	transcription factor Sox-3	3.20	3.49E-01	-

in many representative genomes of arowana and some other bony fishes. A total of nine members from the *dmrt* gene family were identified in arowana (Figure 5A) and named based on the nomenclature of teleost *dmrt* genes. All of these identified DMRT proteins contain a conserved DM (Doublesex and Mab-3) domain (near the N-terminus in Figure 5B). Meanwhile, Dmrt1 has an additional DMRT1 domain, as well as Dmrt3, Dmrt3L, DmrtA1 and DmrtA2 own a special DAM domain (middle in Figure 5B). Furthermore, we observed a somehow conserved *dmrt1-dmrt3-dmrt2* cluster in the examined 20 species (see the middle panel in Figure 5C), while those species of Osteoglossus and Elopomorpha have another *dmrt2-dmrt3-dmrt1* gene cluster (see the left panel in Figure 5C). Based on the neighboring genes of both gene clusters, we infer that the gene clusters of these two groups are *dmrt1-dmrt3l-dmrt2l* and *dmrt3-dmrt2a*, respectively (Figure 5C).

The expression patterns of *dmrt* genes were investigated among testes and ovaries. Obviously, the transcription levels of *dmrt1* and *dmrtB1* in the testes were significantly higher than those in ovaries, *dmrt3l* was upregulated in the testes, while *dmrt2l* was not detectable in transcription (Figure 3A). Surprisingly, in the *dmrt3-dmrt2a* gene cluster (middle panel of Figure 5C), the *dmrt3* is slightly expressed in the testis but undetectable in the ovary, while *dmrt2a* is highly transcribed in the ovary (Figure 3A).

Discussion

Transcriptomics is an effective method to obtain gene regulatory networks among different individuals. In this study, we employed the RNA-Seq technology for transcriptomics analysis, and many genes related to gonad development were identified and characterized to improve our understanding of their sex-related functions from a molecular perspective. Generally speaking, the fish sex determination mechanism usually involves complex biological processes, including a series of genes that promote or maintain the development of gonads into testes or ovaries. The sexual dimorphism in expression identified in this study, especially for those previously reported candidate genes that are related to sex determination and sexual differentiation (Kitano et al., 2023), may be effective indicators for sex prediction in Asian arowana.

The Wnt signaling is a crucial pathway to regulate cell proliferation, differentiation, embryonic development, and folliculogenesis. Recent genetic research has identified the *wnt4* as a vital regulator of the Wnt/ β -catenin signaling pathway during sex determination (Kossack et al., 2018; Farhadi et al., 2021). The most important member *wnt4* gene has been identified in zebrafish (Kossack et al., 2018), medaka (LI et al., 2012) and various other teleost fishes. Due to the teleost-specific whole-genome duplication (WGD) events, there are two paralogs in most teleost, *wnt4a* and

