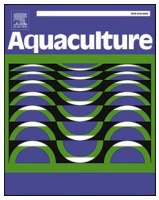
Aquaculture 575 (2023) 739771

Contents lists available at ScienceDirect

Aquaculture

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



Comparative study of candidate sex determination regions in snakeheads (*Channa argus* and *C. maculata*) and development of novel sex markers

Donglei Sun [a], Haishen Wen [a], Xin Qi [a], Chao Li [b], Chaonan Sun [a], Lingyu Wang [a], Mingxin Zhu [a], Tianyu Jiang [a], Xiaoyan Zhang [b], Yun Li [a,*]

[a] Key Laboratory of Mariculture (Ocean University of China), Ministry of Education (KLMME), Fisheries College, Ocean University of China, Qingdao 266003, China
[b] School of Marine Science and Engineering, Qingdao Agricultural University, Qingdao 266109, China

ARTICLE INFO

Keywords:
Snakeheads
GWAS
Sex determination
Sex markers

ABSTRACT

Belonging to the Channidae family, *Channa argus* and *C. maculata* are the main cultured snakehead species in China and have important economic value. Both species show distinct sexual dimorphism in growth, with males growing much faster than females. Therefore, the production of all males will majorly improve the efficiency of aquaculture. However, the sex-determination (SD) mechanism of snakeheads remains unclear. Herein, we used whole-genome resequencing data of *C. maculata* and *C. argus* to analyse sex-linked genomic regions in these two species. By using GWAS and Pool-seq analysis methods, their potential SD regions were refined to fewer than 200 kb, containing two protein-coding genes (*rnf144a* and *id2*) without known sex-related function. The gene content on the sex chromosomes of the two snakehead species was basically consistent, and gonadal transcriptome analysis results showed that the mRNA expression level of *id2* in the testis was significantly lower than that in the ovary in both species, whereas *rnf144a* showed inconsistent expression trends in the two species. Moreover, we assembled the male reference genomes and identified a 4–5 kb male-specific noncoding fragment for both species, allowing us to design universal sex-specific markers for *C. argus*, *C. maculata*, and their hybrids (*C. maculata* ♀ × *C. argus* ♂). Our findings will facilitate the understanding of the molecular basis of the SD mechanism in snakeheads, and further promote their all-male breeding progress in the industry.

1. Introduction

The northern snakehead, *Channa argus*, belonging to the Channidae family, is known as an economically important freshwater species native to East Asia. Owing to its strong growth ability, hypoxia tolerance and disease resistance, as well as the advantage of lacking intermuscular bones that are easily processed for consumers, *C. argus* has become extremely popular for aquaculture in China, with annual production exceeding 500,000 tons (Sun et al., 2023). It has been well documented that *C. argus* displays sexual dimorphism in growth in that the body size of males is approximately two times larger than that of females at 2–3 years old (Liu et al., 2011; Wang et al., 2019). The related blotched snakehead (*C. maculata*) also exhibits sexual dimorphism, with males growing significantly faster than females (Ou et al., 2018; Zhao et al., 2021). Moreover, the commercial hybrids of these two snakehead species, which present obvious heterosis and are widely produced in breeding, show similar sexual growth dimorphism as expected (Ou et al., 2018; Zhao et al., 2021). Therefore, mono-sex culture of the all-male population is a very effective way to achieve higher production and incomes for snakehead industry (Ou et al., 2021a, 2021b; Zhao et al., 2021). This requires the development of sex-specific markers for the establishment of artificial mono-sex breeding technology and raises an interesting question about the sex determination (SD) mechanisms of these species.

Previous studies have been conducted on the SD system of *C. argus* and *C. maculata*, and demonstrated that they possessed morphologically undifferentiated sex chromosomes with XY male heterogametic SD patterns (Wu et al., 1986; Ou et al., 2017; Wang et al., 2019; Yang et al., 2020). To identify the potential SD region and sex-determining genes, quantitative trait locus (QTL) mapping studies based on linkage mapping have been performed for *C. argus*, which uncovered a sex chromosome carrying the major sex-associated region (Wang et al., 2019). For example, the dd-RAD sequencing approach was applied for 74 F1 progenies, which identified a major QTL with a length of 6.18 Mb at

* Corresponding author.
E-mail address: yunli0116@ouc.edu.cn (Y. Li).

https://doi.org/10.1016/j.aquaculture.2023.739771
Received 31 March 2023; Received in revised form 3 June 2023; Accepted 8 June 2023
Available online 12 June 2023

LG16 as the sex-linked region, containing 291 protein-coding genes (Wang et al., 2019). Later, by using 160 F1 progenies, the 2b-RAD sequencing method revealed a major sex-associated interval at LG5 (78.837–79.236 cM), and several genes were annotated in this region (Liu et al., 2020). Interestingly, none of the genes identified in both studies have been reported as necessary for sex determination, such as the Y-specific DM domain gene (*dmy*) in medaka (Matsuda et al., 2002, 2003), anti-Müllerian hormone (*amhy*) in black rockfish (*Sebastes schlegelii*) (Song et al., 2021), doublesex and mab-3-related transcription factor 1 (*dmrt1*) in Chinese tongue sole (*Cynoglossus semilaevis*) (Chen et al., 2014; Cui et al., 2017) and fighting fish (*Betta splendens*) (Kwon et al., 2022; Wang et al., 2022; Zhang et al., 2022), and breast cancer anti-resistance 1 (*bcar1*) in channel catfish (*Ictalurus punctatus*) (Bao et al., 2019), which have been reported as master SD genes in fish species with XY SD systems. This suggests that the master sex-determining factor in *C. argus* should be a novel gene or a regulatory element. Additionally, sex-linked insertion and deletions (InDels) have been reported for both *C. argus* and *C. maculata* (Ou et al., 2017; Wang et al., 2019), but their length was restricted to a few to hundred base pairs and no large DNA variants were observed between sexes in these data. Because the methods employed by these studies were based on the reduced-representation sequencing platform, and the genome assembly of XX female was used as the reference, it is difficult to determine whether larger forms of Y-chromosome variants (male-specific insertion) exist.

In this study, we generated whole-genome resequencing data of females and males of *C. argus*, and took advantage of the available whole-genome resequencing data of *C. maculata* to investigate the sex-linked genomic variations in these two species. By implementing a genome-wide association study (GWAS) and pooled sequencing analysis (Pool-seq) strategy, their candidate SD region was narrowed to <200 kb, which contained two protein coding genes without known sex-related functions. Additionally, the gene contents and sexual bias expressions on sex chromosomes were compared between the two snakeheads. We further assembled the male reference genomes and identified a male-specific noncoding insertion approximately 4–5 kb in length for both species, which enabled us to design universal sex-specific genetic markers for *C. argus*, *C. maculata* and their hybrids. Our findings will contribute to a better understanding of the molecular basis of sex determination in *C. argus* and *C. maculata* and facilitate the development of all-male breeding in the snakehead industry.

2. Materials and methods

2.1. Ethics statement

All procedures involved in handling and treatment of fish in this study were approved by Animal Research and Ethics Committees of Ocean University of China (Permit Number: 20141201). The field studies did not involve endangered or protected species.

2.2. Fish sampling, DNA extraction

30 females and 30 males of adult *C. argus* (body length: 50.1 ± 4.6 cm; body weight: 1.41 ± 0.36 kg) were captured from Dechang fishery Co. Ltd. in Jining (JN), Shandong, China, and another 30*C. argus* individuals (15 females and 15 males (body length: 20.7 ± 3.8 cm; body weight: 141.8 ± 36.4 g)) were collected from Lianyungang (LYG), Jiangsu. Moreover, a total of 30 matured *C. maculata* (15 females and 15 males (body length: 37.0 ± 5.2 cm; body weight: 781.6 ± 167.4 g)) and 30 hybrid individuals (*C. maculata* ♀ × *C. argus* ♂) (15 females and 15 males (body length: 46.3 ± 5.8 cm; body weight: 1.05 ± 0.22 kg)) were obtained from Foshan (FS), Guangdong. The phenotypic sex of these fish was then identified through visual assessment and histological examination of their gonads. The caudal fin tissue was collected from all samples mentioned above and fixed in 100% ethanol. Using the TIANamp Marine Animals DNA Kit (TIANGEN, China), DNA was extracted according to the manufacturer's instructions. The quality of DNA was examined by QuickDrop spectrophotometer (Spectramax, USA) and 1.0% agarose gel electrophoresis.

2.3. Whole genome resequencing

From the JN population, 40*C. argus* samples (18 females and 22 males) were chosen for whole genome resequencing (WGS). The 150 bp paired-end sequencing libraries with a 350 bp insert size were produced using the TruSeq Library Construction Kit (Illumina, USA) and sequenced using the DNBSEQ-T7 platform. Raw reads generated by WGS were processed using fastp (v 0.20.0) with default parameters to eliminate low-quality sequences. The WGS data have been deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) with the accession number PRJNA834927.

2.4. Genome-wide association study (GWAS)

After quality control of the sequencing data, the clean reads of each library were separately mapped to the female reference genome of the *C. argus* (GenBank: GCA_027943205.1) using BWA (v 0.7.17) mem mode (settings: mem -t 4 -k 32 -M -R) (Li and Durbin, 2009). Using SAMtools (v 1.10) (settings: -bS -t), alignment files were converted to BAM files (Li et al., 2009). Additionally, potential PCR duplications were removed using SAMtools command "rmdup". Variants calling were performed for all BAM files by the HaplotypeCaller protocol in Genome Analysis Toolkit (GATK, v 4.0) (http://www.broadinstitute.org/gatk/download) (McKenna et al., 2010). The single nucleotide polymorphisms (SNPs) were filtered using GATK VariantFiltration with parameter setting as: –filterExpression "QD < 4.0, FS > 60.0, MQ < 40.0 ", −G_filter "GQ < 20″. We further filtered the variants with "–min-meanDP 5, –max-meanDP 100, –max-missing 0.95" using VCFtools (v 0.1.16) (Danecek et al., 2011). Genome-wide association mapping was conducted using GEMMA (v 0.98.5) with linear mixed models (LMMs). The Wald test is used to estimate the association strength (*P*-values), and the genome-wide significance level is set as $-\log_{10}(0.01/N)$, where N represents the total number of SNPs used for association test. The Manhattan plot was created using the CMplot R package (https://github.com/YinLiLin/CMplot). The significant SNPs were annotated by ANNOVAR (Wang et al., 2010) based on the GFF files of the reference genome.

