Whole-Genome Resequencing Reveals the Genetic Diversity and Adaptive Mechanism of *Mastacembelus armatus* in the Dongjiang and Ganjiang River Sources

Wu Bin , Fang Yuan , Zeng Qingxiang , Li Han , Wang Haihua *

Posted Date: 17 April 2026

doi: 10.20944/preprints202604.1282.v1

Keywords: *Mastacembelus armatus*; whole-genome resequencing; genetic diversity; Dongjiang river source; Ganjiang river source

 doi:10.20944/preprints202604.1282.v1

Article

Whole-Genome Resequencing Reveals the Genetic Diversity and Adaptive Mechanism of *Mastacembelus armatus* in the Dongjiang and Ganjiang River Sources

Wu Bin [1], Fang Yuan [2], Zeng Qingxiang [2], Li Han [1] and Wang Haihua [1,*]

[1] Jiangxi Fisheries Research Institute, Nanchang, China

[2] Ganzhou Animal Husbandry and Fisheries Research Institute, Ganzhou, China

* Correspondence: haihuawang998@sina.com.cn

Simple Summary

Wild zig-zag eel (*Mastacembelus armatus*) populations are declining due to environmental damage and overfishing, making their conservation urgent. This study used whole-genome resequencing to examine the genetic diversity and adaptation of three zig-zag eel groups from the Dongjiang River source and Ganjiang River source in southern China. We found that all three groups belong to the same evolutionary lineage but show clear genetic differences. The Xunwushui and Jiuqu River groups have high genetic diversity and close genetic relationships, while the Taojiang River group has lower genetic diversity and distinct genetic features. These differences are mainly caused by geographic isolation and limited gene flow between river systems. Our results provide genome-level guidance for protecting the genetic resources of zig-zag eels, developing effective conservation plans, and supporting the sustainable management of this important freshwater fish.

Abstract

To explore the genetic diversity and adaptive evolutionary mechanism of *Mastacembelus armatus* in the Dongjiang and Ganjiang River Sources, whole-genome resequencing was performed on three populations of *M. armatus* from Xunwushui (XW) and Jiuqu River (DN) in the Dongjiang River Source, and Taojiang (XF) in the Ganjiang River Source. Population genetics methods were integrated to analyze their genetic structure, differentiation characteristics and selection signals. The results showed that a total of 209.05 Gbp of Clean Data was obtained from the three populations, with the Q30 base percentage reaching 94.42% and the average mapping rate to the reference genome being 97.85%, indicating high reliability of the sequencing data. A mean of 7,459,686 single nucleotide polymorphisms (SNPs) were detected, with a transition/transversion ratio of 1.52 and a heterozygous SNP ratio of 2.22%. The total number of genome-wide insertions and deletions (InDels) was 1,902,722±23,247. Gene Ontology (GO) functional annotation revealed a consistent variation pattern of core genes among the three populations. Phylogenetic tree, Admixture and principal component analysis (PCA) confirmed that the three populations belonged to a single evolutionary clade and shared a genetic origin from two ancestral populations (the lowest cross-validation error at K=2), while significant genetic differentiation was observed among populations: XW and DN populations had similar genetic backgrounds and closer genetic relationships, both biased towards the blue ancestral component, whereas XF population was inclined to the red ancestral component, with the DN population showing the highest degree of genetic admixture. Individuals within the XF population had more distant genetic relationships and the longest linkage disequilibrium (LD) decay distance, which was speculated to be associated with its small population size and low recombination rate; in contrast, the XW population had the shortest LD decay distance, corresponding to the characteristics of large population size and high recombination rate. Analysis of population genetic diversity indicated that XW and DN populations were classified as the high-diversity group (with

 doi:10.20944/preprints202604.1282.v1

more than 440,000 polymorphic markers, expected heterozygosity >0.31 and polymorphism information content (PIC) ≈0.25), while the XF population was the low-diversity group (with 342,646 polymorphic markers, expected heterozygosity of 0.2608 and PIC of 0.2073). Only the minor allele frequency (MAF) of the XF population (0.2829) was slightly higher than that of the other two populations. This study systematically elucidated the characteristics of genetic differentiation and diversity differences of *M. armatus* in the Dongjiang and Ganjiang River Sources, providing a genome-level scientific basis for the conservation of genetic resources, development of molecular markers and analysis of environmental adaptive mechanisms of this species.

Keywords: *Mastacembelus armatus*; whole-genome resequencing; genetic diversity; Dongjiang river source; Ganjiang river source

1. Introduction

Mastacembelus armatus (zig-zag eel) belongs to the genus *Mastacembelus* of the family Mastacembelidae in the order Synbranchiformes. As an important freshwater economic fish in southern China, it possesses high edible and nutritional value as well as ecological value [1]. In recent years, affected by the deterioration of aquatic ecological environment, overfishing and other factors, the resources of wild *M. armatus* populations have been declining sharply. It has been listed as a provincial key protected wild animal in many provinces including Guangdong and Fujian, making the conservation and sustainable utilization of its germplasm resources a research hotspot in the fields of aquaculture and aquatic biodiversity conservation [2,3].

Focusing on the resource conservation and industrial application of *M. armatus*, scholars have carried out systematic research from multiple dimensions including cross breeding, growth regulation, nutritional feed, disease immunity, sex determination and differentiation, and population genetics. In the field of cross breeding, as a core approach for germplasm improvement, the mechanism of heterosis formation in distant hybridization has been extensively elucidated: Chen et al. [1] confirmed through the hybridization of *M. armatus* and *Mastacembelus lateralis* that the phenotypic differentiation of hybrid offspring is closely related to the differentially expressed genes in key pathways such as the insulin-like growth factor (IGF) signaling pathway and skeletal muscle development; Yang et al. [4] clarified the correlation between heterosis of hybrid offspring and the pathways of growth regulation and material metabolism by combining transcriptome and metabolome analysis; Sun et al. [5] revealed that the core causes of low hatching rate and high malformation rate in the hybrid offspring of *M. armatus* and *Sinobdella sinensis* are the disordered expression of NAD(P)H-related genes, which provides an important basis for the selection of hybrid parents. In the research on growth regulation, Han et al.[6] identified multiple growth-related SNP loci and candidate genes through genome-wide association study (GWAS), and Luo et al. [7] developed 42 microsatellite markers, providing technical support for molecular marker-assisted selection; Xue et al. [8] and Zhong et al. [9] further analyzed the regulatory functions of the growth hormone/insulin-like growth factor 1 (GH/IGF-1) axis and growth hormone-releasing hormone (GHRH). From the perspective of feed optimization, Xue et al. [10] proposed the improvement effect of hydrolysable tannin on the growth of juvenile fish, which improved the technical pathway for enhancing growth performance.