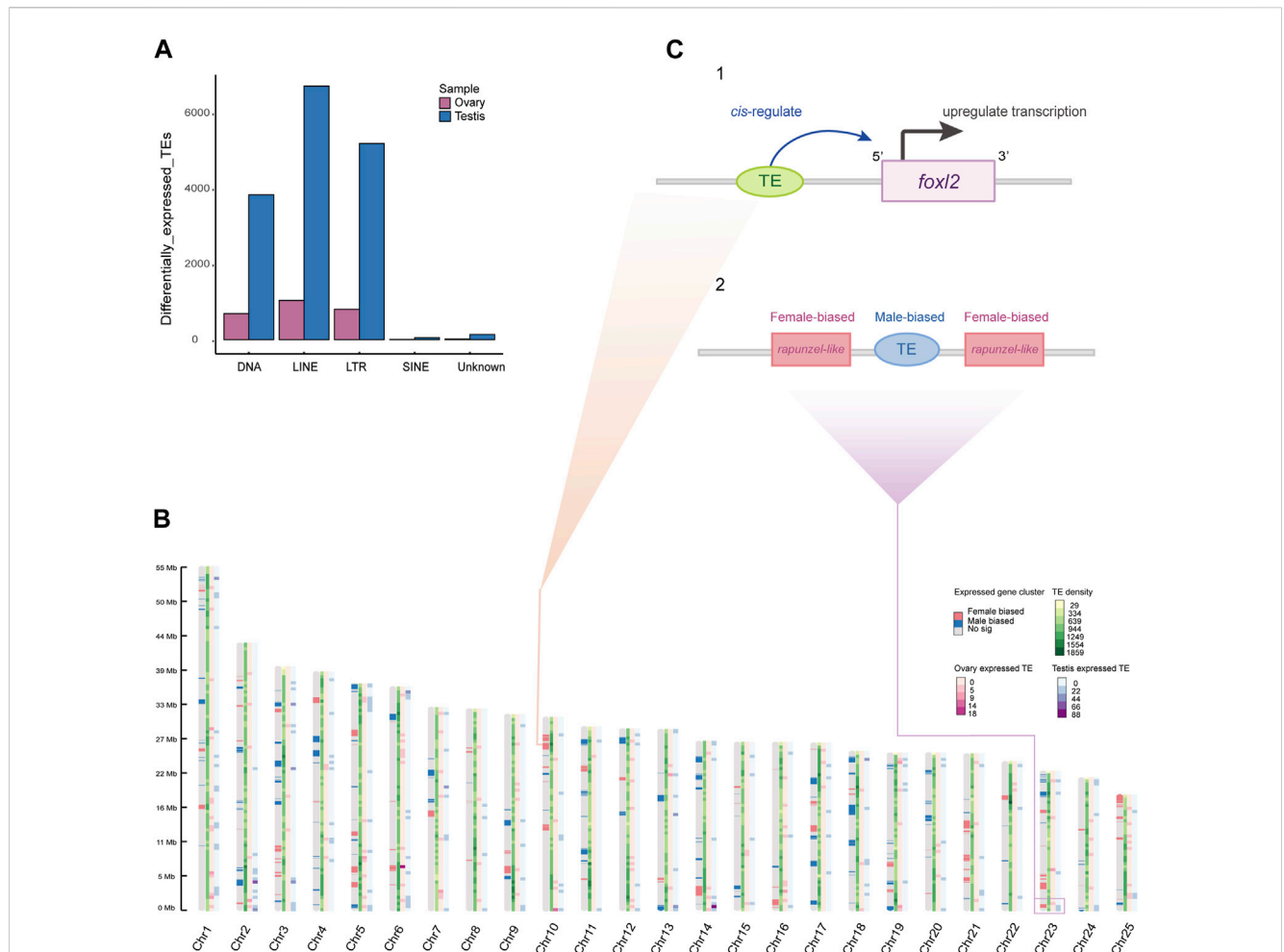
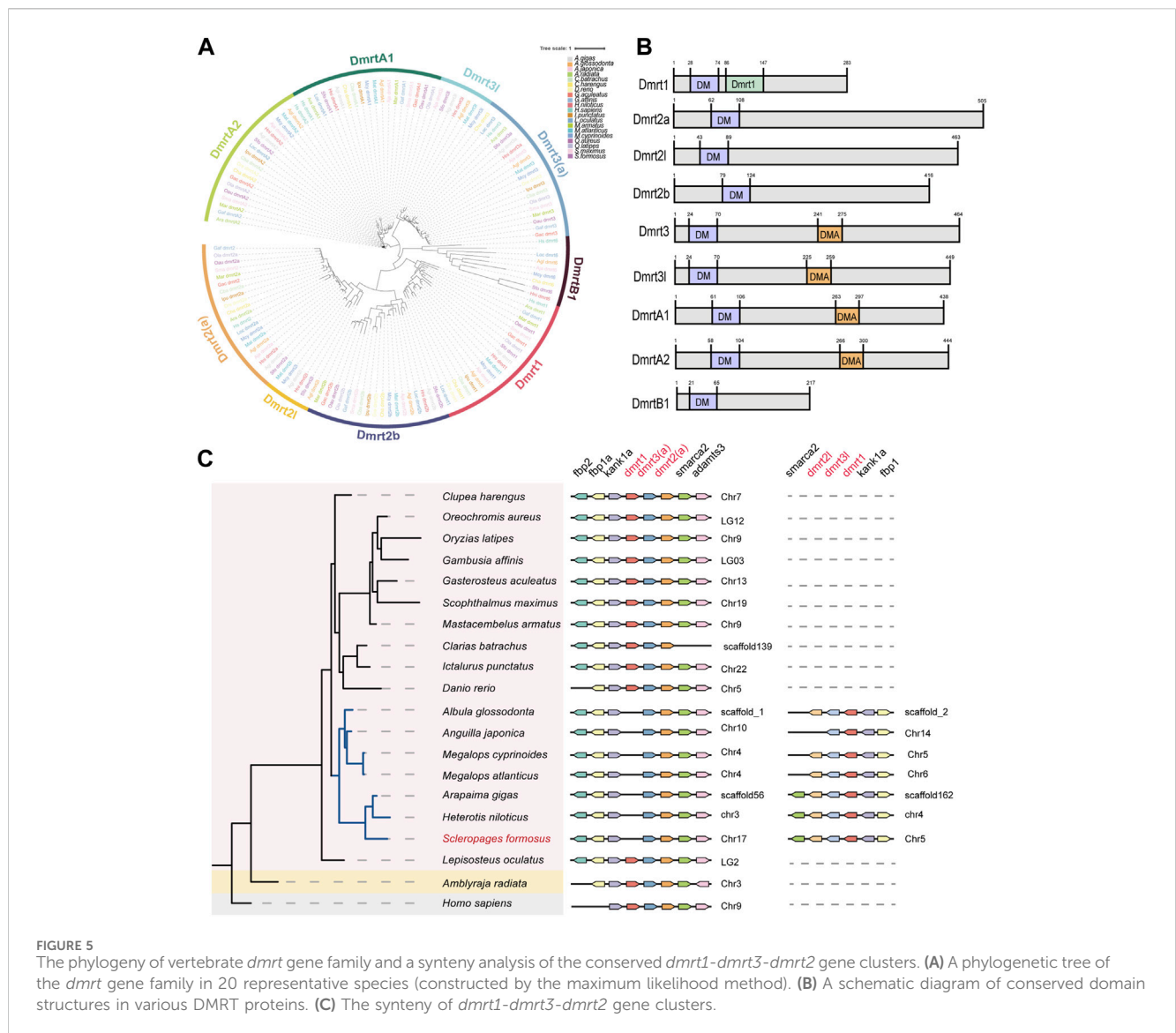


