Aquaculture 615 (2026) 743606

Contents lists available at ScienceDirect

Aquaculture

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

Genomic insights into *Hypophthalmichthys nobilis* in National Fish Stock Farms (NFSF): population structure, adaptive divergence and germplasm conservation

Qianhong Gu [a,1], Dengqiang Wang [b,1], Weitong Xu [b], Ziyi Wang [a], Liang Guo [a], Hang Sha [b], Hongwei Liang [b], Chuangju Li [b], Xinbin Duan [b,c,*]

[a] *Engineering Research Center of Polyploid Fish Reproduction and Breeding of the State Education, Ministry, College of Life Sciences, Hunan Normal University, Changsha 410081, Hunan, People's Republic of China*

[b] *Yangtze River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Wuhan 430223, Hubei, People's Republic of China*

[c] *Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River Key Laboratory of Sichuan Province, Yangtze River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Neijiang 641100, Sichuan, People's Republic of China*

ARTICLE INFO

Keywords:
Whole-genome resequencing
Population structure
Genetic diversity
Signatures of selection
Hypophthalmichthys nobilis

ABSTRACT

The ecologically and economically vital *Hypophthalmichthys nobilis* has been bred in National Fish Stock Farm (NFSF) in China since the 1960s to conserve genetics and combat overfishing. This long-term domestication has likely caused selection-driven genomic changes. To investigate this genetic diversity and selection pressure, a whole-genome resequencing was conducted on 341 NFSF *H. nobilis* from 13 populations and 21 wild counterparts from 3 populations (WP). Population genomics delineated eight genetically distinct populations (AN1–AN8), with six endemics to five NFSF populations and two (AN5, AN8) widespread in both NFSF and WP populations. Low inter-population F_{ST} versus high inter-strain F_{ST} implies NFSF's key role in preserving *H. nobilis* genetic diversity, particularly evident in higher within-NFSF diversity levels compared to natural populations. The divergence in these genetic populations predated the structured breeding and domestication efforts that began during the Tang Dynasty, indicating natural lineage divergence in *H. nobilis*. Genomic scans detected hundreds to thousands of selective sweeps per strain versus WP, including 30 divergent genes associated with environmental stress adaptation and immune defense. Furthermore, 75 genes linked to growth, development, and immunity—notably *mTOR*, *eif4g*, *kdm4a*, *atp9b*, *fgf12*, and interleukin receptors (*il21r*, *il4r*)—showed selection signals in domestication and improvement. Population genomics reveals NFSF-sustained adaptive divergence in *H. nobilis* driven by environmental stressors and immune adaptation, with candidate genes enabling genetic improvement of key traits (growth, disease resistance, and production). These findings establish a conservation genetic paradigm: NFSF populations sustain adaptive diversity in ex situ-managed *H. nobilis*, guiding vital fish germplasm stewardship.

1. Introduction

The establishment of living germplasm bank for indigenous species of notable wild species is an essential ecological protection strategy. This measure is especially crucial in preserving aquatic biodiversity and ensuring the sustainable utilization of fishery resources, given the widespread impact of environmental changes and anthropogenic activities on fish populations (Danet et al., 2024; Su et al., 2021a, 2021b; Xiang et al., 2023). The creation of a secure habitat by the fish stock farm has safeguarded certain fish species against external harmful elements. The practice of restocking or enhancing aquaculture in fish farms not only aids in the preservation of germplasm resources, but also assists in population recovery, and contributes to the maintenance of genetic diversity through enhancement and release (Bartley and Bell, 2008; Bell et al., 2008; Ruiz-Luna et al., 2023). However, sustained monoculture could potentially lead to a decrease in fish genetic diversity. This phenomenon occurs when a small number of captive parents generate a large quantity of offspring, leading to heightened inbreeding and a

* Corresponding author at: Yangtze River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Wuhan 430223, Hubei, People's Republic of China.
E-mail address: duan@yfi.ac.cn (X. Duan).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.aquaculture.2025.743606
Received 30 August 2025; Received in revised form 24 December 2025; Accepted 27 December 2025
Available online 29 December 2025
0044-8486/© 2025 Published by Elsevier B.V.

diminished effective population size in the medium to long term (Laikre et al., 2010; Ryman and Laikre, 1991). For instance, the genetic diversity of the wild Chinese sturgeon and large yellow croaker has diminished significantly due to habitat destruction and overfishing, while only a small fraction of artificial propagation was conducted for enhancement and release (Li et al., 2024a, 2024b; Liu et al., 2018).