In addition, we obtained WGS data of 24*C. maculata* samples from the NCBI SRA repository (Accession No. SRR11310364-SRR11310381, SRR13165338-SRR13165344). In total, the sequencing data from 12 female and 12 male samples were used to GWAS analysis to identify the SD region of *C. maculata*. Clean reads were separately mapped to the female reference genome of *C. maculata* (GenBank: GCA_020496755.1). The detailed data processing and analysis procedure was consistent with that of the *C. argus* as described above.

2.5. Pool-sequencing analysis (Pool-seq)

To further verify the genomic loci potentially linked to sex-determination, the female and male pools were separately concatenated from clean data of 40*C. argus* samples (18 females and 22 males) described above using seqtk (v1.2) (https://github.com/lh3/seqtk). Clean reads (from both the male and female pools) were separately mapped to the female reference genome of the *C. argus*, resulting in BAM files using the above processing method. After that, using PSASS (Pooled Sequencing Analysis for Sex Signal, https://github.com/SexGenomicsToolkit/PSASS) software command "pileup", we generated a pileup file by the two sex BAM files. To identify sex-biased SNPs in enriched regions, a 100 kb sliding window with an output point every 10 kb was utilized with PSASS. The parameters were as follows: minimum depth of a site was set to 50 (−-min-depth 50), range of frequency

for a sex-linked SNP in the heterogametic sex was set to 0.5 ± 0.15 (−-freq-het 0.5, −-range-het 0.15), and frequency of a sex-linked SNP in the homogametic sex set to >0.95 (−-freq-hom 1, −-range-hom 0.05). The final genomic information was visually present using ggplot2 R package (https://github.com/tidyverse/ggplot2).

Similarly, the female and male pools were generated from the above-mentioned 12 female and 12 male WGS data of *C. maculata*. Clean reads (male and female pools) were mapped to the female reference genome of the *C. maculata* separately to obtain the BAM file. For the Pool-seq method, the SD region of *C. maculata* was identified using the same analysis process as for *C. argus* described above.

2.6. Male genome assembly and sex-specific sequence identification

We used WGS data from 40*C. argus* and 24*C. maculata* to screen sex-specific sequences. After data filtering, clean reads from male pools of *C. argus* and *C. maculata* were separately assembled using masurca software (v 4.0.5) to construct male reference genomes of *C. argus* and *C. maculata*, respectively (Zimin et al., 2013). The following parameters were set apart from default: jellyfish hash size (JF_SIZE = 15,000,000,000), paired-end insert size and standard deviation (350, 50). Then the male assembly were utilized for sex-specific sequence identification. First, male and female clean data were mapped to the male reference genome using BWA mem mode and default settings, and the genomic region that was covered by all male samples without any read from female samples was deemed as male-specific sequence. The coverage and depth information for each locus of the BAM file was obtained using "SAMtools depth" command. The overall workflow for screening the sex-specific sequences was showed in Fig. S1.

2.7. RNA-Seq analysis

For the transcriptome analysis of gonads, four-month-old *C. argus* (body length: 22.8 ± 2.7 cm; body weight: 156.3 ± 66.1 g) were reared in fresh water (18.0 ± 1.5 °C) at Daqiang fisheries Co., Ltd. in Linyi, Shandong province, China for transcriptome sequencing. The gonads from 18*C. argus* (9 females and 9 males) were immediately frozen in liquid nitrogen for subsequent RNA isolation. Total RNA was isolated from the testis and ovary tissues using TRIzol Reagent (Invitrogen, USA) with manufacturer's protocol. The quality of RNA was examined by QuickDrop spectrophotometer (Spectramax, USA) and 1.0% agarose gel electrophoresis. Equal amounts of RNA from 3*C. argus* individuals were pooled for each library, and a total of 6 RNA-Seq libraries (3 libraries for each sex, named as CA_F1–3, CA_M1–3) were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's instructions. The libraries were sequenced on Illumina NovaSeq 6000 platforms with PE150 strategy in Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). Moreover, a total of 6 RNA-Seq datasets (CM_F1–3, CM_M1–3) were obtained from Ou et al. (Accession No. SAMN13677127–13677138) for the transcriptome analysis of the *C. maculata* gonad (Ou et al., 2020). The raw sequencing data was first processed by removing adapters and low-quality reads using Trimmomatic (v 0.39). The resulting clean paired-end reads were mapped to the female reference genome of *C. argus* and *C. maculata* using Hisat2 (v 2.2.1) (Kim et al., 2015), respectively. Then, the read number mapped to each gene was counted using featureCounts (v 2.0.1) (Liao et al., 2013). Differential gene expression analysis was performed using the DESeq2 R package (v 1.38.2) (Love et al., 2014). The differential expressed genes (DEGs) were determined with the thresholds of q-value (adjusted *P*-value using the Benjamini & Hochberg method) < 0.05 and |log2(Fold Change)| > 1, and the transcripts per million (TPM) value was used for measuring gene expression levels.

2.8. Syntenic analysis of id2 and rnf144a

Syntenic analysis of *id2* and *rnf144a* among snakeheads and other teleost fishes, including spotted gar, stickleback, fugu, tilapia, Amazon molly, medaka, zebrafish and cave fish was performed. The gene locations of *id2* and *rnf144a* were obtained from the genomic annotations, which were downloaded from the NCBI.

2.9. Confirmation of male-specific sequences with PCR amplification and sanger sequencing

The sex-specific sequences were validated in *C. argus* (n = 90), *C. maculata* (n = 30) and hybrid snakehead (*C. maculata* ♀ × *C. argus* ♂) (n = 30), respectively. Two primer pairs were designed using Primer 5.0 based on the sex-specific sequences (Table S1). The PCR amplification was carried out in a total of 20 μl reaction system, including 10 μl of Taq PCR Mix 2× (Vazyme, China), 1 μl of each primer (10 uM), 1 μl of DNA template and 7 μl RNAase-free water. PCR amplification was performed on T100TM Thermal Cycl er (Bio-Rad, Germany) as follows: initial denaturation at 95 °C for 3 min, followed by 35 cycles at 95 °C for 30 s, 50–55 °C for 1.5–3.5 min (1 min/kb) and 72 °C for 10 min. Gel electrophoresis was used to detect the PCR products by 1.0% agarose gel. PCR products with expected sizes were purified by the FastPure Gel DNA Extraction Mini Kit (Vazyme, China), and were subsequently sequenced using the Sanger technique (BGI company, China).

3. Results

3.1. Summary of whole genome resequencing data

Whole genome sequencing of 18 females and 22 males *C. argus* produced 336.95 Gb data in total, with the estimated depths ranging from 10.00× to 19.65× for each sample (Table S2). Approximately, 99.47% ~ 99.82% reads from 40 individuals were mapped to female reference genome (GenBank: GCA_027943205.1) (Table S2). Subsequently, the female and male pools were obtained by combining the sequencing data of each sex, which yielded 1,135,437,838 and 1,310,702,078 pairs of clean reads, respectively (Table S2). With the female *C. argus* genome as the reference, the mapping rate of clean reads from male pool and female pool were about 99.66% and 99.58%, and the estimated average depths were 239.63× and 272.37×, respectively (Table S2).

For *C. maculata*, totally 796.12 Gb WGS data were obtained from12 females and 12 males, and the estimated depths ranged from 20.05× to 107.06× for each sample (Table S3). Using the female *C. maculata* (GenBank: GCA_020496755.1) as the reference genome, the mapping rate of clean reads from 24 individuals were ranged from 98.13% to 99.26% (Table S3). After combining the sequencing data of each sex, 2,693,377,142 and 2,583,853,519 pairs of clean reads were generated for female and male pools, respectively (Table S3). The mapping rate of clean reads from female pool and male pool to the female reference genome of *C. maculata* were 99.00% and 98.93%, with the average depths calculated as 641.31× and 614.43×, respectively (Table S3).

3.2. Identification of sex chromosomes and sex-determination (SD) regions in C. argus and C. maculata

Two methods were used to identify the sex chromosomes and SD regions in snakeheads. First, GWAS analysis was performed by using 698,310 high-quality SNPs with linear mixed models (LMMs), which revealed a single genome-wide peak and a 5.38-Mb region on chromosome 18 (Chr18: 8,616,487-13,995,663) in the female genome of *C. argus* (Fig. 1A). Second, we used the Pool-seq analysis for the female pool and male pool to map against the female genome, and searched for sex-biased signals, including sex coverage differences and sex-biased SNPs. As a result, whole genome analysis of SNP distribution revealed an ~190 kb (Chr18: 11,390,000-11,580,000) region on Chr18 exhibiting a strong sex-linked signal in males (Fig. 1B-D), with a large number of observed SNPs (2408 SNPs) being heterozygous in the male pool and