FIGURE 4 Sex-biased expressed TEs and gene expression clusters. **(A)** A bar plot of differentially expressed TEs in ovaries and testes (categorized by TE family). **(B)** Genome-wide sex-biased gene clusters and TEs density across sliding windows. The color pattern on each chromosome represents gene expression cluster, TE density, ovary-biased expressed TE density, and testis-biased expressed TE density, respectively. **(C)** The predicted regulative pattern of expressed TE at upstream of the *foxl2* (1). The testis-biased TE located in neighborhood of ovary-biased expressed genes (2).

wnt4b, both of which play essential roles in the development of female gonads, ovarian growth, early male development, and sex transition (Hu et al., 2014). We found that several *wnt* genes showed sex-biased expression patterns in arowana. For example, the expression of *wnt4a* ($\log_2FC = -7.26$) and *wnt11* ($\log_2FC = -5.46$) that are related to ovary differentiation were upregulated in ovaries; the expression of *wnt6* ($\log_2FC = 7.95$), which could initiate Wnt/ β -catenin signaling to control the proliferation of undifferentiated spermatogonia in mouse (Takase and Nusse, 2016), was also testis-biased. *rspo1* showed a sexually dimorphic expression pattern with significantly higher expression in testes over ovaries in our present study ($\log_2FC = 5.41$). It is a potential female-determining gene in mammals that can regulate the important Wnt/ β -catenin signaling pathway; loss of *rspo1* can lead to female sex reversal. However, the expression pattern of *rspo1* in fish seems not very conservative. For instance, in many teleost fishes such as zebrafish, medaka, and half-smooth tongue, *rspo1* shows a female bias, while in some ancient fishes (like *Acipenser stephensi*, coelacanth, and lungfish) and some sex-

changed teleost fishes (like *Notolabrus celidotus*) shows a male bias (Biscotti et al., 2018; Zhang et al., 2020; Muncaster et al., 2023), implying that *rspo1* may have other gender-related functions in various fishes.

The TGF- β signaling pathway is a large family comprising many members responsible for regulation of various processes including tissue and organ formation, reproductive development, cell growth, proliferation, and differentiation (Pan et al., 2021). In fish, this critical pathway often plays a key role in regulating important physiological processes, such as embryonic development, tissue regeneration, immune regulation, chondrogenesis, as well as sex determination and sexual differentiation (Pan et al., 2021; Kitano et al., 2023). Some members belonged to the TGF- β superfamily, such as *amh*, *amhr2*, and *gsdf* were highly expressed in males (Figure 3A), which indicates that these pathways should play a significant role in sexual differentiation of Asian arowana. During ovarian development, the interaction among granulosa cells, theca cells, and oocytes is crucial for follicular development, steroidogenesis, and female oogenesis (Richards, 2018). The



forkhead-box L2 (*foxl2*) transcription factor is one of the earliest known markers of granulosa cell differentiation (Cocquet et al., 2002), and it has been determined to be a major gene of sex determination and maintenance in various teleost (Cui et al., 2017; Zhang et al., 2017; Yuan et al., 2021). In addition, the *foxl2* can directly bind to the promoter region of the aromatase gene *cyp19a1a* to promote its expression, or indirectly regulated expression of *cyp19a1a* through interaction with *sfl* (splicing factor 1), resulting in estrogen (E2) production to promote ovarian differentiation (Wang et al., 2007; Zhang et al., 2017; Yuan et al., 2021). In our present study, *foxl2* (log2FC = -3.67) and *cyp19a1* (log2FC = -6.97) were found to be highly upregulated in the ovary compared with the testis (Figure 3A), showing a similar expression pattern to other fish species (Zhang et al., 2019; Mustapha et al., 2022). Through a PPI analysis, we found that most of the potential sex-related genes among the DEGs exhibit interaction relationships, suggesting that these genes may play functional roles in sex differentiation and development,

potentially being directly or indirectly involved in gonadal development or differentiation. Notably, *dmrt1*, *amh*, as well as *cyp19a1* (and its upstream regulator *foxl2*) interact with multiple detected potential sex-related genes, indicating that these gene play central regulatory roles in sex differentiation in Asia arowana. In mammals, *Dmrt1* directly represses *Foxl2* expression in the testis, and *Foxl2* is required for repressing *Dmrt1* expression in the ovary (Matson et al., 2011). Previous studies on tilapia and zebrafish found that *dmrt1* has an antagonistic effect on the expression of *foxl2* and *cyp19a1* (Wang et al., 2010). *Dmrt1* can directly repress the expression of *foxl2* and *cyp19a1* in somatic cells of tilapia, while in females *foxl2* activates the expression of *cyp19a1* and inhibits expression of *dmrt1*. In addition, studies have found that *foxl2* has ovary-enriched expression in zebrafish, just like in mammals, and *dmrt1* is required for its downregulation. Therefore, it is speculated that the antagonistic relationship between *foxl2* and *dmrt1* in the sex determination process may be conserved among vertebrates (Webster et al., 2017). We have observed that *dmrt1* not only

interacts with male-biased gene, such as *amh* and *sox9*, but also engages with *cyp19a1* and *foxl2*, suggesting that the antagonism may also exist in Asian arowana. However, due to limited research on Asian arowana, some protein-protein interactions may require further experimental validation (like *Dmrt2a*). According to the sex-biased DEGs detected in our present study and the sex relevant pathways in amphioxus and vertebrates (Pan et al., 2021; Curzon et al., 2023; Huang et al., 2023), we infer that *wnt4a*, *foxl2* and *cyp19a1* may play the major role in promoting ovarian differentiation of Asia arowana, while the expression and regulation of *sox9*, *dmrt1*, *gsdf*, *amh* and *amhr2* possibly play an important role in development and differentiation of testis. Based on this inference, a potential regulatory network of Asian arowana sexual development was proposed (Figure 3B).