Population genetic diversity and population structure are the core foundations of species resource conservation. Existing studies have conducted preliminary explorations by means of mitochondrial genome, microsatellite markers and other tools. Wang et al. [11] revealed the population structure and historical dynamics of *M.armatus* in southern China through mitochondrial genome analysis, and confirmed the driving effect of geographical isolation on population differentiation; Tingting et al. [2] further verified the influence of river systems on the population genetic structure using microsatellite markers; Wu et al. [3] provided data support for regional conservation through the resource assessment of the *M. armatus* population in the Taojiang River of

 doi:10.20944/preprints202604.1282.v1

Jiangxi Province. In addition, in the research on reproductive biology, Alam et al. [12] and Rashid et al. [13] supplemented the basic data on growth and reproduction of different geographical populations, and Moosa et al. [14] provided a warning for the health management of wild populations through pathogen detection. Despite laying a foundation for the research on *M. armatus*, the existing achievements still have obvious limitations: first, population genetic research based on low-throughput molecular markers is difficult to fully analyze genetic variations at the genomic level; second, there is a lack of systematic genomic analysis of populations in key regions such as the Dongjiang River Source and Ganjiang River Source, which cannot support precise conservation at the regional scale; third, the molecular mechanism of environmental adaptive evolution has not been thoroughly explained, which restricts the efficient utilization of genetic resources. Based on this, this study selected three typical M. armatus populations from Xunwushui (XW) and Jiuqu River (DN) in the Dongjiang River Source, and Taojiang River (XF) in the Ganjiang River Source as research objects to carry out whole-genome resequencing analysis. By systematically analyzing the population genetic structure, differentiation characteristics and selection signals, this study clarified the differences in genetic diversity and evolutionary rules of *M. armatus* between the Dongjiang River Source and Ganjiang River Source, and further identified the selected genes and explored their environmental adaptive mechanisms. The research results not only provide genomic data support for the analysis of genetic resource diversity and the development of molecular markers for important traits of *M. armatus*, but also offer a scientific basis for the regional population conservation, germplasm resource excavation and sustainable utilization of this species.

2. Materials and Methods

2.1. DNA Extraction

Muscle samples of *M. armatus* populations were collected, with detailed sample information shown in Table 1. The muscle samples were immediately preserved in liquid nitrogen on site, and genomic DNA was extracted from the muscle tissues using the Marine Animal Tissue Genomic DNA Extraction Kit (DP324, Tiangen Biotech Co., Ltd.). After extraction, the quality of the isolated DNA was detected by 1.2% agarose gel electrophoresis, and the concentration and purity of DNA were determined using a Nanodrop D2000 ultramicro spectrophotometer. This study was approved by the Institutional Animal Care and Use Committee (IACUC) of Jiangxi Fisheries Research Institute. All experimental procedures were strictly performed in accordance with the ethical norms and the formulated rules and regulations of Jiangxi Fisheries Research Institute.

Table 1. Sample Information of the *M. armatus* populations.

Population	Latitude	Longitude	Drainage	River
Xunwu (XW)	25.0085	115.7439	Headwaters of the Dongjiang Rive	Xunshui River
Dingnan (DN)	24.9129	115.2231	Headwaters of the Dongjiang Rive	Jiuqu River
Xinfeng (XF)	25.4649	114.9815	Headwaters of the Ganjiang River	Taojiang River

2.2. Sample Sequencing and Sequence Alignment

Samples were sent to Beijing Biomarker Technologies Co., Ltd. for sequencing. Qualified DNA samples were randomly fragmented into 300-500 bp lengths using ultrasonic fragmentation. Library construction was completed through a series of steps including end repair, poly-A tailing, adapter ligation, purification, and PCR amplification. The constructed resequencing libraries were sequenced on the Illumina HiSeq 2500 platform.

Raw data underwent quality control (QC) filtering, during which adapter sequences, low-quality reads (MQ ≥ 6), and short fragments (< 500 bp) were removed to generate clean reads. The

 doi:10.20944/preprints202604.1282.v1

clean reads were aligned against the *M. armatus* genome[15] using BWA software[16]. Samtools[17] was employed to calculate the alignment rate and coverage.

2.3. Variant Detection and Structural Annotation

Variant detection for single nucleotide polymorphisms (SNPs) and insertions/deletions (INDELs) was performed using GATK[18]. The filtering criteria applied were as follows:

SNP: QD < 2.0, MQ < 40.0, FS > 60.0, MQRankSum < -12.5, ReadPosRankSum < -8.0; INDEL: QD < 2.0, FS > 200.0, ReadPosRankSum < -20.0. Structural variation detection was carried out using lumpy[19], and copy number variation (CNV) detection was implemented using CNVnator [20]. Subsequently, ANNOVAR[21] and snpEff [22] software were utilized to perform functional and positional annotation for the SNPs, INDELs, structural variations (SVs), and copy number variations (CNVs).

2.4. Population Evolutionary Analysis

Population structure and phylogenetic relationships among the populations were investigated using population clustering analysis and principal component analysis (PCA). Phylogenetic trees were constructed with IQ-TREE[23] and visualized using iTOL [24]. PCA was performed with PLINK[25], and the plots were generated using the ggplot2 package in R language. Nucleotide diversity of each population was calculated using VCFtools[26].

2.5. Gene Enrichment Analysis

Gene function annotation was performed using eggNOG-mapper[27]. The *M. armatus* OrgDB database was constructed using AnnotationForge (https://bioconductor.org/packages/release/bioc/html/AnnotationForge.html). GO and KEGG enrichment analyses were carried out using the enrichGO and enricher functions implemented in the clusterProfiler package [28].

3. Results

3.1. Whole-Genome Resequencing Data Analysis

After quality control, a total of 209.05 Gbp of clean data was generated from the resequencing data of the three populations, yielding 697,876,043 clean reads. The Q30 value reached 94.42%, indicating a high sequencing quality. The average mapping rate against the reference genome was 97.85%, with an average coverage depth of 10×. The genome coverage was 91.40% (covered by at least one base) (Table 2). These results demonstrate that the sequencing data is of reliable quality and suitable for subsequent analyses.

Table 2. Quality and Alignment Summary of Whole Genome Resequencing Data.

Population	Bases/bp	GC Content /%	Q20/%	Q30/%	Reads	Alignment Rate/%	Sequencing Depth
Xunwu (XW)	70075162040	40.23	97.90	94.36	233938669	97.85	10.20
Dingnan (DN)	72130154974	40.23	97.94	94.42	240800045	97.86	10.30
Xinfeng (XF)	66844317454	40.21	97.95	94.49	223137329	97.84	9.70

3.2. SNP Detection and Statistical Annotation of Results

A total of 7,459,686 SNPs were detected on average across the three populations, including 4,510,018 transition SNPs and 2,949,668 transversion SNPs, with a transition/transversion ratio of

 doi:10.20944/preprints202604.1282.v1

1.52. The number of heterozygous SNPs was 165,357 and homozygous SNPs was 7,294,329, accounting for a heterozygous SNP ratio of 2.22% (Table 3).

Table 3. Statistics of Single Nucleotide Polymorphisms (SNPs) obtained from sequence detection.

Population	SNP number	SNP number (transition)	SNP number (transversion)	Transition/transversion SNP ratio	SNP number (heterozygous)	SNP number (homozygous)	Heterozygous SNP percentage (%)
Xunwu (XW)	7467441±61589^{ab}	4514787±37776^{ab}	2952654±23815^{ab}	1.52±0^{a}	173325±10282^{a}	7294117±53593^{a}	2.32±0.1476^{a}
Dingnan (DN)	7483236±64472^{a}	4524336±39454^{a}	2958900±25021^{a}	1.52±0^{a}	180267±10456^{a}	7302968±55468^{a}	2.43±0.1252^{a}
Xinfeng (XF)	7428382±45245^{b}	4490931±27732^{b}	2937451±17517^{b}	1.52±0^{a}	142478±7749^{b}	7285903±43635^{a}	1.91±0.0994^{b}
Total (ZT)	7459686±60435	4510018±36997	2949668±23441	1.52±0	165357±19090	7294329±49873	2.22±0.2578

Note: Data are presented as mean ± standard deviation. Different superscript letters in the same column indicate significant differences at the level of $p < 0.05$.