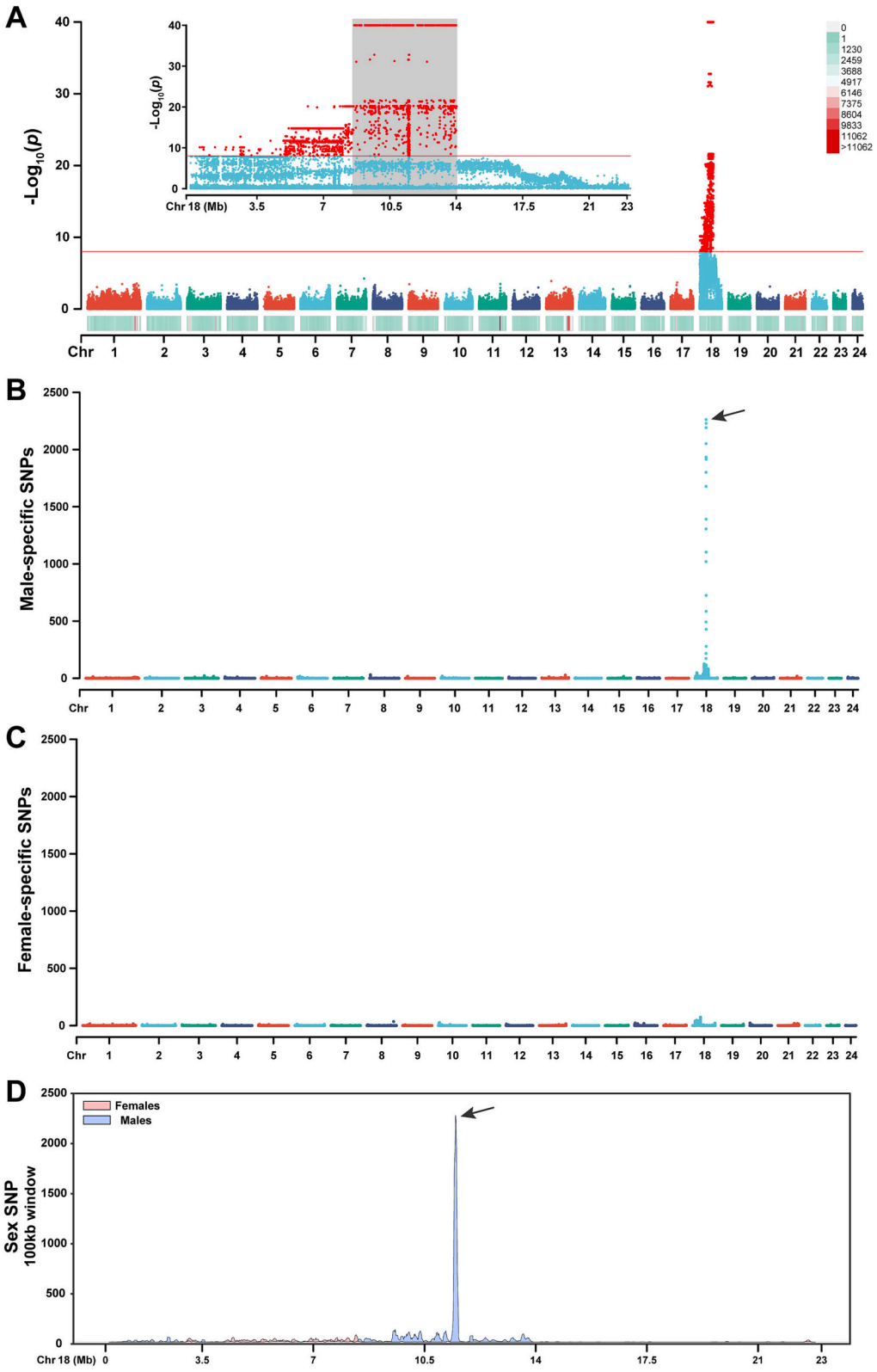


Fig. 1. Sex determination (SD) region of *C. argus* identified by GWAS and Pool-seq analyses. A. GWAS analysis using linear mixed models (LMMs) revealed that the genome-wide peak (5.38-Mb) at Chr18 was recognized to play sex determinate function in *C. argus*. B. The distribution of SNPs showed that male-specific SNPs were enriched in a region on Chr18 that was regarded as the SD region in *C. argus*. SNPs were counted using 100 kb sliding window with an output point every 500 bp. All the 24 chromosomes (Chrs) were labelled with their Chr number. C. The chromosome distribution of female-specific SNPs was present on Manhattan plot in *C. argus*. D. A significant ~190 kb SD region was identified on Chr18, which was indicated by the arrow. The female-specific and male-specific SNPs were respectively indicated by red and blue colour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

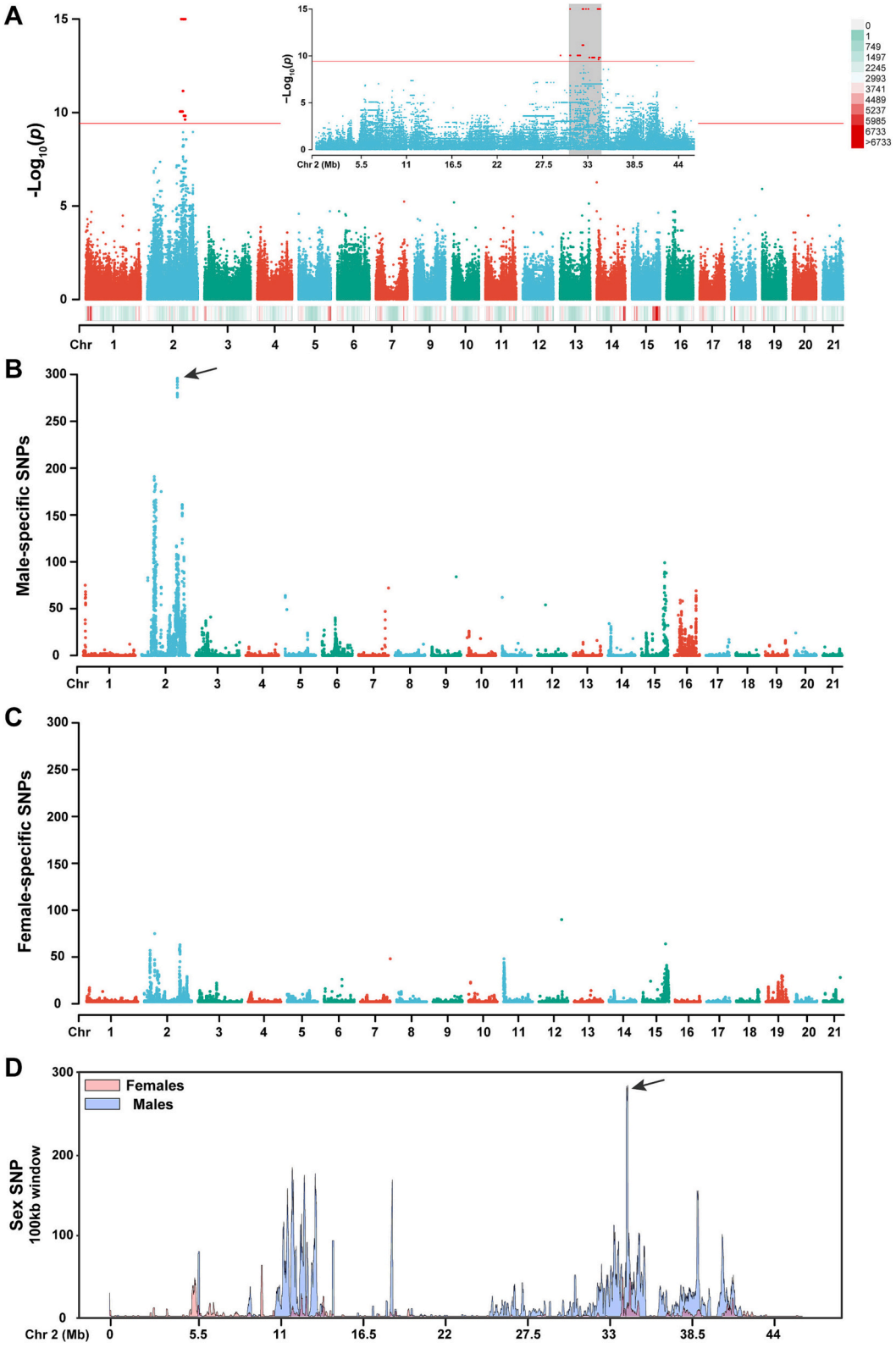


Fig. 2. Sex determination (SD) region of *C. maculata* identified by GWAS and pool-sequencing analysis. A. GWAS analysis using linear mixed models revealed that only one genome-wide peak (5.20-Mb) at Chr2 was significantly associated with sex trait in *C. maculata*. B. The distribution of SNPs showed that male-specific SNPs were enriched in a region on Chr2 that was regarded as the potential SD region in *C. maculata*. SNPs were counted using 100 kb sliding window with an output point every 500 bp. All the 21 chromosomes (Chrs) were labelled with their Chr number. C. The chromosome distribution of female-specific SNPs. D. A significant ~180 kb potential SD region was identified on Chr2, which was indicated by the arrow. The female-specific and male-specific SNPs were respectively indicated by red and blue colour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Information of assembled male reference genomes of *C. argus* and *C. maculata*.

Features	*C. argus*	*C. maculata*
Assembled Size (Mb)	605.69	628.52
Seq Num	25,732	33,772
Average length (bp)	23,538.58	18,610.83
Largest length (bp)	611,525	740,730
N50	70,808	92,189
Mapping ratio (female pool)	99.32%	99.34%
Mapping ratio (male pool)	99.21%	99.23%

homozygous in the female pool (Table S4). These findings indicated that Chr18, possessing sex-linked genomic sequences was recognized to play a sex determinate function in *C. argus*.

Simultaneously, we used the WGS data of 12 females and 12 males and applied strategies similar to those used in *C. argus* to identify the potential sex chromosome and SD region in *C. maculata*. A total of 1,587,458 filtered SNPs were retained and applied for GWAS analysis. The GWAS result showed that a 5.20-Mb region on Chr2 (Chr2: 29,681,280-34,884,623) of the female reference genome was significantly associated with the sex trait of *C. maculata* (Fig. 2A). Pool-seq analysis further confirmed the distribution of male-biased SNPs on Chr2 and identified the ~180 kb (Chr2: 34,260,000-34,440,000) region as the potential SD region of *C. maculata* (Fig. 2B-D).

3.3. Identification of sex-specific sequences in C. argus and C. maculata

Furthermore, we assembled the male reference genome of *C. argus* based on the WGS clean reads of males to explore the male-specific region. The total assembled genome sequence was 605.69 Mb, containing 25,732 contigs with an N50 of 70,808 bp (Table 1). Then the male genome assembly was used as the reference, and sequences from the above female and male pools were separately aligned with it. The mapping ratio of reads in each library to the male reference genome was >99%, indicating that the new reference genome was assembled well (Table 1). Notably, by using Integrative Genomics Viewer (IGV), two genomic segments in assembled contig jcf7180001325712 with lengths of ~4.6 kb and ~ 1.6 kb respectively were identified as male-specific (Y-chromosome specific) insertions where no read could be detected from the female pool covering this area (Fig. 3A), however, no protein coding gene was identified in these insertion regions. Moreover, an ~1.7 kb region in contig jcf7180001322946 with depth of coverage difference between the male and female pools was identified, corresponding to the sex-linked pattern in which the male pool harbored approximately half of the sequence depth in the female pool (Fig. 3B), which was considered the X-chromosome-specific genomic fragment. We then mapped the two sex-linked contigs against the female reference genome, revealing that they are located at Chr18, corresponding to Chr18: 11,469,793 to 11,520,052 bp (Fig. 3C).