It is well known that TEs could regulate gene expression (Feschotte, 2008; Sundaram and Wysocka, 2020; Gebrie, 2023), and previous studies reported that the closer TE copy to the target gene, the higher correlation of their expressions (Dechaud et al., 2021). In mammalian genomes, TEs are an important source of various cis-regulatory sequences; 20% of the cis-regulatory elements (CREs) in the human genome may have been taken from TEs (Sundaram and Wysocka, 2020), and TEs often contribute to zebrafish cis-regulatory elements, tissue-specific expression and alternative promoters (Lee et al., 2022). Researchers have found that the fourth exon (Ex4) of the sex-determining gene *dm-W* in the African clawed frog (*Xenopus laevis*) originated from a non-coding fragment of the *hAT-10* family of DNA transposons (Hayashi et al., 2022). An experimental evidence from sablefish (*Anoplopoma fimbria*) demonstrated that a TE insertion in the promoter region of *gsdfy* produced allelic diversification by bringing a cis-regulatory module, leading to transcriptional reprofiling and generating a new sex-determined gene for this species (Herpin et al., 2021). The regulatory elements of many genes contain TE sequences, which are involved in the regulation of gene expression. Some studies have reported that TEs can be expressed in multiple tissues (Brunet et al., 2018; Chang et al., 2022; Xu et al., 2023). However, because they align to multiple genomic loci, many RNA-seq reads derived from TEs are often discarded before data analysis. To resolve this problem, several computational tools have been developed (Jin et al., 2015; Jeong et al., 2017; Yang et al., 2019; Lanciano and Cristofari, 2020). We therefore observed a DNA transposon and an LTR transposon located at the promoter region of *foxl2* expressed in the arowana ovary (Figure 4C), suggesting co-regulation of the cis-regulatory elements to enhance the neighboring sex-biased gene(s) and/or sex-biased TEs themselves on these two types of sequences, as well as the *foxl2* may play a significant functional role in ovarian differentiation of arowana.

Dmrt1, an important member of the *dmrt* gene family, is essential for maintaining male-specific germ cells and testis differentiation (Matson and Zarkower, 2012; Cui et al., 2017). It not only participates in sex regulation in mammals, but also acts as a ubiquitous conserved sex regulation factor in other non-mammalian vertebrates including fishes. A recently published paper reported no orthologous gene of *dmrt1* in amphioxus (Huang et al., 2023), while previous studies on the *dmrt* family of diverse animals found that *dmrt1* and *dmrtB1* are vertebrate-specific genes, and they may have arisen in vertebrates through WGD events (Mawaribuchi et al.,

2019). Various studies on non-mammalian vertebrates have proved that *dmrt1* or its paralogous genes control gonadal sex determination and differentiation with different mechanisms (Smith et al., 2009; Yoshimoto et al., 2010; Cui et al., 2017; Mawaribuchi et al., 2019). From the TF annotation of arowana transcriptomes, we identified multiple copies of *dmrt*, such as *dmrt3* and *dmrt2a*. Through gene annotation and genome scanning of one elasmobranch and 18 bony fishes, we observed the conserved *dmrt1-dmrt3-dmrt2* gene cluster among various vertebrates, and some species of the order Osteoglossus and Elopomorpha also have another *dmrt3-dmrt2* gene cluster (Figure 5C). Previous studies have shown that *dmrt1* may have emerged after a WGD event in vertebrates, and based on our findings we speculate that the *dmrt* gene clusters in these two basal teleost species are likely to come from the teleost-specific WGD event (3R-WGD).

It seems that the expanded *dmrt* gene cluster remains in this clade for continuous function after the 3R-WGD. This confirms that the two groups are sister branch to each other, and this branch forms as sister groups to all the other teleost (Parey et al., 2023). The gene cluster identified in the Asian arowana was named as *dmrt1-dmrt3l-dmrt2l*, based on the names of related genes published in the NCBI, and another cluster is named as *dmrt3-dmrt2a* (Figure 5C). Interestingly, *dmrt1-dmrt3l* in the *dmrt1-dmrt3l-dmrt2l* gene cluster was upregulated in the arowana testis ($\log_2FC = 2.50$, $\log_2FC = 1.06$), while *dmrt2l* was undetectable in both testes and ovaries; in the *dmrt3-dmrt2a* gene cluster, the *dmrt3* gene is slightly expressed in the testes but without expression in the ovary, while *dmrt2a* is highly expressed in the ovary ($\log_2FC = -3.39$) (Figure 3A). These expression trends are very similar to those in other teleost fishes, indicating that *dmrt1* and *dmrt2* located in different clusters may be critical players in the sexual differentiation of Asian arowana: one of the gene clusters may be associated with male development, while another is potentially associated with female development.