As a result, the diversification of living germplasm banks is of utmost importance in fish preservation and artificial breeding. The process of screening representative individuals to establish a core gene pool can significantly mitigate the risk of genetic drift. For example, numerous National Fish Stock Farms (NFSF) dedicated to the preservation of the original genetic characteristics of the "Four Domesticated Fish" have been established in China since the 1960s. Historically, these fish were the predominant species in the Yangtze River. However, their breeding populations have significantly diminished over the years, with numbers dropping to a staggering low of less than 10 % by the 1960s (Jian et al., 2021). As one of the "Four Domesticated Fish", *H. nobilis* is a key farmed fish in China, possessing significant economic and ecological importance (Qiu et al., 2009; Zhang et al., 2012; Zhang et al., 2023). Regrettably, due to environmental alterations and anthropogenic activities, the natural populations of this species have witnessed a substantial decline. It was found that *Hypophthalmichthys nobilis* populations from the middle and lower Yangtze River basin exhibited heterozygote deficiency and inbreeding degradation (Li et al., 2024a, 2024b). These populations exhibited lower genetic diversity compared to those in the upper stream of the Yangtze River.

The investigation into piscine conservation and aquaculture constitutes a vitally significant focal point within the realms of contemporary ecology and hydrobiology. This is especially the case when considering the backdrop of global climate alteration and the continuous attrition of biodiversity. Among the available tools, genomic technology offers robust techniques for preserving fish germplasm resources, evaluating population genetic diversity, and facilitating breeding. With the rapid development of whole-genome resequencing (WGR) technology, it enables researchers to conduct in-depth examinations of the genetics in various species (Lv et al., 2024; Song et al., 2023). This technology has been widely applied in various fields including agriculture, livestock farming, fisheries, and ecological conservation, and enabled researchers to uncover the genetic diversity and evolutionary history of various species, while also providing crucial insights into population structure, selective pressure, and adaptability (Andrews et al., 2023; Wang et al., 2022a, 2022b; Wei et al., 2021). Notably, research on fish species, such as the siamese fighting fish and cichlid (Ciezarek et al., 2022; Wang et al., 2022a, 2022b), has been particularly influential. These studies elucidate both the population structure and genetic diversity of these fish, while simultaneously contributing valuable genomic resources for their protection and breeding efforts.

With the breakthrough in the artificial breeding technology of the "Four Domesticated Fish" in 1950s, numerous fish stock farms for *H. nobilis* have been successively established along the Yangtze River basin. The NFSF formed the fundamental infrastructure for the aquatic seed industry, encompassing five key dimensions: germplasm preservation, breed quality assurance, industrial advancement, ecological rehabilitation, and research innovation. The NFSF promotes a "preservation-selection-promotion" continuum that ensures the security of aquatic product supply and the sustainable development of fishery. However, long-term artificial breeding in the NFSF might have affected the productive traits of *H. nobilis*. This provides an ideal model for examining the relationships between domestic animals and their wild ancestors or potential domestication events (Larson and Burger, 2013). While several investigations have been undertaken regarding the genetic diversity of wild *H. nobilis* populations, a comprehensive understanding of the germplasm resources of *H. nobilis* within the NSFS remains elusive. This study uses whole-genome resequencing (WGR) to analyze the population structure, divergent selection, and domestication signals of *H. nobilis* reared in NFSF. It aims to clarify the core role of NFSF in germplasm conservation, identify candidate genes associated with growth and immunity, and provide genomic-level theoretical support for the sustainable conservation and efficient utilization of fish germplasm resources. Meanwhile, it offers references for the genetic improvement of the 'Four Domesticated Fish'.