The male reference genome of *C. maculata* was also assembled using the WGS data generated from males, yielding a 628.52 Mb sequencing assembly with an N50 of 92,189 bp (Table 1). After alignment by female and male Pool-seq reads, an ~4.0 kb male-specific insertion was observed in assembled contig jcf7180002641401, which corresponded to the female genome location of Chr2: 34,086,515-34,304,867 (Fig. 3D-E). Similarly, there is no protein-coding gene predicted in this male specific genomic segment. Therefore, these results confirmed that Chr18 of *C. argus* and Chr2 of *C. maculata* were potential sex chromosomes harboring the candidate SD region.

3.4. Comparison of SD region and sexual gene expression between C. argus and C. maculata

We compared the gene contents and sequences of sex chromosomes between *C. argus* and *C. maculata* to explore the difference in the SD region in the two species. As previously reported in a genome study, these two snakehead species harbored different chromosome numbers (24 and 21 in *C. argus* and *C. maculata*, respectively), among which Chr17 and 18 of *C. argus* correspond to Chr2 of *C. maculata* (Ou et al., 2021b). Through manual annotation and collinearity analysis, we showed that the sex chromosome, Chr18 of *C. argus* shared highly conserved genes with part of Chr2 in *C. maculata* (Fig. 4A). Moreover, we compared the sequences of male-specific insertions between contig jcf7180001325712 of *C. argus* and contig jcf7180002641401 of *C. maculata* through global alignment, which indicated that the ~4.6 kb insertion in *C. argus* showed high sequence identity with the ~4.0 kb insertion of *C. maculata*, with only a 5.86% difference (Fig. 4B).

Furthermore, RNA-seq reads generated from ovary and testis samples of *C. argus* and *C. maculata* were mapped to their own reference genome to detect the sexually differential gene expression pattern on sex chromosomes. In total, 128,762,479 reads were generated from six cDNA libraries of *C. argus*, and their mapping rates ranged from 96.00% to 96.89% (Table S5). The results showed that within the 5.38-Mb GWAS-informative genomic region on Chr18 of *C. argus*, a total of 36 genes were expressed significantly higher in the testis than in the ovary samples, while 36 genes were expressed higher in the ovary than in the testis (Fig. 5A). For *C. maculata*, a total of six transcriptomic sequencing libraries of ovary and testis containing 121,896,151 reads were used for gene expression analysis, revealing that 60 genes were up-regulated and 46 genes were down-regulated in testis compared with those in ovary within the 5.20-Mb GWAS-informative region on Chr2 (Fig. 5B). Among them, 56 genes showed significantly different expression in the ovary and testis in both *C. argus* and *C. maculata* (Fig. 5C, Table S6). Notably, among these genes that fell into the GWAS-informative region, many displayed highly consistent sexually differential expression patterns between the two species (Fig. 5D).

In the case of the Pool-seq results, only two protein coding genes, including ring finger protein 144a (*rnf144a*) and inhibitor of DNA binding 2 (*id2*) were identified within the ~190 kb and ~ 180 kb male-specific regions in *C. argus* and *C. maculata*, respectively, which were considered the most significantly sex-associated genes in the two snakeheads (Fig. 4A). The mRNA expression of *id2* in the testis was significantly lower than that in the ovary for both species, whereas the expression of *rnf144a* showed different trends in the two species, with a higher expression level in the testis of *C. argus* but a lower expression level in the testis of *C. maculata* (Fig. 5D). In addition, a highly conserved syntenic relationship was observed among snakeheads and other teleosts for this male-specific region (Fig. 5E), although their roles in sex determination have not been reported.

3.5. Development of sex-specific markers in C. argus, C. maculata and the hybrid (C. maculata ♀ × C. argus ♂)

As the sequences of male-specific regions between *C. argus* and *C. maculata* were highly conserved, it was possible to develop universal sex markers applicable to both species, as well as their commercial hybrid species (*C. maculata* ♀ × *C. argus* ♂). Based on the universal male-specific insertion fragments, we designed two pairs of sex-specific primers that could efficiently distinguish genetic males and females in these snakehead species (Table S1). These primers were initially verified by PCR-based tests using 45 phenotypic females and 45 males of *C. argus* collected from two different cultured populations (JN and LYG), 15 females and 15 males of *C. maculata* from FS, and 15 females and 15 males of hybrid snakeheads (*C. maculata* ♀ × *C. argus* ♂) generated from FS. Part of the PCR results are displayed in Fig. 6, which shows that one pair of primers (Marker 1) produced a single amplification product in males of *C. argus*, *C. maculata*, and hybrids, respectively, while no band was detected in all females (Fig. 6A). The other pair of primers (Marker 2) amplified two bands of PCR products in males of *C. argus* and *C. maculata*, while a single amplicon was detected in females (Fig. 6B). Interestingly, an additional band was amplified in both males (total of three bands) and females (total of two bands) of hybrids by using the

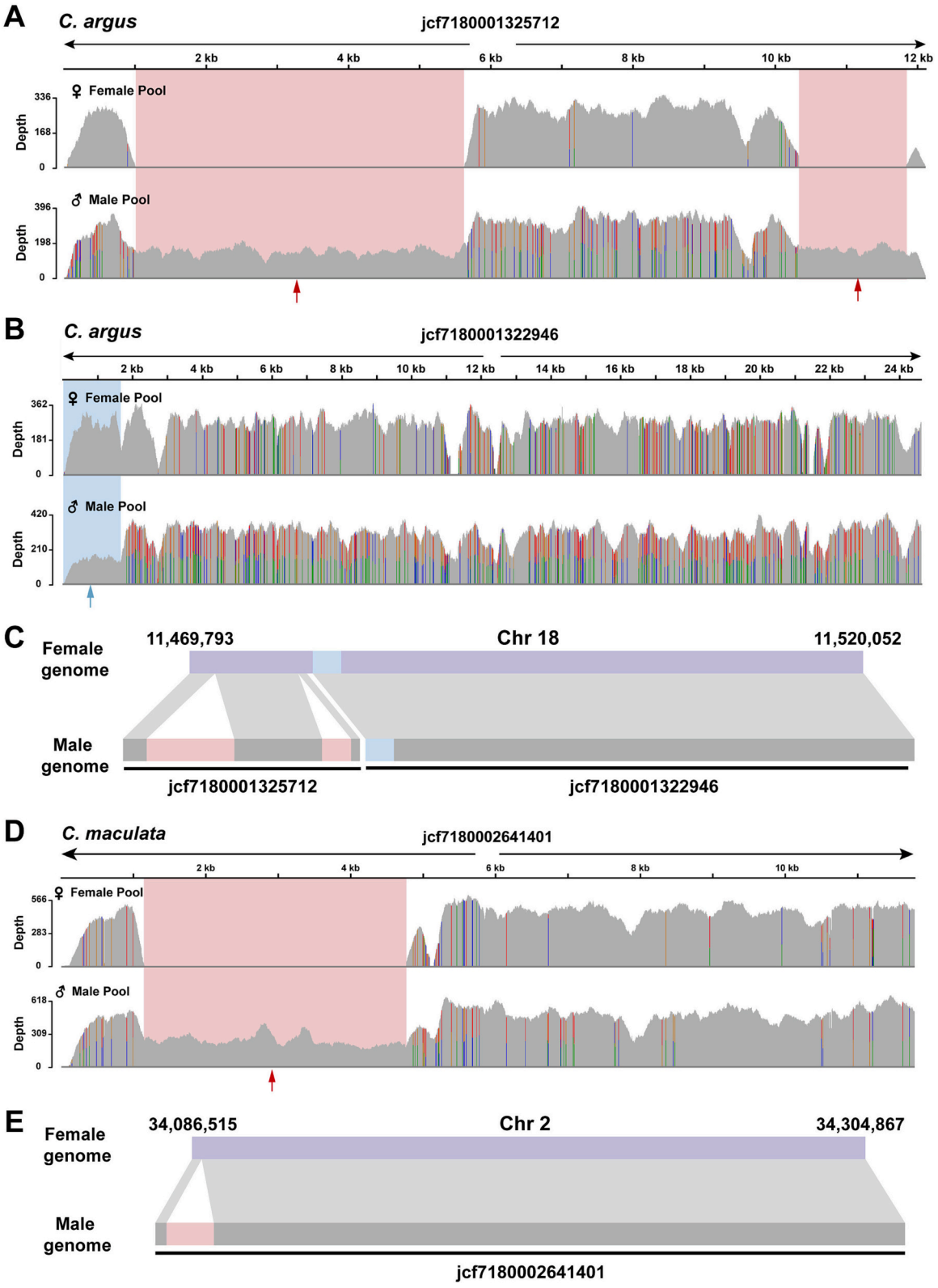


Fig. 3. Characterization of the sex-specific sequences in *C. argus* and *C. maculata*. A. Two male-specific (Y-chromosome specific) insertions (red shaded area) were observed in *C. argus*, with a region is considered male-specific if it is covered by none female reads and by male reads at a depth close to half of the genome average. B. ~1.7 kb X-chromosome specific genomic region (blue shaded area) in contig jcf7180001322946 was identified in *C. argus*, exhibiting sex-linked pattern that the sequence depth of male pool was nearly half that of female pool. C. Homology plot showing the difference of sex-chromosome (Chr18) and two sex-linked contigs between female and male genomes in *C. argus*. Male-specific insertions and the X-chromosome specific region are highlighted in red and blue, respectively. D. A ~ 4.0 kb male-specific insertion (red shaded area) was identified in contig jcf7180002641401 of the male genome in *C. maculata*. E. Homology plot showing the difference of sex-chromosome (Chr2) and sex-linked contigs between female and male genomes in *C. maculata*. The location of Chr2: 34086515–34,304,867 in the female genome of *C. maculata* corresponds to male-specific insertions, which are highlighted in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