Conclusion

Based on gonadal transcriptome sequencing and transcriptomic comparisons, we characterized a set of differentially expressed genes and pathways for potential involvement in sex determination or differentiation in Asia arowana. This sexual dimorphism is valuable for ascertaining the differential expression patterns of sex-related genes and enriching the genetic resources of this ornamental fish species. Our transcriptome data on sex-related genes can also promote the exploration of molecular mechanisms of gonadal development and sex determination in Asian arowana, which provides valuable references for practical gender identification and one-to-one breeding programs so as to expand fish number for a rapid elevation of worldwide economic value.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession

number(s) can be found below: NCBI under accession numbers: SRR20631749-SRR20631754.

was supported by National Key Research and Development Program of China (no. 2022YFE0139700).

Ethics statement

The animal study was approved by the Animal Care and Use Committee of BGI. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

CZ: Writing—original draft, Writing—review and editing, Formal Analysis, Methodology. CB: Writing—original draft, Writing—review and editing, Conceptualization, Methodology. XM: Investigation, Resources, Writing—review and editing, Conceptualization, Supervision. XZ: Methodology, Writing—review and editing. QS: Conceptualization, Writing—original draft, Writing—review and editing, Supervision.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1381832/full#supplementary-material>

References

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556
- Austin, C. M., Tan, M. H., Croft, L. J., Hammer, M. P., and Gan, H. M. (2015). Whole genome sequencing of the Asian arowana (*Scleropages formosus*) provides insights into the evolution of ray-finned fishes. *Genome Biol. Evol.* 7 (10), 2885–2895. doi:10.1093/gbe/evv186
- Bian, C., Hu, Y., Ravi, V., Kuznetsova, I. S., Shen, X., Mu, X., et al. (2016). The Asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts. *Sci. Rep.* 6 (1), 24501. doi:10.1038/srep24501
- Biémont, C., and Vieira, C. (2006). Genetics: junk DNA as an evolutionary force. *Nature* 443 (7111), 521–524. doi:10.1038/443521a
- Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. *Genome Res.* 14 (5), 988–995. doi:10.1101/gr.1865504
- Biscotti, M. A., Adolfini, M. C., Barucca, M., Forconi, M., Pallavicini, A., Gerdol, M., et al. (2018). A comparative view on sex differentiation and gametogenesis genes in lungfish and coelacanth. *Genome Biol. Evol.* 10 (6), 1430–1444. doi:10.1093/gbe/evy101
- Brunet, F., Roche, A., Chalopin, D., Naville, M., Klopp, C., Vizziano-Cantonnet, D., et al. (2018). Analysis of transposable elements expressed in the gonads of the siberian sturgeon. *Sib. Sturgeon (Acipenser baerii, Brandt, 1869) Volume 1-Biology*, 115–130.
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12 (1), 59–60. doi:10.1038/nmeth.3176
- Chalopin, D., Volff, J.-N., Galiana, D., Anderson, J. L., and Scharl, M. (2015). Transposable elements and early evolution of sex chromosomes in fish. *Chromosome Res.* 23, 545–560. doi:10.1007/s10577-015-9490-8
- Chang, K. (2009). “Molecular analysis of the breeding biology of the Asian arowana (*Scleropages formosus*),” (Singapore: National University of Singapore). [PhD thesis].
- Chang, N.-C., Rovira, Q., Wells, J., Feschotte, C., and Vaquerizas, J. M. (2022). Zebrafish transposable elements show extensive diversification in age, genomic distribution, and developmental expression. *Genome Res.* 32 (7), 1408–1423. doi:10.1101/gr.275655.121
- Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., et al. (2017). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* 7 (1), 1–6. doi:10.1093/gigascience/gix120
- Cocquet, J., Pailhoux, E., Jaubert, F., Servel, N., Xia, X., Pannetier, M., et al. (2002). Evolution and expression of *FOXL2*. *J. Med. Genet.* 39 (12), 916–921. doi:10.1136/jmg.39.12.916
- Cui, Z., Liu, Y., Wang, W., Wang, Q., Zhang, N., Lin, F., et al. (2017). Genome editing reveals *dmrt1* as an essential male sex-determining gene in Chinese tongue sole (*Cynoglossus semilaevis*). *Sci. Rep.* 7 (1), 42213. doi:10.1038/srep42213
- Curzon, A. Y., Shirak, A., Ron, M., and Seroussi, E. (2023). Master-key regulators of sex determination in fish and other vertebrates—a review. *Int. J. Mol. Sci.* 24 (3), 2468. [Online]. doi:10.3390/ijms24032468
- Dechaud, C., Miyake, S., Martinez-Bengochea, A., Scharl, M., Volff, J.-N., and Naville, M. (2021). Clustering of sex-biased genes and transposable elements in the genome of the medaka fish *Oryzias latipes*. *Genome Biol. Evol.* 13 (11), evab230. doi:10.1093/gbe/evab230
- Dechaud, C., Volff, J.-N., Scharl, M., and Naville, M. (2019). Sex and the TEs: transposable elements in sexual development and function in animals. *Mob. DNA* 10, 42–15. doi:10.1186/s13100-019-0185-0
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (1), 15–21. doi:10.1093/bioinformatics/bts635
- Dong, X., Chen, S., and Ji, X. (2010). Molecular cloning and expression analysis of *Dmrt3* gene in half-smooth tongue sole (*Cynoglossus semilaevis*). *J. Fish. China* 34 (6), 649–655.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 (5), 1792–1797. doi:10.1093/nar/gkh340
- Farhadi, A., Fang, S., Zhang, Y., Cui, W., Fang, H., Ikhwanuddin, M., et al. (2021). The significant sex-biased expression pattern of Sp-Wnt4 provides novel insights into the ovarian development of mud crab (*Scylla Paramamosain*). *Int. J. Biol. Macromol.* 183, 490–501. doi:10.1016/j.ijbiomac.2021.04.186
- Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* 9 (5), 397–405. doi:10.1038/nrg2337
- Gebrie, A. (2023). Transposable elements as essential elements in the control of gene expression. *Mob. DNA* 14 (1), 9. doi:10.1186/s13100-023-00297-3
- Han, C., Wang, C., Ouyang, H., Zhu, Q., Huang, J., Han, L., et al. (2021). Characterization of *dmrts* and their potential role in gonadal development of Mandarin fish (*Siniperca chuatsi*). *Aquacult. Rep.* 21, 100802. doi:10.1016/j.aqrep.2021.100802