Functional annotation of the identified SNPs across the three populations revealed their distribution in different genomic regions as follows: 1,592,482 intergenic region SNPs, 3,979,280 intronic SNPs, 576,424 SNPs in the upstream gene region (within 5K), 511,922 SNPs in the downstream gene region (within 5K), 599,355 5′ untranslated region (5′ UTR) SNPs, 260,461 3′ untranslated region (3′ UTR) SNPs, 393 splice acceptor site SNPs, 386 splice donor site SNPs, 41,042 splice site region SNPs, 3,320 SNPs with start codon gain, 230 SNPs with start codon loss, 265,627 synonymous coding SNPs, 157,746 non-synonymous coding SNPs, 200 synonymous stop codon SNPs, 270 SNPs with stop codon gain, and 127 SNPs with stop codon loss (Table 4).

Table 4. Functional Annotation Statistics of Single Nucleotide Polymorphisms (SNPs) Identified by Sequence Detection.

Population	Intergenic region	Intron	Upstream gene region (within 5K)	Downstream gene region (within 5K)	5′ untranslated region (5′ UTR)	3′ untranslated region (3′ UTR)	Splice acceptor site	Splice donor site
Xunwu (XW)	1594206±13286^{ab}	3983366±32141^{ab}	577187±4910^{ab}	512281±4234^{ab}	60066±555^{ab}	260706±1848^{ab}	393±6^{a}	384±8^{b}
Dingnan (DN)	1597961±13795^{a}	3991511±33812^{a}	578180±5018^{a}	513580±4421^{a}	60145±641^{a}	261191±1863^{a}	396±8^{a}	390±5^{a}
Xinfeng (XF)	1585280±9677^{b}	3962964±23565^{b}	573904±3636^{b}	509905±3075^{b}	59595±482^{b}	259487±1304^{b}	390±5^{a}	383±5^{b}
Total (ZT)	1592482±13121	3979280±31573	576424±4780	511922±4118	599355±597	260461±1788	393±7	386±7
Population	Splice site region	Start codon gain	Start codon loss	Synonymous coding	Non-synonymous coding	Synonymous stop codon	Stop codon gain	Stop codon loss
Xunwu (XW)	41096±408^{a}	3330±28^{a}	229±2^{a}	265892±2762^{a}	158011±1569^{ab}	199±2^{ab}	271±5^{a}	126±2^{a}
Dingnan (DN)	41185±425^{a}	3328±35^{a}	231±3^{a}	266604±2883^{a}	158358±1707^{a}	202±4^{a}	272±6^{a}	127±2^{a}
Xinfeng (XF)	40846±293^{a}	3302±27^{a}	230±4^{a}	264386±2025^{a}	156868±1200^{b}	198±3^{b}	267±7^{a}	127±2^{a}

 doi:10.20944/preprints202604.1282.v1

Total (ZT)	41042±395	3320±32	230±3	265627±2665a	157746±1592	200±3	270±6	127±2

Note: Data are presented as mean ± standard deviation. Different superscript letters in the same column indicate significant differences at the level of $p < 0.05$.

3.3. Detection and Result Annotation of Small Fragment Insertions and Deletions

The overall numbers of coding region insertions, coding region deletions, homozygous coding region indels, heterozygous coding region indels, total coding region indels, genome-wide insertions, genome-wide deletions, homozygous genome-wide indels, heterozygous genome-wide indels and total genome-wide indels detected in the three populations were 4402±73, 5298±88, 8788±107, 913±69, 9700±159, 917918±11690, 984804±11582, 1827240±17482, 75482±9024 and 1902722±23247, respectively (Table 5).

Table 5. Statistics of insertions and deletions in the whole genome and coding regions.

Population	Coding region insertions	Coding region deletions	Homozygous indels in coding regions	Heterozygous indels in coding regions	Total indels in coding regions	Whole-genome insertions	Whole-genome deletions	Homozygous indels in whole genome	Heterozygous indels in whole genome	Total indels in whole genome
Xunwu (XW)	4400±72[a]	5295±84[a]	8776±106[a]	919±53[a]	9695±154[a]	918905±12529[a]	986020±11990[a]	1826216±19229[a]	78709±5940[a]	1904925±24504[a]
Dingnan (DN)	4426±85[a]	5329±101[a]	8793±125[a]	962 ±65[a]	9755±184[a]	921581±12935[a]	989080±12719[a]	1828376±19543[a]	82285±6392[a]	1910661±25638[a]
Xinfeng (XF)	4380±60[a]	5270±75[a]	8794±98[a]	856±44[b]	9650±132[a]	913268±8720[a]	979312±8438[a]	1827127±15182[a]	65453±3387[b]	1892580±17142[a]
Total (ZT)	4402±73	5298±88	8788±107	913±69	9700±159	917918±11690	984804±11582	1827240±17482	75482±9024	1902722±23247

Note: Data are presented as mean ± standard deviation. Different superscript letters in the same column indicate significant differences at the level of $p < 0.05$.

3.4. Functional Annotation of Mutant Genes at the DNA Level

The GO classification systems of the samples from the three water areas were highly consistent. For cellular components (CCs), 17 categories were identified, ranging from the highly abundant cell/cell part (7796/7238) to the low-abundant nucleoid (1/1), with the categories and their corresponding gene numbers being completely identical across all samples. For molecular functions (MFs), 14 categories were found, dominated by binding (10172/9552) and catalytic activity (5525/5251), and the gene numbers of specific functions such as toxin activity (1/1) were fixed. For biological processes (BPs), 23 categories were detected, with the core concentrated in basic processes including cellular process (8883/8409) and biological regulation (6265/5927), and consistent characteristics observed in low-abundant processes like cell aggregation (8/7).

Across all water area samples, the proportion of mutant genes in each GO category relative to the total genes was maintained at 85%–95%, with no abnormally high or low variation detected in any functional gene set. For example, the variation proportion of membrane-related genes in cellular components was approximately 93% (membrane: 7374/6868, membrane part: 6605/6136); that of signal transducer activity genes in molecular functions was about 90% (1502/1358); and that of response to stimulus genes in biological processes was roughly 95% (3218/3051), which reflected a consistent variation pattern of core genes among the three populations.

All samples from the three water areas exhibited the characteristic of "active basic functions and conserved specific functions". Specifically, active functions (e.g., binding, catalytic activity, and cellular processes) were supported by a large number of genes with stable variation, serving as the core underpinning for the life activities of the populations. In contrast, conserved functions (e.g., toxin

 doi:10.20944/preprints202604.1282.v1

activity, translational regulator activity, and nucleoid) showed identical numbers of total genes and mutant genes (e.g., 1/1 for toxin activity). These genes are speculated to be species-essential conserved genes that are not affected by environmental differences among the water areas.

3.5. Genetic and Phylogenetic Analysis

The phylogenetic tree revealed that all *M. armatus* samples from Xunwu Water (XW), Jiuqu River (DN) and Taojiang River (XF) clustered within a single major phylogenetic clade, indicating a high level of species-level consistency among the *M. armatus* populations from the three water areas. These populations share a common ancestor, with no obvious species differentiation observed and close genetic relationships among them.