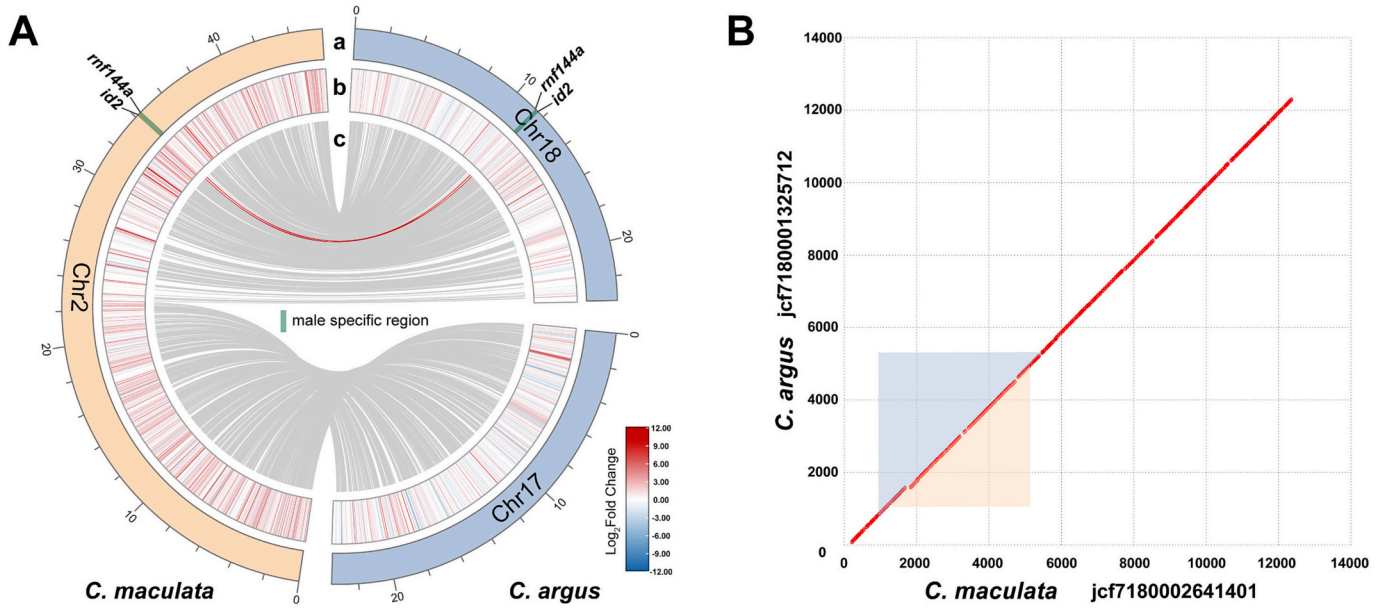


Fig. 4. Comparison of the gene content and sex-specific sequence in sex chromosomes between *C. argus* and *C. maculata*. A. Circos plots showing the syntenic relationships between genes in the sex chromosome of *C. argus* and *C. maculata*. The outer circle (a) represents the different chromosomes of two snakeheads. The line of the center circle (b) shows the expression difference of genes in the testis compared to the ovaries. Lines (c) linking two chromosomes indicate the location of homologues gene. Male specific region was highlighted by green colour. B. The dot plot shows the contig sequences including male-specific insertions compared between the male genome assemblies of *C. argus* and *C. maculata*. Male-specific insertions of *C. argus* (~4.6 kb) and *C. maculata* (~4.0 kb) are shaded in blue and orange, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

same primers, showing that this pair of primers (Marker 2) can not only be used to distinguish genetic sexes in snakeheads but can also be applied for distinguishing hybrid individuals from pure parental species (Fig. 6B). Overall, these findings confirmed the effectiveness and accuracy of our sex-specific primers and markers.

4. Discussion

Belonging to the Channidae family, *C. argus* and *C. maculata* are the main cultured snakehead species with important economic value in China. As males grow much faster than females, the production of all-male snakeheads has major economic implications in aquaculture (Ou et al., 2021a). Therefore, understanding their sex determination mechanism will not only promote scientific research on the evolution of sex chromosomes in vertebrates, but also facilitate mono-sex breeding progress which could dramatically increase profits in the snakehead industry. In the current study, we identified the most likely candidate SD region which was restricted to a genomic segment <200 kb for both *C. argus* and *C. maculata*, and established an efficient PCR test for sex genotyping in these two parental species and their hybrid (*C. maculata* ♀ × *C. argus* ♂) based on the sequence differences between sexes.

In comparison with the relatively conserved SD systems in mammals (XX/XY system) and birds (ZZ/ZW system), teleosts have evolved a great variety of SD mechanisms with distinct SD systems including genetic SD (GSD), environmental SD (ESD) and a combination of GSD and ESD (Guiguen et al., 2018), and frequent turnover of master SD (MSD) genes. For *C. argus*, previous QTL mapping studies on sex traits revealed a major QTL for sex determination with a percentage variance explained (PVE) over 90% (Liu et al., 2020; Wang et al., 2019), indicating the sex of snakehead was controlled by a major genetic factor. Additionally, SNP markers within the sex-linked QTL region exhibited heterozygous patterns in males and homozygous patterns in females, suggesting an XX/XY SD system for snakehead (Liu et al., 2020). These findings were also confirmed by GWAS using WGS data in our study. In addition, in combination with the newly assembled chromosome-level genome of *C. argus* and *C. maculata* in the last two years, we further narrowed the candidate SD region of both snakehead species down to a genomic segment <200 kb by Pool-Seq analysis, which has been successfully applied in characterizing the SD region in largemouth bass (*Micropterus Salmoides* L.) (Wen et al., 2022a), channel catfish (Wen et al., 2023), goldfish (*Siniperca chuatsi*) (Wen et al., 2020), mandarin fishes (*Siniperca chuatsi*) (Wen et al., 2022b), Pacific halibut (*Hippoglossus stenolepis*) (Jasonowicz et al., 2022) and mosquitofish (*Gambusia affinis* and *G. holbrooki*) (Kottler et al., 2020).

The fast evolution of the sex chromosome was suggested to depend on the origin and fixation of a new MSD gene (Schartl, 2004). With the rapid development of high-throughput sequencing technologies, several candidate MSD genes have been identified in various teleosts, such as *dmy*, *gsdfy*, *amhy* and anti-Müllerian hormone receptor 2 (*amhr2*) in species with the XX/XY SD system (Edvardsen et al., 2022; Hattori et al., 2022; Matsuda et al., 2002, 2003; Song et al., 2021; Zheng et al., 2022), as well as *dmrt1* and 17b-hydroxysteroid dehydrogenase 1 (*hsd17b1*) in fishes with the ZZ/ZW system (Chen et al., 2014; Cui et al., 2017; Fan et al., 2021). Previously, almost all MSD genes were single gene duplicates or allelic variants of genes known to be related to sex determination and differentiation (Herpin and Schartl, 2015). In recent years, genes that have not been implicated in sexual development were characterized as novel candidate MSD genes, such as the Y-chromosome (*sdY*) gene in salmonids (Bertho et al., 2018), the *bcar1* gene in channel catfish (Bao et al., 2019), and the *rnf183* gene in little yellow croaker (*Larimichthys polyactis*) (Xie et al., 2022). Within the ~200 kb candidate SD region in snakehead genomes generated by Pool-Seq analysis, only two protein-coding genes, *rnf144a* and *id2*, were annotated, and their function in the SD system has never been demonstrated. Although the mRNA of *rnf144a* and *id2* displayed differential expression patterns in immature gonad tissues between sexes, it is still too early to determine whether both or one of them are MSD genes in snakeheads.

Id2 protein belongs to the inhibitor of DNA binding (ID) family, members of which inhibit the functions of basic helix-loop-helix (bHLH) transcription factors in a dominant-negative manner by suppressing their heterodimerization partners, and may play a role in negatively regulating cell differentiation (Norton et al., 1998; Yang et al., 2015).