- Hayashi, S., Suda, K., Fujimura, F., Fujikawa, M., Tamura, K., Tsukamoto, D., et al. (2022). Neofunctionalization of a noncoding portion of a DNA transposon in the coding region of the chimerical sex-determining gene *dm-W* in *Xenopus* frogs. *Mol. Biol. Evol.* 39 (7), msac138. doi:10.1093/molbev/msac138
- Herpin, A., Scharlt, M., Depincé, A., Guiguen, Y., Bobe, J., Hua-Van, A., et al. (2021). Allelic diversification after transposable element exaptation promoted *gsdf* as the master sex determining gene of sablefish. *Genome Res.* 31 (8), 1366–1380. doi:10.1101/gr.274266.120
- Hu, Q., Zhu, Y., Liu, Y., Wang, N., and Chen, S. (2014). Cloning and characterization of *wnt4* gene and evidence for positive selection in half-smooth tongue sole (*Cynoglossus semilaevis*). *Sci. Rep.* 4 (1), 7167. doi:10.1038/srep07167
- Hu, Y., Wang, B., and Du, H. (2021). A review on *sox* genes in fish. *Rev. Aquacult.* 13 (4), 1986–2003. doi:10.1111/raq.12554
- Huang, Z., Xu, L., Cai, C., Zhou, Y., Liu, J., Xu, Z., et al. (2023). Three amphioxus reference genomes reveal gene and chromosome evolution of chordates. *Proc. Natl. Acad. Sci.* 120 (10), e2201504120. doi:10.1073/pnas.2201504120
- Jeong, H.-H., Yalanchili, H. K., Guo, C., Shulman, J. M., and Liu, Z. (2017). “An ultra-fast and scalable quantification pipeline for transposable elements from next generation sequencing data,” in *Bioinformatics 2018* (Singapore: World Scientific), 168–179.
- Jin, Y., Tam, O. H., Paniagua, E., and Hammell, M. (2015). TETranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* 31 (22), 3593–3599. doi:10.1093/bioinformatics/btv422
- Kamiya, T., Kai, W., Tasumi, S., Oka, A., Matsunaga, T., Mizuno, N., et al. (2012). A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu). *PLoS Genet.* 8 (7), e1002798. doi:10.1371/journal.pgen.1002798
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27
- Kitano, J., Ansai, S., Takehana, Y., and Yamamoto, Y. (2023). Diversity and convergence of sex determination mechanisms in teleost fish. *Annu. Rev. Anim. Biosci.* 12, 233–259. doi:10.1146/annurev-animal-021122-113935
- Kobayashi, T., Matsuda, M., Kajiura-Kobayashi, H., Suzuki, A., Saito, N., Nakamoto, M., et al. (2004). Two DM domain genes, *DMY* and *DMRT1*, involved in testicular differentiation and development in the medaka, *Oryzias latipes*. *Dev. Dyn.* 231 (3), 518–526. doi:10.1002/dvdy.20158
- Kossack, M. E., High, S. K., Hopton, R. E., Yan, Y.-I., Postlethwait, J. H., and Draper, B. W. (2018). Female sex development and reproductive duct formation depend on *Wnt4a* in zebrafish. *Genetics* 211 (1), 219–233. doi:10.1534/genetics.118.301620
- Lanciano, S., and Cristofari, G. (2020). Measuring and interpreting transposable element expression. *Nat. Rev. Genet.* 21 (12), 721–736. doi:10.1038/s41576-020-0251-y
- Lee, H. J., Hou, Y., Maeng, J. H., Shah, N. M., Chen, Y., Lawson, H. A., et al. (2022). Epigenomic analysis reveals prevalent contribution of transposable elements to cis-regulatory elements, tissue-specific expression, and alternative promoters in zebrafish. *Genome Res.* 32 (7), 1424–1436. doi:10.1101/gr.276052.121
- Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49 (W1), W293–W296. doi:10.1093/nar/gkab301
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 323–416. doi:10.1186/1471-2105-12-323
- Li, J.-Z., Liu, Q., Wang, D.-S., Zhou, L.-Y., Sakai, F., and Nagahama, Y. (2012). Molecular cloning and identification of two *Wnt4* genes from the medaka (*Oryzias latipes*). *Acta Hydrobiol. Sin.* 36 (5), 983–986. doi:10.3724/SP.J.1035.2012.00983
- Li, Q., Zhou, X., Guo, Y., Shang, X., Chen, H., Lu, H., et al. (2008). Nuclear localization, DNA binding and restricted expression in neural and germ cells of zebrafish *Dmrt3*. *Biol. Cell* 100 (8), 453–463. doi:10.1042/BC20070114
- Li, S., Li, W., Jiang, S., Jing, Y., Xiao, L., Yu, Y., et al. (2023). Mechanisms of sex differentiation and sex reversal in hermaphrodite fish as revealed by the *Epinephelus coioides* genome. *Mol. Ecol. Resour.* 23 (4), 920–932. doi:10.1111/1755-0998.13753
- Li, X.-Y., and Gui, J.-F. (2018). Diverse and variable sex determination mechanisms in vertebrates. *Sci. China Life Sci.* 61, 1503–1514. doi:10.1007/s11427-018-9415-7
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550–621. doi:10.1186/s13059-014-0550-8
- Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., and Zarkower, D. (2011). *DMRT1* prevents female reprogramming in the postnatal mammalian testis. *Nature* 476 (7358), 101–104. doi:10.1038/nature10239
- Matson, C. K., and Zarkower, D. (2012). Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat. Rev. Genet.* 13 (3), 163–174. doi:10.1038/nrg3161
- Mawaribuchi, S., Ito, Y., and Ito, M. (2019). Independent evolution for sex determination and differentiation in the *DMRT* family in animals. *Biol. Open* 8 (8), bio041962. doi:10.1242/bio.041962