Some *M. armatus* samples from different water areas showed interleaved clustering in the phylogenetic tree and did not form independent clades strictly according to their water area origins. For instance, several clades containing samples from Xunwu Water and Taojiang River were closely adjacent or even interspersed, suggesting the possibility of genetic exchange between *M. armatus* populations across different water areas.

However, *M. armatus* samples from Jiuqu River exhibited a relatively concentrated clustering pattern and formed several relatively independent small clades. In contrast, samples from Xunwu Water and Taojiang River were distributed more dispersedly; although local relatively aggregated small clades were present in each population, the interspersion between them was more pronounced (see Figure 1 for details).

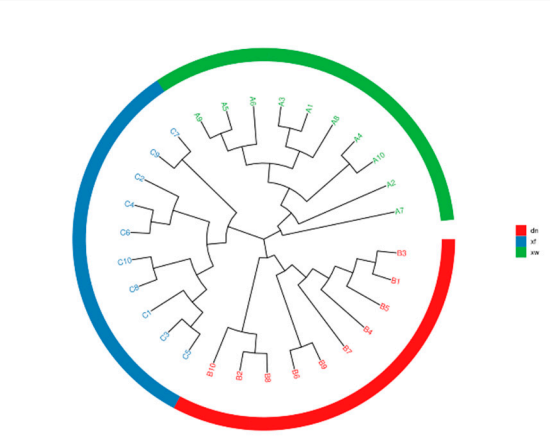

Figure 1. Phylogenetic tree.

3.6. Population Genetic Structure and Diversity Analysis

Based on the cross-validation (CV) error curve from Admixture analysis, the CV error reached a local minimum at K = 2 (marked by the red dot), indicating that the model provided the most reliable inference of the population genetic structure of *M. armatus* at this value, which is thus the core K for resolving population structure (see Figure 2 for details). The genetic composition of M. armatus individuals from Xunwu Water (XW) and Jiuqu River (DN) was dominated by the blue component with an extremely low proportion of the red component, suggesting that these two populations had a genetic background highly biased toward one ancestral population, with strong intraspecific genetic consistency and significant genetic differentiation from the population of the other water area. In contrast, the genetic composition of Taojiang River (XF) individuals was dominated by the red component with a minimal blue component. Similar to XW and DN, the XF population was also highly biased toward a single ancestral population, but the ancestral population

 doi:10.20944/preprints202604.1282.v1

it leaned toward (red) differed from that of XW and DN (blue), indicating obvious genetic differentiation between XF and the other two populations. Notably, DN individuals exhibited a mixed pattern of blue and red genetic components, suggesting that the DN population harbored genetic contributions from both ancestral populations and represented the population with the highest level of genetic admixture (see Figure 3 for details).

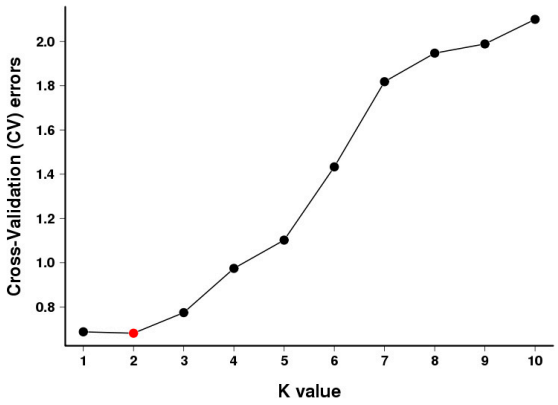

Figure 2. Cross-validation error curve.

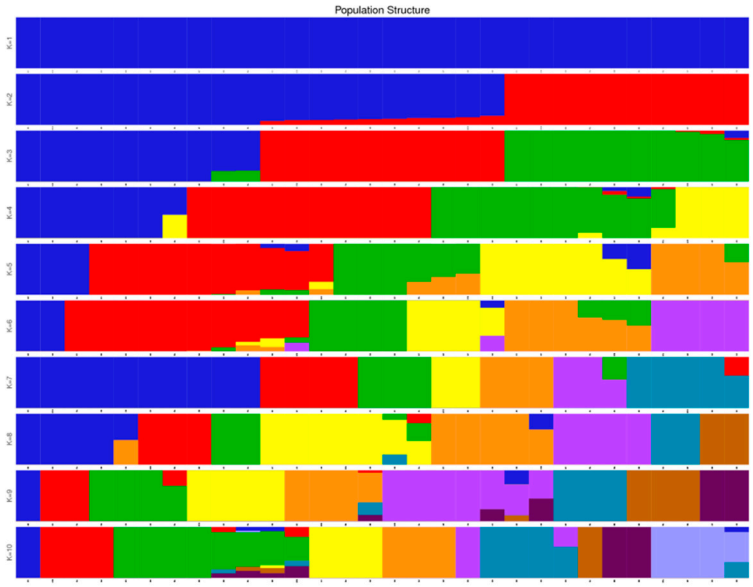

Figure 3. Sample clustering results corresponding to different K values.

From the clustering results of Principal Component Analysis (PCA), PC1 (≈11.94%), PC2 (≈6.27%) and PC3 (≈4.16%) collectively explained approximately 22.37% of the total genetic variation. Although not covering all variation, these principal components reflected the major patterns of genetic differentiation among populations. The XW, DN and XF populations all showed strong intraspecific genetic consistency but significant interspecific genetic differentiation, each forming an independent genetic cluster. Genetic differentiation between DN and XF was bidirectional and complete; XW had a relatively close genetic distance to DN but significant differentiation from XF (see Figure 4 for details).

 doi:10.20944/preprints202604.1282.v1

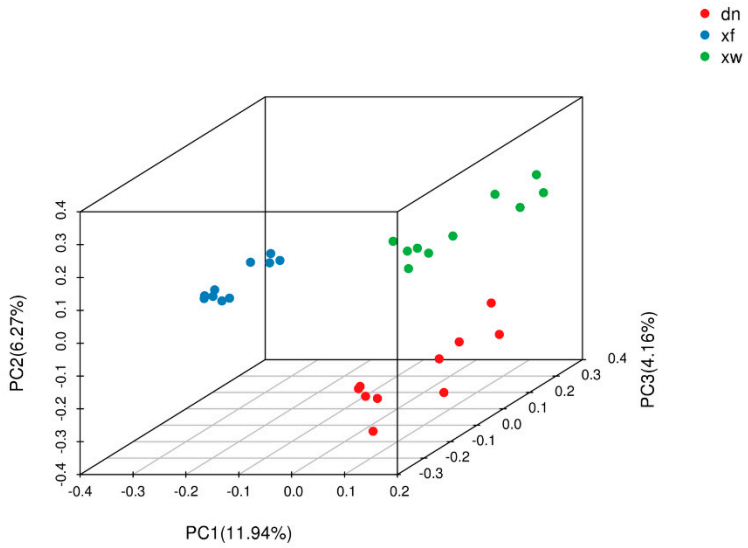

Figure 4. Clustering results of Principal Component Analysis (PCA).