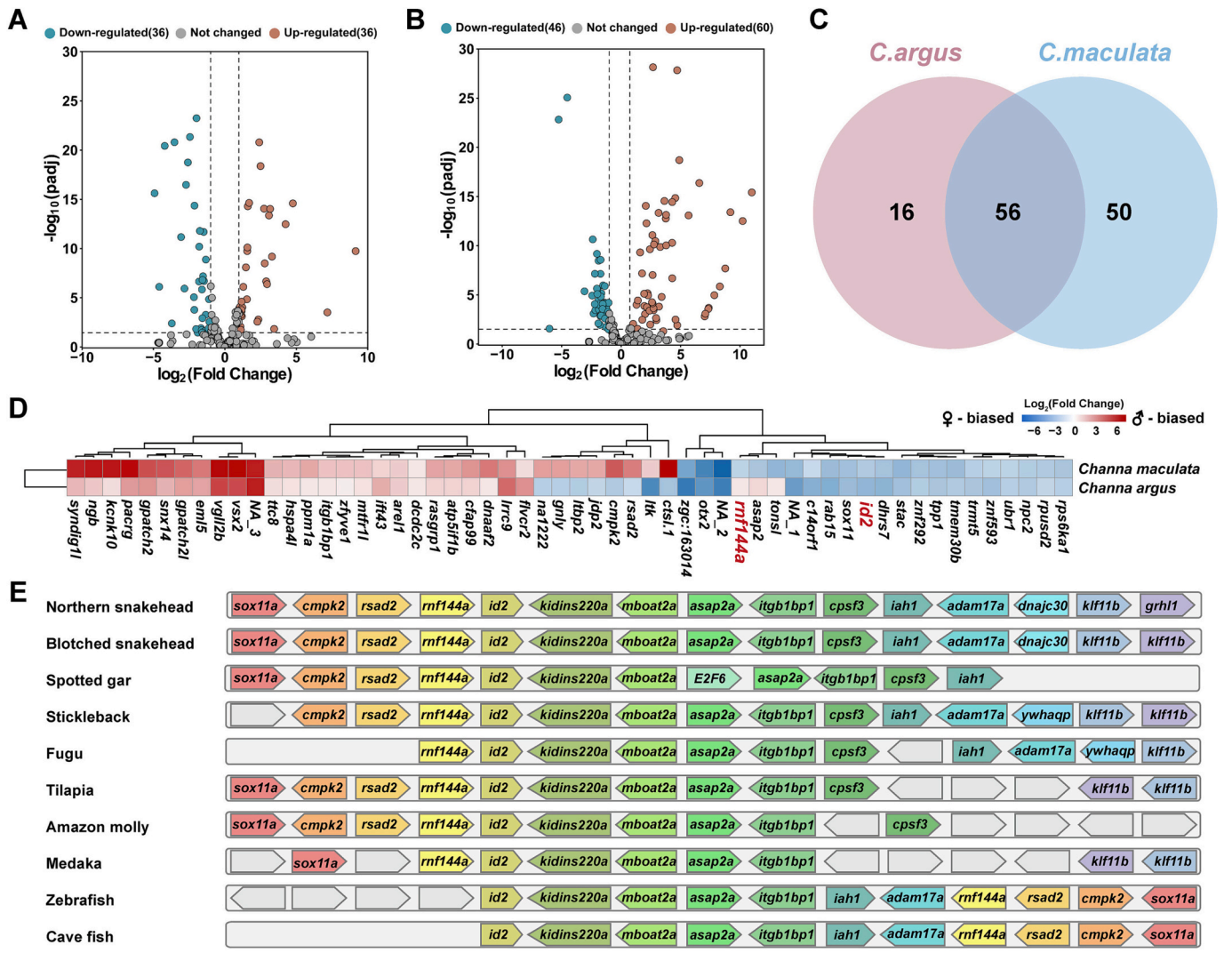


Fig. 5. Sexually differential gene expression patterns on sex chromosomes. A. Volcano plot showing DEGs on the GWAS-informative genomic region between ovary and testis in *C. argus*. B. Volcano plot showing DEGs on the GWAS-informative genomic region between ovary and testis in *C. maculata*. Red dots and blue dots indicate the significantly up-regulated and down-regulated genes in testis, respectively. Grey dots indicate genes that were not expressed differently between ovary and testis. C. Venn diagram displays number of DEGs between ovary and testis locating in the GWAS-informative genomic region in *C. argus* and *C. maculata*. There are 56 DEGs shared between the two species, which are regarded as sex-biased expression genes. D. Heatmaps and hierarchical clustering of sex-biased expression genes in two snakeheads. Blue represents genes expressing higher in ovary, and red represents higher expression genes in testis. E. Syntenic analyses of *rnf144a*, *id2*, and their adjacent genes among snakeheads and other teleosts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Recently, in a study of arapaima (*Arapaima gigas*), the *id2bbY* gene (a duplicated copy of *id2* gene on the Y chromosome) was recognized as a candidate MSD gene and a male-specific marker, although its role in the SD system has not been confirmed by genetic manipulation experiments. The authors hypothesized that *id2bbY* might function as the downstream gene of the TGF-β signalling pathway containing several sex-determining components (Adolfi et al., 2021). Additionally, several studies have shown that ID proteins are involved in spermatogenesis, granulosa cell differentiation, ovarian follicle differentiation and ovary maturation (Clelland and Kelly, 2011; da Silveira et al., 2014; Johnson et al., 2008; Li et al., 2005). Another candidate MSD gene of snakeheads, *rnf144a*, encodes an E3 ubiquitin ligase for DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and can promote DNA damage-induced cell apoptosis (Ho et al., 2015). Although studies on its roles in teleosts have rarely been reported, it is clear that RNF proteins act as regulators and engage in diverse aspects of biological activities, including growth, development and stress responses (Cai et al., 2022; Joazeiro and Weissman, 2000). A member of the RNF family, *rnf183* was speculated to be involved in sex determination in little yellow croaker. Through syntenic analysis, the gene contents surrounding the sex-specific region of snakeheads were highly conserved among different teleost species but not homologous to the SD locus of any fish species that have been reported to have sex chromosomes. Therefore, it is reasonable to assume that the *id2* and/or *rnf144a* gene might be a novel candidate MSD gene in teleosts, but the role and regulatory mechanism in the SD system must be investigated by gene-editing experiments and studying gene expression profiles in gonads during sex determination and differentiation stages.

Alternatively, we identified a 4–5 kb male-specific genomic insertion in both *C. argus* and *C. maculata*. Although it does not contain any protein-coding genes, the role of this sex-specific noncoding genomic region cannot be underestimated, and we have predicted transposable elements (TEs) existing in this fragment. Understanding the role of epigenetic regulation, such as TEs and noncoding RNAs (ncRNAs), in sex determination and differentiation has made substantial strides in recent years (Akagi et al., 2016; Chalopin et al., 2015; Wang et al., 2022; Weber

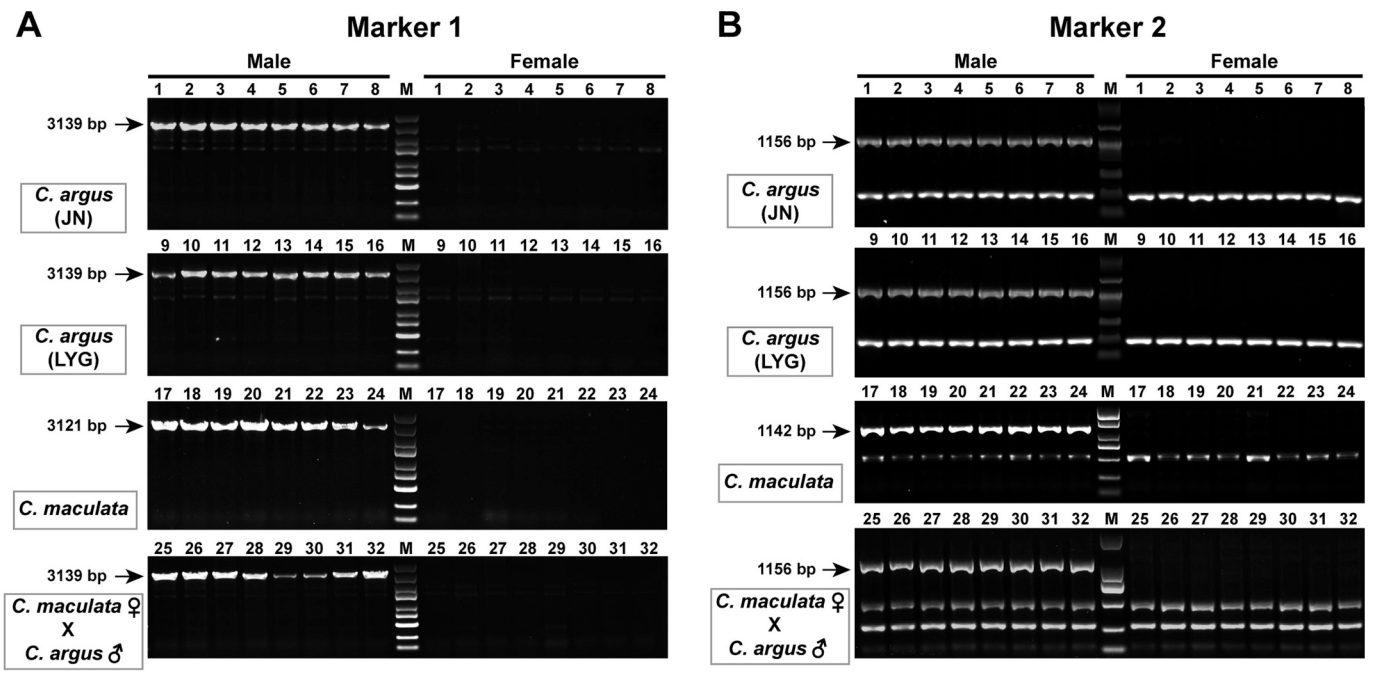


Fig. 6. PCR detection of universal sex-specific markers in *C. argus*, *C. maculata* and their hybrids species (*C. maculate* ♀ × *C. argus* ♂). A. In male individuals of *C. argus*, *C. maculata*, and hybrids, Marker 1 produced single amplification product, and no band was detected in all females. B. In male *C. argus* and *C. maculata* individuals, Marker 2 amplified two bands of the PCR products, whereas only one amplicon was detected in females. Meanwhile, additional band was amplified in both males (total of three bands) and females (total of two bands) of their hybrids. JN and LYG represent two geographically distinct populations, namely Jining and Lianyungang.

et al., 2020). Repetitive elements, especially transposons, have been shown to play a central role in SD by causing insertion and duplication (Bachtrog, 2005; Bachtrog et al., 2008; Ding et al., 2021; Natri et al., 2013). Comparative analyses of the SD loci in the three salmonid species revealed that a 4.1 kb orthologous region shared by all three species contains the genetic material required for masculinization, along with transposable elements (Faber-Hammond et al., 2015). In melon (*Cucumis melo L.*), TE-induced methylation of the promoter of the transcription factor *CmWIP1* indirectly represses the expression of the andromonoecious gene *CmACS-7* to allow the development of stamens (Martin et al., 2009). Recently, similar mechanisms have been found in teleosts in which the transposable element *drbx1* is inserted into an intron of the X-linked region encoding the SD gene *dmrt1* in fighting fish. During the crucial stage of sex determination, this structural modification was associated with a shift in the epigenetic silencing of X-*dmrt1* (Wang et al., 2022). Piwi-interacting RNAs (piRNAs) are 24–31 nucleotide-long small noncoding RNAs expressed in the germline and derived from long RNAs that contain TE sequences (Iwasaki et al., 2015). A unique and interesting case of an mRNA-regulating piRNA has been reported in the silkworm (*Bombyx mori*), confirming that the piRNA molecule is critical in the primary sex determination mechanism (Kiuchi et al., 2014). Moreover, epigenetic modifications, including genomic DNA methylation and histone modification, have been confirmed to control sex determination by regulating the expression of sex-determining genes in fish species (Ge et al., 2018; Shao et al., 2014; Weber et al., 2020). Therefore, the existence of a novel sex determination gene/element driving the evolution of the SD system in the male-specific genome insertion fragment of snakehead fish cannot be ruled out.