- Mu, X., Liu, Y., Liu, C., Zhao, C., Li, R., You, X., et al. (2022). Identification of candidate sex-specific genomic regions in male and female Asian arowana genomes. *GigaScience* 11, giac085. doi:10.1093/gigascience/giac085
- Mu, X., Wang, X., Song, H., Yang, Y., Luo, D., Gu, D., et al. (2012). Mitochondrial DNA as effective molecular markers for the genetic variation and phylogeny of the family Osteoglossidae. *Gene* 511 (2), 320–325. doi:10.1016/j.gene.2012.09.087
- Muncaster, S., Goikotxea, A., Lokman, P. M., De Farias e Moraes, C. E., Damsteegt, E. L., Edgcombe, J., et al. (2023). Genes involved in sex differentiation, epigenetic reprogramming, and cell fate regulate sex change in a wrasse. *Rev. Fish. Biol. Fish.* 33 (1), 281–294. doi:10.1007/s11160-022-09755-2
- Mustapha, U. F., Peng, Y.-X., Huang, Y.-Q., Assan, D., Zhi, F., Shi, G., et al. (2022). Comparative transcriptome analysis of the differentiating gonads in *Scatophagus argus*. *Front. Mar. Sci.* 9, 962534. doi:10.3389/fmars.2022.962534
- Myosho, T., Otake, H., Masuyama, H., Matsuda, M., Kuroki, Y., Fujiyama, A., et al. (2012). Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias latipes*. *Genetics* 191 (1), 163–170. doi:10.1534/genetics.111.137497
- Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., et al. (2002). A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc. Natl. Acad. Sci.* 99(18), 11778–11783. doi:10.1073/pnas.182314699
- Pan, Q., Kay, T., Depincé, A., Adolphi, M., Scharlt, M., Guiguen, Y., et al. (2021). Evolution of master sex determiners: TGF- β signalling pathways at regulatory crossroads. *Phil. Trans. R. Soc. B* 376 (1832), 20200091. doi:10.1098/rstb.2020.0091
- Parey, E., Louis, A., Montfort, J., Bouchez, O., Roques, C., Iampietro, C., et al. (2023). Genome structures resolve the early diversification of teleost fishes. *Science* 379(6632), 572–575. doi:10.1126/science.abq4257
- Richards, J. S. (2018). “Chapter one - the ovarian cycle,” in *Vitamins and hormones*. Editor G. Litwack (United States: Academic Press), 1–25.
- Shen, X., Kwan, H., Thevasagayam, N., Prakkai, S., Kuznetsova, I., Ngoh, S., et al. (2014). The first transcriptome and genetic linkage map for Asian arowana. *Mol. Ecol. Resour.* 14 (3), 622–635. doi:10.1111/1755-0998.12212
- Smith, C. A., Roeszler, K. N., Ohnesorg, T., Cummins, D. M., Farlie, P. G., Doran, T. J., et al. (2009). The avian Z-linked gene *DMRT1* is required for male sex determination in the chicken. *Nature* 461 (7261), 267–271. doi:10.1038/nature08298
- Stachowiak, M., Nowacka-Woszek, J., Szabelska-Berezewicz, A., Zyprych-Walczak, J., Krzeminska, P., Sosinski, O., et al. (2024). A massive alteration of gene expression in undescended testicles of dogs and the association of *KAT6A* variants with cryptorchidism. *Proc. Natl. Acad. Sci.* 121 (7), e2312724121. doi:10.1073/pnas.2312724121
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22 (21), 2688–2690. doi:10.1093/bioinformatics/btl446
- Su, G., Morris, J. H., Demchak, B., and Bader, G. D. (2014). Biological network exploration with CytoScape 3. *Curr. Protoc. Bioinforma.* 47 (1), 8.13.11–24. doi:10.1002/0471250953.bi0813s47
- Sundaram, V., and Wysocka, J. (2020). Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Phil. Trans. R. Soc. B* 375 (1795), 20190347. doi:10.1098/rstb.2019.0347
- Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., et al. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51 (D1), D638–D646. doi:10.1093/nar/gkac1000
- Takase, H. M., and Nusse, R. (2016). Paracrine Wnt/ β -catenin signaling mediates proliferation of undifferentiated spermatogonia in the adult mouse testis. *Proc. Natl. Acad. Sci.* 113 (11), E1489–E1497. doi:10.1073/pnas.1601461113
- Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinforma.* 25 (1), Unit 4.10–14.10.14. doi:10.1002/0471250953.bi0410s25
- Toma, G. A., dos Santos, N., dos Santos, R., Rab, P., Kretschmer, R., Ezaz, T., et al. (2023). Cytogenetics meets genomics: cytotaxonomy and genomic relationships among color variants of the Asian arowana *Scleropages formosus*. *Int. J. Mol. Sci.* 24 (10), 9005. doi:10.3390/ijms24109005
- Toubiana, W., Armisen, D., Dechaud, C., Arbore, R., and Khila, A. (2021). Impact of male trait exaggeration on sex-biased gene expression and genome architecture in a water strider. *BMC Biol.* 19 (1), 89–17. doi:10.1186/s12915-021-01021-4
- Wang, D.-S., Kobayashi, T., Zhou, L.-Y., Paul-Prasanth, B., Jjiri, S., Sakai, F., et al. (2007). Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as interacting with Ad4 binding protein/steroidogenic factor 1. *Mol. Endocrinol.* 21 (3), 712–725. doi:10.1210/me.2006-0248
- Wang, D.-S., Zhou, L.-Y., Kobayashi, T., Matsuda, M., Shibata, Y., Sakai, F., et al. (2010). Doublesex- and mab-3-related transcription factor-1 repression of aromatase transcription, a possible mechanism favoring the male pathway in *Tilapia*. *Endocrinology* 151 (3), 1331–1340. doi:10.1210/en.2009-0999