A heatmap based on kinship values was used to compare the genetic relatedness of M. armatus populations from the three water areas (XW: Series A, DN: Series B, XF: Series C). Kinship values reflect the degree of genetic relatedness between individuals, with values closer to 1 indicating closer relatedness and values closer to 0 indicating greater genetic distance. In the heatmap, kinship values among XW individuals showed a cluster of medium-to-high values (orange-red in color); for example, the kinship values between A1 and A2, A3 and other XW individuals were significantly higher than those between XW individuals and individuals from other water areas. This indicated close genetic relatedness within the XW population, strong intraspecific genetic consistency, and potential inbreeding or sufficient genetic exchange within the small population. Kinship values among DN individuals also showed a concentration of medium-to-high values (orange-red in color), with high values between B1 and B2, B3 and other DN individuals, suggesting close relatedness within the DN population and a relatively stable and tightly linked internal genetic structure. Kinship values among XF individuals were dominated by medium-to-low values (blue in color); the kinship values between C1 and C2, C3 and other XF individuals were significantly lower than those within the XW and DN populations, reflecting relatively distant genetic relatedness within the XF population and richer genetic diversity at the individual relatedness level. Kinship values between XW (Series A) and DN (Series B) individuals were generally in the medium-to-low range (blue in color), indicating distant genetic relatedness and significant genetic differentiation between the two populations, which may result from restricted genetic exchange due to geographic isolation or ecological environmental differences. Kinship values between XW (Series A) and XF (Series C) individuals were also dominated by medium-to-low values (blue in color), suggesting distant relatedness and obvious differences in genetic background between these two populations. Similarly, kinship values between DN (Series B) and XF (Series C) individuals fell in the medium-to-low range (blue in color), reflecting loose genetic relatedness and prominent genetic differentiation between the DN and XF populations (see Figure 5 for details).

 doi:10.20944/preprints202604.1282.v1

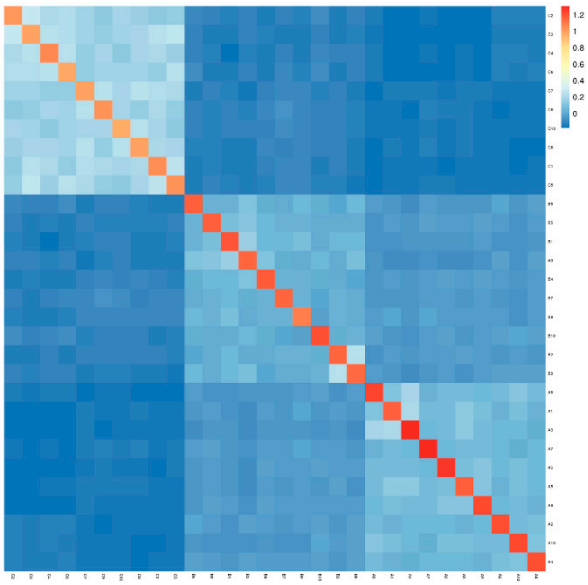

Figure 5. Heatmap of genetic relationship distribution.

Based on the linkage disequilibrium (LD) decay curves (mean R^2 vs. physical distance) of M. armatus from the three water areas (XW, DN, XF), all three curves showed a trend of a rapid decrease in mean R^2 with increasing physical distance followed by a gradual plateau, consistent with the universal law of genomic LD decay: the shorter the physical distance between loci, the higher the level of LD; the longer the distance, the more easily recombination events break the linkage, leading to lower LD. The XF population had the highest initial mean R^2 value and the slowest decay rate within the short distance range (0–100 kb), with its curve consistently at the top. The DN population had a lower initial mean R^2 than XF, with a decay rate faster than XF but slower than XW within 0–100 kb, and its curve lay between those of XF and XW. The XW population had the lowest initial mean R^2 and the fastest decay rate within 0–100 kb, with its curve consistently at the bottom (see Figure 6 for details).

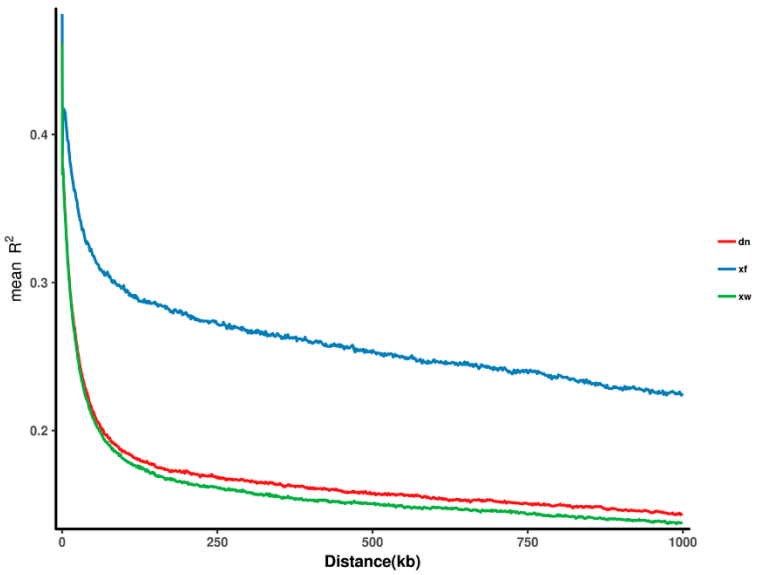

Figure 6. Decay distance of linkage disequilibrium.

All populations had an Average minor allele frequency (Average_MAF) in the range of 0.26–0.28 without extreme high or low values, indicating a relatively balanced distribution of alleles within

 doi:10.20944/preprints202604.1282.v1

each population and no over-dominance of a single allele. The values of expected heterozygosity and observed heterozygosity were close in all populations (e.g., DN: 0.3211 vs. 0.3213; XW: 0.3185 vs. 0.3088), with no significant deviation. A large number of polymorphic markers were detected in all three populations (Number_of_poly_marker > 340,000 for all), and the polymorphism information content (PIC) ranged from 0.20 to 0.26, representing a moderate level of polymorphism.The three populations could be clearly divided into a high-diversity group (DN, XW) and a low-diversity group (XF). The DN and XW populations showed highly similar values in all 8 core diversity indicators, all of which were significantly higher than those of the XF population. Both populations had over 440,000 polymorphic markers (DN: 448,746; XW: 446,349), an observed number of alleles close to 1.92 (DN: 1.922253; XW: 1.917326), and an expected number of alleles of approximately 1.54 (DN: 1.5482; XW: 1.5433). Both populations had an expected heterozygosity > 0.31 (DN: 0.3211; XW: 0.3185), a Nei's diversity index > 0.33 (DN: 0.3387; XW: 0.336), a Shannon-Wiener index > 0.47 (DN: 0.4815; XW: 0.4778), and a PIC of approximately 0.25 (DN: 0.2572; XW: 0.2552), with all indices at a relatively high level. Low-diversity group (XF),All core diversity indicators of the XF population were significantly lower than those of DN and XW. XF had only 342,646 polymorphic markers (≈106,000 fewer than DN), an observed number of alleles of 1.704198 (0.218 lower than DN), and an expected number of alleles of 1.4529 (0.095 lower than DN). XF had an expected heterozygosity of 0.2608 (0.0603 lower than DN), a Nei's diversity index of 0.2752 (0.0635 lower than DN), a Shannon-Wiener index of 0.3864 (0.0951 lower than DN), and a PIC of 0.2073 (0.0499 lower than DN), with all indices at the lowest level among the three populations. Notably, the Average_MAF of XF (0.2829) was slightly higher than that of DN (0.2616) and XW (0.2606), which was the only indicator that did not follow the low-diversity pattern (see Table 6 for details).