In the past decade, an increasing number of sex-specific markers have been developed in economically important aquaculture species using a variety of techniques (Han et al., 2023; Liu et al., 2023; Purcell et al., 2018; Wang et al., 2019; Wen et al., 2023). Although a few studies have tried to identify DNA markers for sexing in the two snakehead species, the reported sex-specific markers were either population-specific or not available for verification (Liu et al., 2011; Ou et al., 2017; Wang et al., 2019). Furthermore, there is no universal sex marker that can be used for both *C. argus* and *C. maculata*. We developed, for the first time, a universal sex-specific marker that can effectively distinguish the sexes of *C. argus* and *C. maculata*, as well as their hybrids. Moreover, we generated a pair of primers that was not only able to distinguish genetic sexes in snakeheads but also able to distinguish hybrid individuals from their parental species. Therefore, the successful development of genetic markers in this study will accelerate the progress of mono-sex breeding of snakeheads for commercial production, and help to accurately identify hybrids and pure snakehead species.

5. Conclusions

In this study, we investigated sex-linked genomic variations in both *C. maculata* and *C. argus* and identified the most likely SD region candidates, which were restricted to a genomic segment of <200 kb and contained two protein-coding genes (*rnf144a* and *id2*). In addition, a 4–5 kb male-specific noncoding genomic insertion within the sex-linked region was identified in both species, which enabled us to generate a universal sex-specific marker that can effectively distinguish the sexes of two snakeheads, as well as their hybrids. Our study will not only facilitate understanding of the mechanism of sex determination in snakeheads, but also promote mono-sex breeding progress that would increase the profits in their aquaculture industry.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2023.739771.

CRediT authorship contribution statement

Donglei Sun: Writing – original draft, Methodology, Formal analysis, Software, Validation, Visualization. **Haishen Wen:** Conceptualization, Supervision. **Xin Qi:** Conceptualization, Supervision. **Chao Li:** Funding acquisition, Investigation. **Chaonan Sun:** Validation, Visualization. **Lingyu Wang:** Investigation, Resources. **Mingxin Zhu:**

Investigation, Resources. **Tianyu Jiang:** Investigation, Resources. **Xiaoyan Zhang:** Investigation, Resources. **Yun Li:** Conceptualization, Methodology, 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

Whole genome resequencing reads from male and female C. argus have been deposited in the Sequence Read Archive (SRA), under BioProject PRJNA834927. Data will be made available on request.

Acknowledgements

This work was supported by Shandong Technical System of Fish Industry (SDAIT-12-03). We thank professor Jian Zhao from Chinese Academy of Fishery Sciences, Pearl River Fisheries Research Institute for providing the transcriptome data of *C. maculata*.

References

Adolfi, M.C., Du, K., Kneitz, S., Cabau, C., Zahm, M., Klopp, C., Feron, R., Paixão, R.V., Varela, E.S., de Almeida, F.L., de Oliveira, M.A., Nóbrega, R.H., Lopez-Roques, C., Iampietro, C., Lluch, J., Kloas, W., Wuertz, S., Schaefer, F., Stöck, M., Guiguen, Y., Schartl, M., 2021. A duplicated copy of *id2b* is an unusual sex-determining candidate gene on the Y chromosome of arapaima (*Arapaima gigas*). Sci. Rep. 11 (1), 21544.

Akagi, T., Henry, I.M., Kawai, T., Comai, L., Tao, R., 2016. Epigenetic regulation of the sex determination gene MeGI in polyploid persimmon. Plant Cell 28 (12), 2905–2915.

Bachtrog, D., 2005. Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in Drosophila. Genome Res. 15 (10), 1393–1401.

Bachtrog, D., Hom, E., Wong, K.M., Maside, X., de Jong, P., 2008. Genomic degradation of a young Y chromosome in *Drosophila miranda*. Genome Biol. 9, 1–10.

Bao, L.S., Tian, C.X., Liu, S.K., Zhang, Y., Elaswad, A., Yuan, Z.H., Khalil, K., Sun, F.Y., Yang, Y.J., Zhou, T., et al., 2019. The Y chromosome sequence of the channel catfish suggests novel sex determination mechanisms in teleost fish. BMC Biol. 17, 1–16.

Bertho, S., Herpin, A., Branthonne, A., Jouanno, E., Yano, A., Nicol, B., Muller, T., Pannetier, M., Pailhoux, E., Miwa, M., Yoshizaki, G., Schartl, M., Guiguen, Y., 2018. The unusual rainbow trout sex determination gene hijacked the canonical vertebrate gonadal differentiation pathway. Proc. Natl. Acad. Sci. U. S. A. 115 (50), 12781–12786.

Cai, C., Tang, Y.D., Zhai, J., Zheng, C., 2022. The RING finger protein family in health and disease. Signal Transduct Target Ther 7 (1), 300.

Chalopin, D., Volff, J.N., Galiana, D., Anderson, J.L., Schartl, M., 2015. Transposable elements and early evolution of sex chromosomes in fish. Chromosom. Res. 23 (3), 545–560.

Chen, S.L., Zhang, G.J., Shao, C.W., Huang, Q.F., Liu, G., Zhang, P., Song, W.T., An, N., Chalopin, D., Volff, J.N., et al., 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat. Genet. 46 (3), 253–260.

Clelland, E.S., Kelly, S.P., 2011. Exogenous GDF9 but not Activin A, BMP15 or TGFβ alters tight junction protein transcript abundance in zebrafish ovarian follicles. Gen. Comp. Endocrinol. 171 (2), 211–217.

Cui, Z.K., Liu, Y., Wang, W.W., Wang, Q., Zhang, N., Lin, F., Wang, N., Shao, C.W., Dong, Z.D., Li, Y.Z., 2017. Genome editing reveals *dmrt1* as an essential male sex-determining gene in Chinese tongue sole (*Cynoglossus semilaevis*). Sci. Rep. 7 (1), 42213.

da Silveira, J.C., Carnevale, E.M., Winger, Q.A., Bouma, G.J., 2014. Regulation of ACVR1 and ID2 by cell-secreted exosomes during follicle maturation in the mare. Reprod. Biol. Endocrinol. 12, 44.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al., 2011. The variant call format and VCFtools. Bioinformatics 27 (15), 2156–2158.

Ding, M., Li, X.Y., Zhu, Z.X., Chen, J.H., Lu, M., Shi, Q., Wang, Y., Li, Z., Zhao, X., Wang, T., et al., 2021. Genomic anatomy of male-specific microchromosomes in a gynogenetic fish. PLoS Genet. 17 (9), e1009760.

Edvardsen, R.B., Wallerman, O., Furmanek, T., Kleppe, L., Jern, P., Wallberg, A., Kjærner-Semb, E., Mæhle, S., Olausson, S.K., Sundström, E., et al., 2022. Heterochiasmy and the establishment of *gsdf* as a novel sex determining gene in Atlantic halibut. PLoS Genet. 18 (2), e1010011.

Faber-Hammond, J.J., Phillips, R.B., Brown, K.H., 2015. Comparative analysis of the shared sex-determination region (SDR) among salmonid fishes. Genome Biol. Evol. 7 (7), 1972–1987.

Fan, B., Xie, D.Z., Li, Y.W., Wang, X.L., Qi, X., Li, S.S., Meng, Z.N., Chen, X.H., Peng, J.Y., Yang, Y.J., et al., 2021. A single intronic single nucleotide polymorphism in splicing site of steroidogenic enzyme *hsd17b1* is associated with phenotypic sex in oyster pompano, *Trachinotus anak*. Proc. Royal Soc. B 288 (1963), 20212245.

Ge, C.T., Ye, J., Weber, C., Sun, W., Zhang, H.Y., Zhou, Y.J., Cai, C., Qian, G.Y., Capel, B., 2018. The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. Science 360 (6389), 645–648.

Guiguen, Y., Fostier, A., Herpin, A., 2018. Sex Determination and Differentiation in Fish: Genetic, Genomic, and Endocrine Aspects. Sex Control in Aquaculture. Wiley-Blackwell, Hoboken, NJ, pp. 35–63.

Han, C., Huang, W.W., Peng, S.H., Zhou, J.W., Zhan, H.W., Gui, L., Li, W.J., Li, Q., 2023. Screening and characterization of sex-specific markers by NGS sequencing in *Spinibarbus hollandi* with implication of XY sex determination system. Aquaculture 565, 739147.

Hattori, R.S., Kumazawa, K., Nakamoto, M., Nakano, Y., Yamaguchi, T., Kitano, T., Yamamoto, E., Fuji, K., Sakamoto, T., 2022. Y-specific amh allele, amhy, is the master sex-determining gene in Japanese flounder *Paralichthys olivaceus*. Front. Genet. 13, 1007548.

Herpin, A., Schartl, M., 2015. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 16 (10), 1260–1274.

Ho, S.R., Lee, Y.J., Lin, W.C., 2015. Regulation of RNF144A E3 ubiquitin ligase activity by self-association through its transmembrane domain. J. Biol. Chem. 290 (38), 23026–23038.

Iwasaki, Y.W., Siomi, M.C., Siomi, H., 2015. PIWI-interacting RNA: its biogenesis and functions. Annu. Rev. Biochem. 84, 405–433.

Jasonowicz, A.J., Simeon, A., Zahm, M., Cabau, C., Klopp, C., Roques, C., Iampietro, C., Lluch, J., Donnadieu, C., Parrinello, H., et al., 2022. Generation of a chromosome-level genome assembly for Pacific halibut (*Hippoglossus stenolepis*) and characterization of its sex-determining genomic region. Mol. Ecol. Resour. 22 (7), 2685–2700.

Joazeiro, C.A., Weissman, A.M., 2000. RING finger proteins: mediators of ubiquitin ligase activity. Cell 102 (5), 549–552.

Johnson, A.L., Haugen, M.J., Woods, D.C., 2008. Role for inhibitor of differentiation/deoxyribonucleic acid-binding (id) proteins in granulosa cell differentiation. Endocrinology 149 (6), 3187–3195.

Kim, D., Langmead, B., Salzberg, S.L., 2015. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12 (4), 357–360.

Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y., Ishihara, G., Kawaoka, S., Sugano, S., Shimada, T., Suzuki, Y., Suzuki, M.G., Katsuma, S., 2014. A single female-specific piRNA is the primary determiner of sex in the silkworm. Nature 509 (7502), 633–636.

Kottler, V.A., Feron, R., Nanda, I., Klopp, C., Du, K., Kneitz, S., Helmprobst, F., Lamatsch, D.K., Lopez-Roques, C., Lluch, J., Journot, L., Parrinello, H., Guiguen, Y., Schartl, M., 2020. Independent origin of XY and ZW sex determination mechanisms in mosquitofish sister species. Genetics 214 (1), 193–209.