- Webster, K. A., Schach, U., Ordaz, A., Steinfeld, J. S., Draper, B. W., and Siegfried, K. R. (2017). *Dmrt1* is necessary for male sexual development in zebrafish. *Dev. Biol.* 422 (1), 33–46. doi:10.1016/j.ydbio.2016.12.008
- Xie, Y., Li, H., Luo, X., Li, H., Gao, Q., Zhang, L., et al. (2022). IBS 2.0: an upgraded illustrator for the visualization of biological sequences. *Nucleic Acids Res.* 50 (W1), W420–W426. doi:10.1093/nar/gkac373
- Xu, A., Teefy, B. B., Lu, R. J., Nozownik, S., Tyers, A. M., Valenzano, D. R., et al. (2023). Transcriptomes of aging brain, heart, muscle, and spleen from female and male African turquoise killifish. *Sci. Data* 10 (1), 695. doi:10.1038/s41597-023-02609-x
- Yang, W. R., Ardeljan, D., Pacyna, C. N., Payer, L. M., and Burns, K. H. (2019). SQuIRE reveals locus-specific regulation of interspersed repeat expression. *Nucleic Acids Res.* 47 (5), e27. doi:10.1093/nar/gky1301
- Ye, J., McGinnis, S., and Madden, T. L. (2006). BLAST: improvements for better sequence analysis. *Nucleic Acids Res.* 34 (Web Server issue), W6–W9. doi:10.1093/nar/gkl164
- Yoshimoto, S., Ikeda, N., Izutsu, Y., Shiba, T., Takamatsu, N., and Ito, M. (2010). Opposite roles of DMRT1 and its W-linked paralogue, DM-W, in sexual dimorphism of *Xenopus laevis*: implications of a ZZ/ZW-type sex-determining system. *Development* 137 (15), 2519–2526. doi:10.1242/dev.048751
- Yuan, Z., Shen, X., Yan, H., Jiang, J., Liu, B., Zhang, L., et al. (2021). Effects of the thyroid endocrine system on gonadal sex ratios and sex-related gene expression in the pufferfish *Takifugu rubripes*. *Front. Endocrinol.* 12, 674954. doi:10.3389/fendo.2021.674954
- Zhang, H.-M., Liu, T., Liu, C.-J., Song, S., Zhang, X., Liu, W., et al. (2015). AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Res.* 43 (D1), D76–D81. doi:10.1093/nar/gku887
- Zhang, K., Xu, J., Zhang, Z., Huang, Y., Ruan, Z., Chen, S., et al. (2019). A comparative transcriptomic study on developmental gonads provides novel insights into sex change in the protandrous black porgy (*Acanthopagrus schlegelii*). *Genomics* 111 (3), 277–283. doi:10.1016/j.ygeno.2018.11.006
- Zhang, X., Li, M., Ma, H., Liu, X., Shi, H., Li, M., et al. (2017). Mutation of foxl2 or cyp19a1a results in female to male sex reversal in XX Nile Tilapia. *Endocrinology* 158 (8), 2634–2647. doi:10.1210/en.2017-00127
- Zhang, X., Zhou, J., Li, L., Huang, W., Ahmad, H. I., Li, H., et al. (2020). Full-length transcriptome sequencing and comparative transcriptomic analysis to uncover genes involved in early gametogenesis in the gonads of Amur sturgeon (*Acipenser schrenckii*). *Front. Zool.* 17 (1), 11–21. doi:10.1186/s12983-020-00355-z