Table 6. Population Genetic Diversity.

Population	Average_MAF	Expected_allele_number	Expected_heterozygous_number	Nei_diversity_index	Number_of_poly_marker	Observed_allele_number	Observed_heterozygous_number	Polymorphysm_information_content	Shnnon_Wiener_index
Xunwu (XW)	0.2606	1.5433	0.3185	0.336	446349	1.917326379	0.3088	0.2552	0.4778
Dingnan (DN)	0.2616	1.5482	0.3211	0.3387	448746	1.922252639	0.3213	0.2572	0.4815
Xinfeng (XF)	0.2829	1.4529	0.2608	0.2752	342646	1.704198316	0.2581	0.2073	0.3864

4. Discussion

Studies on Megalobrama have confirmed that the blockage of gene flow caused by geographical isolation is the core driving factor for its population differentiation: the northern Megalobrama populations exhibit low genetic diversity, intensified inter-population differentiation, high levels of linkage disequilibrium, and a tendency toward a homogeneous genetic structure. In contrast, the Poyang Lake Megalobrama population has higher genetic diversity and Tajima's D values, lower linkage disequilibrium, and genetic characteristics closer to the ancestral state, making it a key reserve group for the primitive genetic traits of Megalobrama[29]. This study also found that key genes related to reproduction, somatic development, and muscle metabolism in northern Megalobrama populations are under significant positive selection, which serves as the molecular basis for their adaptation to geographically isolated habitats and the genetic root of phenotypic and physiological differentiation from sympatric populations[29]. Based on GO enrichment statistics from whole-genome resequencing, this study revealed no significant differences in gene function classification, the number of functional genes, or the proportion of gene variations of *M. armatus* among the three water systems of Xunwu River, Jiuqu River, and Tao River in Ganzhou. This phenomenon is speculated to stem from two factors: first, Xunwu River (XW) and Jiuqu River (DN), both tributaries

 doi:10.20944/preprints202604.1282.v1

of the headwater system of the Dongjiang River, have relatively good aquatic connectivity and few physical river barriers, serving as the main channel for gene flow of *M. armatus*. Second, Tao River (XF), a tributary of the upper Ganjiang River, is geographically isolated from Xunwu River and Jiuqu River by a watershed with no direct hydrological connectivity, forming a relatively independent aquatic unit. However, all three water systems are located in Ganzhou and characterized by freshwater river habitats with no obvious differences in core ecological factors such as water temperature, water quality, and food composition, which fail to generate differential natural selection pressure and thus cannot drive significant functional gene differentiation of *M. armatus*. To further clarify population genetic differences, subsequent studies should focus on the specific variant loci of environment-associated genes such as stimulus response and metabolic processes to verify aquatic adaptive variations; meanwhile, detect the expression levels of key functional genes related to immunity and stress resistance to analyze whether aquatic environmental heterogeneity induces differentiation in gene expression regulation.

Relevant studies on the population genetic differentiation of aquatic animals have fully confirmed the close correlation between species genetic structure and geographical distribution, breeding models, and habitat characteristics. Studies on Scylla paramamosain have shown that it exhibits the characteristics of low polymorphism, weak population differentiation, and low inbreeding levels; the results of PCA, Admixture grouping (K=1-5), phylogenetic tree, and LD decay analysis are highly consistent, all indicating no significant genetic differentiation among relevant geographical populations[30]. Studies on the genetic pattern of Channa argus in Shandong have confirmed that its genetic clustering strictly matches geographical distribution, with geographical isolation being the core natural factor for differentiation; wild populations in independent water areas such as the Yellow River Estuary and Dongping Lake have formed distinct genetic clades, while populations in Nansi Lake and Yihe River have closer genetic relationships due to high aquatic connectivity[31]. Relevant studies on Megalobrama amblycephala have identified the genetic structure and selection signals of different geographical populations, confirming their independent ancestral components and limited gene flow among different geographical populations, with geographical isolation as the dominant factor for differentiation, which has laid a genomic foundation for the classification and protection of its germplasm resources[32]. Studies have found that the genetic diversity of farmed populations of Patinopecten yessoensis in China is significantly lower than that of wild populations, with a tendency toward homogeneous genetic structure and significant risks of germplasm degradation[33].

In research on stress-resistant breeding of aquatic animals, candidate genes and QTL loci related to heat tolerance and low-temperature tolerance have been identified in Lateolabrax maculatus and Chlamys farreri, respectively, providing a genomic basis for temperature stress-resistant breeding of aquatic organisms; meanwhile, some QTLs in hybrid scallops have been found to exhibit overdominance effects, which offers a genetic foundation for cultivating new cold-tolerant and fast-growing strains by utilizing interspecific heterosis[34,35]. In the field of genomic selection breeding, machine learning-based whole-genome feature selection strategies have shown excellent application value: they can accurately screen 3-6% of core loci from 8.22 million SNPs in sturgeons, improving the genomic prediction accuracy of core economic traits such as caviar yield, color, and body weight by 3.4-4.6%[36].

Classical studies employing phylogenetic trees, PCA and Admixture analyses have confirmed that geographical isolation is the core driving factor for population differentiation in freshwater fish. Wild populations of Percocypris pingi have formed multiple genetically differentiated units according to their geographical distribution, which is consistent with the findings of studies on Megalobrama and Channa argus[37]. Meanwhile, whole-genome resequencing technology has been applied to research on the dispersal of downstream Coilia brachygnathus into large reservoirs. The dispersed populations exhibit moderate levels of nucleotide diversity (π) and heterozygosity (He), a short LD decay distance, and a recent rapid increase in effective population size (Ne), which conform

 doi:10.20944/preprints202604.1282.v1

to the typical genetic characteristics of dispersed species—rapid proliferation after a bottleneck effect[38].

Multi-dimensional analyses of *M. armatus* from the three major water systems in Ganzhou in this study revealed a unique genetic pattern. The phylogenetic tree showed that all samples from the three locations belong to the same evolutionary clade, with a high degree of consistency at the species level, a common ancestor, no obvious speciation, and overall close genetic relationships. Among them, samples from the Jiuqu River were clustered closely and formed an independent small clade with significant genetic uniqueness, which may be attributed to the stable ecological environment and certain geographical isolation of this water system, leading to the accumulation of specific genetic variations during evolution; samples from the Xunwu River and Tao River were scattered in distribution, with obvious interspersion despite local clustering, reflecting more frequent gene flow between their populations, possibly due to weak selective pressure from the ecological environment resulting in insignificant genetic differences among populations.

Admixture population structure analysis showed that the model had the highest goodness of fit at K=2, confirming that *M. armatus* from the three water systems share a genetic origin based on dual ancestral populations. The Tao River population was significantly biased towards the red ancestral population, while the Jiuqu River and Xunwu River populations were biased towards the blue ancestral population, with significant genetic differentiation between the two major ancestral components. During evolution, the Xunwu River and Tao River populations have specialized towards different ancestral populations respectively, and the Jiuqu River population has retained more primitive mixed genetic characteristics—possibly acting as a genetic bridge for gene flow between the two, or being in a transitional stage of genetic differentiation.