Kwon, Y.M., Vranken, N., Hoge, C., Lichak, M.R., Norovich, A.L., Francis, K.X., Camacho-Garcia, J., Bista, I., Wood, J., McCarthy, S., et al., 2022. Genomic consequences of domestication of the Siamese fighting fish. Sci. Adv. 8 (10), eabm4950.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25 (14), 1754–1760.

Li, B., Nair, M., Mackay, D.R., Bilanchone, V., Hu, M., Fallahi, M., Song, H., Dai, Q., Cohen, P.E., Dai, X., 2005. Ovol1 regulates meiotic pachytene progression during spermatogenesis by repressing Id2 expression. Development 132 (6), 1463–1473.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25 (16), 2078–2079.

Liao, Y., Smyth, G.K., Shi, W., 2013. The subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 41 (10), e108.

Liu, G., Chen, K., Zheng, G., Zhu, X., Zhao, J., Xu, P., Sun, X., 2011. Screening and identification of female-specific DNA fragments in *Channa argus* using SSR-BSA. J. Fish. China 35 (2), 170–175.

Liu, H.Y., Luo, Q., Ou, M., Zhu, X.P., Zhao, J., Chen, K.C., 2020. High-density genetic linkage map and QTL fine mapping of growth and sex in snakehead (*Channa argus*). Aquaculture 519, 734760.

Liu, S.Y., Han, C., Huang, J.J., Zhu, Q.Y., Liu, D.R., Han, L.Q., Li, S.S., Li, G.F., Lin, H.R., Zhang, Y., 2023. Screening and characterization of X chromosome-specific markers in mandarin fish (*Siniperca chuatsi*). Aquaculture 562, 738833.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (12), 550.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C., Bendahmane, A., 2009. A transposon-induced epigenetic change leads to sex determination in melon. Nature 461 (7267), 1135–1138.

Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C.E., Shibata, N., Asakawa, S., Shimizu, N., et al., 2002. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417 (6888), 559–563.

Matsuda, M., Nagahama, Y., Kobayashi, T., Matsuda, C., Hamaguchi, S., Sakaizumi, M., 2003. The sex determining gene of medaka: a Y-specific DM domain gene (DMY) is required for male development. Fish Physiol. Biochem. 28, 135–139.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (9), 1297–1303.

Natri, H.M., Shikano, T., Merilä, J., 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. Mol. Biol. Evol. 30 (5), 1131–1144.
Norton, J.D., Deed, R.W., Craggs, G., Sablitzky, F., 1998. Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol. 8 (2), 58–65.
Ou, M., Yang, C., Luo, Q., Huang, R., Zhang, A.D., Liao, L.J., Li, Y.M., He, L.B., Zhu, Z.Y., Chen, K.C., Wang, Y.P., 2017. An NGS-based approach for the identification of sex-specific markers in snakehead (*Channa argus*). Oncotarget 8 (58), 98733.
Ou, M., Zhao, J., Luo, Q., Hong, X.Y., Zhu, X.P., Liu, H.Y., Chen, K.C., 2018. Characteristics of hybrids derived from *Channa argus*♀× *Channa maculata*♂. Aquaculture 492, 349–356.
Ou, M., Chen, K.C., Gao, D.D., Wu, Y.D., Chen, Z., Luo, Q., Liu, H.Y., Zhao, J., 2020. Comparative transcriptome analysis on four types of gonadal tissues of blotched snakehead (*Channa maculata*). Comp. Biochem. Physiol. Part D Genomics Proteomics 35, 100708.
Ou, M., Chen, K.C., Luo, Q., Liu, H.Y., Wang, Y.P., Chen, B.X., Liang, X.Q., Zhao, J., 2021a. Performance evaluation of XY all-male hybrids derived from XX female Channa argus and YY super-males *Channa maculata*. Aquac. Rep. 20, 100768.
Ou, M., Huang, R., Yang, C., Gui, B., Luo, Q., Zhao, J., Li, Y.M., Liao, L.J., Zhu, Z.Y., Wang, Y.P., Chen, K.C., 2021b. Chromosome-level genome assemblies of *Channa argus* and *Channa maculata* and comparative analysis of their temperature adaptability. GigaScience 10 (10), giab070.
Purcell, C.M., Seetharam, A.S., Snodgrass, O., Ortega-García, S., Hyde, J.R., Severin, A.J., 2018. Insights into teleost sex determination from the Seriola dorsalis genome assembly. BMC Genom. 19 (1), 31.
Schartl, M., 2004. Sex chromosome evolution in non-mammalian vertebrates. Curr. Opin. Genet. Dev. 14 (6), 634–641.
Shao, C.W., Li, Q.Y., Chen, S.L., Zhang, P., Lian, J.M., Hu, Q.M., Sun, B., Jin, L.J., Liu, S. S., Wang, Z.J., et al., 2014. Epigenetic modification and inheritance in sexual reversal of fish. Genome Res. 24 (4), 604–615.
Song, W.H., Xie, Y.H., Sun, M.M., Li, X.M., Fitzpatrick, C.K., Vaux, F., O'Malley, K.G., Zhang, Q.Q., Qi, J., He, Y., 2021. A duplicated *amh* is the master sex-determining gene for Sebastes rockfish in the Northwest Pacific. Open Biol. 11 (7), 210063.
Sun, D.L., Qi, X., Wen, H.S., Li, C., Li, J.L., Chen, J.W., Tao, Z.X., Zhu, M.X., Zhang, X.Y., Li, Y., 2023. The genetic basis and potential molecular mechanism of yellow-albino northern snakehead (*Channa argus*). Open Biol. 13 (2), 220235.
Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variations from next-generation sequencing data. Nucleic Acids Res. 38, e164.
Wang, L., Xie, N., Shen, Y., Ye, B., Yue, G.H., Feng, X., 2019. Constructing high-density genetic maps and developing sexing markers in northern snakehead (*Channa argus*). Mar. Biotechnol. 21 (3), 348–358.
Wang, L., Sun, F., Wan, Z.Y., Yang, Z., Tay, Y.X., Lee, M., Ye, B., Wen, Y., Meng, Z., Fan, B., et al., 2022. Transposon-induced epigenetic silencing in the X chromosome as a novel form of dmrt1 expression regulation during sex determination in the fighting fish. BMC Biol. 20 (1), 1–16.
Weber, C., Zhou, Y., Lee, J.G., Looger, L.L., Qian, G., Ge, C., Capel, B., 2020. Temperature-dependent sex determination is mediated by pSTAT3 repression of Kdm6b. Science 368 (6488), 303–306.
Wen, M., Feron, R., Pan, Q., Guguin, J., Jouanno, E., Herpin, A., Klopp, C., Cabau, C., Zahm, M., Parrinello, H., et al., 2020. Sex chromosome and sex locus characterization in goldfish, *Carassius auratus* (Linnaeus, 1758). BMC Genomics 21 (1), 1–12.
Wen, M., Zhang, Y.X., Wang, S.Y., Hu, F.Z., Tang, C.J., Li, Q., Qin, Q.B., Tao, M., Zhang, C., Zhao, R.R., et al., 2022a. Sex locus and sex markers identification using whole genome pool-sequencing approach in the largemouth bass (*Micropterus Salmoides L.*). Aquaculture 559, 738375.
Wen, M., Zhang, Y.X., Wang, S.Y., Wang, Y.D., Qin, Q.B., Tao, M., Zhang, C., Zhao, R.R., Wang, S., Hu, F.Z., et al., 2022b. Characterization of sex locus and sex-specific sequences in the mandarin fishes. Aquaculture 561, 738650.
Wen, M., Pan, Q., Larson, W., Eché, C., Guiguen, Y., 2023. Characterization of the sex determining region of channel catfish (*Ictalurus punctatus*) and development of a sex-genotyping test. Gene 850, 146933.
Wu, W.X., Chen, H.X., Zhuang, H., 1986. Studies on karyotypes in five species of fish. Acta Scientiarum Naturalium Universitatis Sunyatseni 25 (2), 107.
Xie, Q.P., Zhan, W., Shi, J.Z., Liu, F., Niu, B.L., He, X., Liu, M., Wang, J., Liang, Q.Q., Xie, Y., et al., 2022. Whole-genome assembly and annotation for the little yellow croaker (*Larimichthys polyactis*) provide insights into the evolution of hermaphroditism and gonochorism. Mol. Ecol. Resour. 23 (3), 632–658.
Yang, J., Velikoff, M., Agarwal, M., Disayabutr, S., Wolters, P.J., Kim, K.K., 2015. Overexpression of inhibitor of DNA-binding 2 attenuates pulmonary fibrosis through regulation of c-Abl and twist. Am. J. Pathol. 185 (4), 1001–1011.
Yang, C., Huang, R., Ou, M., Gui, B., Zhao, J., He, L.B., Li, Y.M., Liao, L.J., Chen, K.C., Wang, Y.P., 2020. A rapid method of sex-specific marker discovery based on NGS and determination of the XX/XY sex-determination system in *Channa maculata*. Aquaculture 528, 735499.
Zhang, W.C., Wang, H.R., Brandt, D.Y., Hu, B.J., Sheng, J.Q., Wang, M.G., Luo, H.J., Li, Y.H., Guo, S.J., Sheng, B., et al., 2022. The genetic architecture of phenotypic diversity in the betta fish (*Betta splendens*). Sci. Adv. 8 (38), eabm4955.
Zhao, J., Ou, M., Wang, Y.P., Liu, H.Y., Luo, Q., Zhu, X.P., Chen, B.X., Chen, K.C., 2021. Breeding of YY super-male of blotched snakehead (*Channa maculata*) and production of all-male hybrid (*Channa argus*♀× *C. maculata*♂). Aquaculture 538, 736450.
Zheng, S., Tao, W., Yang, H., Kocher, T.D., Wang, Z., Peng, Z., Jin, L., Pu, D., Zhang, Y., Wang, D., 2022. Identification of sex chromosome and sex-determining gene of southern catfish (*Silurus meridionalis*) based on XX, XY and YY genome sequencing. Proc. Royal Soc. B 289 (1971), 20212645.
Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., Yorke, J.A., 2013. The MaSuRCA genome assembler. Bioinformatics 29 (21), 2669–2677.