PCA analysis further revealed that the Jiuqu River formed an independent cluster in the core genetic dimension, with a high degree of differentiation from the Tao River and Xunwu River; the Tao River exhibited a unique genetic variation pattern, with bidirectional genetic differentiation from the other two water systems; the Xunwu River and Jiuqu River had partially adjacent distribution areas, with a short genetic distance and closer connections, but significant differentiation from the Tao River. Kinship heatmap analysis indicated that within populations, individuals from the Xunwu River and Jiuqu River showed aggregation of medium to high Kinship values, with close genetic relationships and strong genetic consistency, which may be related to inbreeding within populations and sufficient gene flow in small populations; individuals from the Tao River had predominantly medium to low Kinship values, with distant genetic relationships and richer genetic diversity at the individual kinship level. Among populations, Kinship values among the three locations were all medium to low, with no close genetic relationships and prominent differentiation characteristics, which may be caused by restricted gene flow due to geographical isolation and differences in ecological environments. LD decay distance analysis showed that the Tao River had the longest LD decay distance, suggesting a small population size, strong genetic drift and low recombination rate; the Jiuqu River had a moderate LD decay distance, with intermediate levels of effective population size and recombination rate; the Xunwu River had the shortest LD decay distance, indicating a large population size, weak genetic drift and frequent recombination, which can quickly break linked allelic combinations. Population genetic diversity analysis showed that all three locations conform to the regular patterns of normal freshwater fish, without being affected by severe inbreeding drift, with stable genetic structures. More than 340,000 polymorphic markers were detected in all populations, providing a solid foundation for subsequent research. Among populations, they can be divided into a high-diversity group (Jiuqu River, Xunwu River) and a low-diversity group (Tao River). The former had significantly higher core indicators and richer genetic variations; the latter had more homogeneous genomic variant loci and allelic types, with significantly lower diversity. Only the Tao River had a slightly higher Average_MAF, due to the high minor allele frequency at a small number of variant loci, presenting a special characteristic of locally high-frequency variation but overall low diversity.

 doi:10.20944/preprints202604.1282.v1

In summary, the *M. armatus* populations from the three major water systems of Xunwu River, Jiuqu River and Tao River in the Ganzhou water system exhibit strong genetic consistency within each population, each forming an independent genetic cluster, with significant genetic differentiation among populations. Among them, the Jiuqu River and Tao River populations show bidirectional and complete genetic differentiation; the Xunwu River and Jiuqu River populations have a short genetic distance, while the Xunwu River and Tao River populations exhibit significant genetic differentiation. Differences in LD decay distances reflect the divergent characteristics of the three populations in terms of effective population size and recombination rate: the Tao River population is characterized by a small population size and low recombination rate, the Xunwu River population by a large population size and high recombination rate, and the Jiuqu River population by intermediate characteristics between the two.

Based on the results of this study, germplasm resource conservation and research on *M. armatus* can be carried out from four aspects in the follow-up. First, aiming at the genetic correlation between the Xunwu River and Jiuqu River populations, investigate the ecological connectivity (river course connectivity, migratory behavior, etc.) of the two water systems, explore the ecological driving factors of their genetic correlation, and verify the impact of effective population size on LD decay distance combined with population ecological surveys. Second, systematically investigate the ecological environmental factors of the three water systems, analyze the shaping effect of ecological environmental heterogeneity on population genetic structure, and provide a scientific basis for habitat conservation. Third, establish a long-term population dynamic monitoring mechanism to track changes in genetic structure and gene flow, adjust conservation and management strategies in a timely manner, and maintain genetic diversity and ecological balance. Fourth, formulate differentiated conservation strategies, focus on protecting the genetic diversity of the Tao River population to avoid a decrease in its diversity caused by human interference, and continuously monitor the risk of inbreeding in the Xunwu River and Jiuqu River populations, and maintain the genetic health of the populations through reasonable gene flow strategies when necessary.

5. Conclusions

This study performed whole-genome resequencing of three Mastacembelus armatus populations from the Dongjiang River source (Xunwushui, XW; Jiuqu River, DN) and Ganjiang River source (Taojiang River, XF) to reveal their genetic diversity, population structure and adaptive characteristics. All three populations belong to the same evolutionary clade and originate from two ancestral components. The XW and DN populations show high genetic diversity, close genetic relationship and short linkage disequilibrium decay distance, while the XF population exhibits low genetic diversity, distant individual relatedness and long linkage disequilibrium decay distance. Significant genetic differentiation exists among populations, which is mainly driven by geographical isolation and limited gene flow between river systems.

These findings provide genome-level insights into the genetic background and differentiation mechanism of M. armatus in southern river systems. The results support the development of molecular markers, the analysis of environmental adaptation mechanisms, and the formulation of targeted conservation strategies. Priority should be given to protecting the genetic diversity of the XF population and monitoring the inbreeding risk of XW and DN populations to maintain the long-term genetic health and sustainable utilization of this important freshwater fish resource.

Author Contributions: Conceptualization, Bin Wu; methodology, Bin Wu and Yuan Fang; software, Bin Wu and Yuan Fang; validation, Bin Wu and Qingxiang Zeng; formal analysis, Bin Wu and Han Li; investigation, Haihua Wang and Bin Wu; resources, Han Li and Bin Wu; data curation, Bin Wu and Haihua Wang; writing-original draft preparation, Bin Wu; writing-review and editing, Bin Wu and Haihua Wang; visualization, Haihua Wang and Bin Wu; supervision, Bin Wu and Haihua Wang; project administration, Haihua Wang and Bin Wu; funding acquisition, Haihua Wang. All authors have read and agreed to the published version of the manuscript.

 doi:10.20944/preprints202604.1282.v1

Funding: This research was funded by 2025 Jiangxi Provincial Modern Seed Industry Revitalization Project (Aquatic Seed Industry), grant number sczy202511.

Institutional Review Board Statement: The animal study protocol was approved by the Ethics Committee of Jiangxi Provincial Institute of Fisheries Science (JXFRI-EAE-20211008-01).

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this study are available from the corresponding author upon reasonable request for academic research purposes.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen Y ,Zhang L ,Lin X , et al.Growth Performance, Digestive Capacity, and Transcriptomic Analysis of the Hybrid Offspring of *Mastacembelus armatus* ×*Mastacembelus favus*[J].Animals,2025,16(1):11-11.
2. Lin T, Yang J, Yu Z, et al. Genetic diversity and population structure of *Mastacembelus armatus* in the river systems of southern China revealed by microsatellites[J]. SN Applied Sciences, 2023, 5(12): 307.
3. Bin W, Haihua W, Benhe M, et al. Estimation of population parameters and stock assessment of *Mastacembelus armatus* in the Taojiang River, Xinfeng County, Jiangxi Province, China[J]. Indian Journal of Animal Research, 2022, 56(8): 1047-1051.
4. Yang J ,Yu Z ,Feng Y , et al.Comparative transcriptomes and metabolome reveal heterosis in zig zag eel (*Mastacembelus armatus*)[J].Aquaculture Reports,2024,39102529-102529.
5. Sun D, Wang S, Wang C, et al. Transcriptomics reveals that NAD (P) H affects the development of the Zig-zag eel (*Mastacembelus armatus*♀)× Spiny eel (*Sinobdella sinensis*♂) hybrid offspring leading to low hatching rates[J]. Animal Reproduction Science, 2024, 268: 107561.
6. Han C ,Chen K ,Cui J , et al.Identification of Growth-Related SNPs and Candidate Genes in the Genome of Zig-Zag Eel (*Mastacembelus armatus*) by GWAS[J].Marine Biotechnology, 2025,27(5):129-129.
7. Luo L ,Ji L ,Wu C , et al.Isolation and characterization of forty-two polymorphic simple sequence repeat markers from Mastacembelus armatus (Symbranchiformes: Mastacembelidae)[J].Conservation Genetics Resources,2024,16(4):1-2.
8. Xue L ,Gao Y ,Zhang S , et al.Expression pattern of the fused in sarcoma gene and its contextual influence on the density-specific response of the growth hormone/insulin-like growth factor 1 axis in zig-zag eels (Mastacembelus armatus)[J].Frontiers in Marine Science,2024,111461451-1461451.
9. Zhong D, Zhang M, Lan X, et al. Molecular cloning and functional characterization of growth hormone–releasing hormone in Mastacembelus armatus[J]. Fish Physiology and Biochemistry, 2021, 47(1): 69-78.
10. Xue X, Chen Y, Yu Z, et al. Effects of diet supplemented with hydrolyzable tannin on the growth performance, antioxidant capacity, and muscle nutritional quality of juvenile Mastacembelus armatus[J]. Aquaculture Nutrition, 2024, 2024(1): 8266189.
11. Wang Y, Yang Y, Liu Y, et al. Mitogenomic insights into the population structure and demographic history of zig-zag eels (*Mastacembelus armatus*) in southern China[J]. Aquaculture Reports, 2025, 42: 102748.
12. Alam A, Gopinath G, Jha D N, et al. Maturity, breeding cycle, and fecundity of Mastacembelus armatus (Actinopterygii: Synbranchiformes: Mastacembelidae) in the sub-tropical waters of the river Ganga[J]. Acta Ichthyologica et Piscatoria, 2020, 50: 191-199.
13. Rashid M ,Sagir M ,Dobriyal K A .Age and growth analysis of the fish Mastacembelus armatus (Lacepede) from River Nayar, Garhwal Himalaya, Uttarakhand[J].Journal of Applied and Natural Science,2021,13(1):137-144.
14. Abro M M, Birmani N A, Brohi G H, et al. 11. Detection of a pathogenic nematode from tire track eel, Mastacembelus armatus captured in the Indus River, Sindh, Pakistan[J]. Pure and Applied Biology (PAB), 2020, 9(1): 91-95.
15. Xue L, Gao Y, Wu M, Tian T, Fan H, Huang Y, Huang Z, Li D, Xu L. Telomere-to-telomere assembly of a fish Y chromosome reveals the origin of a young sex chromosome pair [J]. Genome biology. 2021 Dec;22:1-20.

 doi:10.20944/preprints202604.1282.v1

16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. Bioinformatics, 2009, 25(14): 1754-1760.
17. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools [J]. Bioinformatics, 2009, 25(16): 2078-2079.
18. McKenna A, Hanna M, Banks E, et al, The genome analysis toolkit: A MapReduce ftamework for analyzing next-generation DNAsequencing data [J].Genome Research,2010,20(9): 1297-1303.
19. Layer R M, Chiang C, Quinlan A R, et a. LUMPY: A probabilistic framework for structural variant discovery [J]. Genome Biology, 2014,15(6):R84.
20. Abyzov A, Urban A E, Snyder M, et al. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs fromfamily and population genome sequencing [J]. Genome Research, 2011,21(6): 974-984.
21. Wang K, Li M Y, Hakonarson 1. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data [J]. NucleiAcids Research,2010,38(16):e164.
22. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3 [J]. fly. 2012, 6(2):80-92.
23. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies[J]. Molecular biology and evolution. 2015, 32(1):268-74.
24. Letunic l, Bork P. Interactive tree of life (iTOL) vs: An online tool for phylogenetic tree display and annotation [J]. Nucleic Acids Research, 2021,49(W1):W293-W296.
25. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses [J]. The American journal of human genetics,2007,81(3):559-575.
26. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools [J]. Bioinformatics, 2011, 27(15): 2156-2158.
27. Cantalapiedra C P, Hernandez-Plaza A, Letunic l, et al. eggNOG-mapper v2: Functional annotation, orthology assignments, and domainprediction at the metagenomic seale [J]. Molecular Biology and Evolution, 2021,38(12): 5825-5829.
28. Wu T Z, Hu E 0, Xu S B, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data [J]. The lnnovation, 2021, 2(3):100141.
29. Ding R ,Yu D ,Yang K , et al.Chromosome-Level Genome Assembly and Whole-Genome Resequencing Revealed Contrasting Population Genetic Differentiation of Black Bream (*Megalobrama skolkovii*) (Teleostei: Cyprinidae) Allopatric and Sympatric to Its Kin Species[J].Ecology and evolution,2025,15(1):e70874.
30. Zhou X ,Ouyang M ,Zhang Y , et al.Whole-Genome Resequencing Provides Novel Insights Into the Genetic Diversity, Population Structure, and Patterns of Runs of Homozygosity in Mud Crab (*Scylla paramamosain*)[J].Evolutionary applications,2025,18(9):e70153.
31. An L ,Meng Q ,Dong X , et al.Unravelling the genetic landscape of Channa argus in Shandong Province: Insights into population differentiation, adaptation and conservation from whole-genome resequencing[J].Journal of fish biology,2025,107(5):1752-1764.
32. Wang L S ,Luo F L ,Yu Y , et al.Whole-genome re-sequencing reveals genetic structure and selection signals of different populations in *Megalobrama mblycephala*[J].Aquaculture, 2025,595(P1):741548-741548.
33. He X ,Wang S ,Teng W , et al.Genetic diversity and differentiation of cultured and wild yesso scallop (Mizuhopecten yessoensis) from China revealed by whole genome resequencing[J]. Aquaculture Reports, 2025,43:102901-102901.
34. Liu C ,Wen H ,Zhang C , et al.Deciphering the genetic basis and genomic prediction of heat tolerance trait from whole-genome resequencing in spotted sea bass (*Lateolabrax maculatus*)[J].Aquaculture,2025,598:741951-741951.
35. Wang K F ,Zhu C P ,Zhang T X , et al.QTL mapping and candidate gene identification for lower temperature tolerance and growth traits using whole genome re-sequencing in *Argopecten scallops*[J].Aquaculture,2025,595(P1):741513-741513.

 doi:10.20944/preprints202604.1282.v1

36. Song H ,Li H ,Wang W , et al. Machine learning–assisted feature selection from whole-genome sequencing data improves trait prediction and reveals candidate genes in sturgeon[J].Aquaculture,2026,618743830-743830.
37. Yan T ,Zheng X ,Chang M , et al. Population structure based on whole-genome resequencing and adaptive evolutionary mechanisms of circadian entrainment in the endangered fish Percocypris pingi. [J]. BMC genomics, 2026, 27 (1): 168-168. DOI:10.1186/S12864-025-12500-1.
38. Magati M T ,Yang P ,Yang K , et al. Proliferation of an invasive fish in a large reservoir from the downstream evidenced by whole-genome resequencing data[J].Hydrobiologia, 2026,(prepublish):1-17. DOI:10.1007/S10750-026-06117-Y.