2. Materials and methods

2.1. Statement on animal subjects

All experimental procedures were conducted in compliance with relevant guidelines and regulations, and were reported following ARRIVE guidelines (https://arriveguidelines.org). Prior to dissection, *H. nobilis* were collected and deeply anesthetized with 100 mg/L MS-222 (Sigma-Aldrich, St. Louis, MO, USA), after which caudal pterygiophore were excised, snap-frozen in liquid nitrogen, and stored at −80 °C. This study received approval from Hunan Normal University's College of Life Sciences Animal Ethics Committee (No. 2024–825) for all piscine experimental protocols. Methodologies strictly adhered to national standards outlined in China's Laboratory Animal Management Principles. After taking small fin ray samples, these anesthetized *H. nobilis* were recovered and released back to their original ponds.

2.2. Fish sampling and genomic DNA extraction

Thirteen populations of *H. nobilis* were obtained from NFSF distributed in Yangtze River system (Ano_01–05, Ano_10–14, Ano_17) and Yellow River system (Ano_06 and Ano_08) (Table 1). In each population, 17 to 30 individuals were sampled and rapidly anesthetized using the tricaine methanesulphonate (MS-222). We measured nine morphological characteristics for each individual (see Additional file1: Table 2), including total length (TL), body length (BL), body height (BH), head length (HL), snout length (SL), eye diameter (ED), interorbital distance (ID), caudal peduncle length (CPL), caudal peduncle height (CPH). And a total of 341 caudal fin specimens were collected and promptly frozen using liquid nitrogen, subsequently stored at −80 °C for DNA extraction. Genomic DNA was isolated from caudal fins utilizing the DNeasy Blood & Tissue Kit (QIAGEN DNeasy®, Shanghai, China), and was subsequently quantified with a spectrophotometer (NanoDrop 2000, Thermo Fisher Scientific, USA) and 1.5 % agarose gels to confirm integrity.

2.3. Whole-genome resequencing

High-qualified DNA samples were used to build paired-end libraries with 350 bp insert size. The libraries were pooled equimolarly and sequenced using Illumina NovaSeq 6000 platform (Illumina Inc., USA), including end repair, A-tailing, adapter ligation, and PCR amplification. Size selection was performed with AMPure XP beads, and final libraries were validated using a Bioanalyzer to ensure insert sizes of 300–500 bp. A minimum sequencing depth of 10 × coverage was targeted for each individual, with raw data output in FASTQ format. Each individual generated an average of 28.12 Gb in reads, which were subsequently utilized for whole-genome genotyping. In order to compare with *H. nobilis* in wild populations, we combined the resequencing data (FASTQ format) deposited in China National GeneBank (CNGB, https://db.cngb.org/cnsa) for *H. nobilis* from Pearl River (Pearl_R, $n = 8$, accession number: CNR0224739- CNR0224746), Yangtze River (Yangtze_R, $n = 9$, CNR0224729- CNR0224738) and Amur River (Amur_R, $n = 4$, CNR0224747-CNR0224750) (Table 1).

Furthermore, we downloaded resequencing data (FASTQ format) of *H. molitrix* (n = 4), *Ctenopharyngodon idella* ($n = 3$), *Megalobrama amblycephala* (n = 3) deposited in China National GeneBank or Genbank (see Additional file1: Table S2) as the outgroup in the following. In addition, we utilized novel resequencing data derived from three *Onychostoma sima* specimens, procured utilizing the methods previously delineated, as an outgroup in phylogenetic analysis.

Table 1
Sample localities and genetic diversity index of *H. nobilis* in this study, 13 populations were collected from National Fish Stock Farms (NFSF), the data of three wild populations (WP) were downloaded from China National GeneBank (CNR0224729- CNR0224750, https://db.cngb.org/cnsa). Eight genetic populations were found for all samples of *H. nobilis*.

	Populations	Latitude	Longitude	Site	N	π	H_O	F_{IS}	AR
	Ano_01	29.835013	112.488212	Hubei Province	21	5.37E-05	0.3056	−0.0185	1.5020
	Ano_02	29.588292	113.018255	Hubei Province	17	5.67E-05	0.3257	−0.0520	1.5366
	Ano_03	30.772996	114.488214	Hubei Province	22	5.37E-05	0.3420	−0.1007	1.5029
	Ano_04	30.331427	112.246258	Hubei Province	28	5.56E-05	0.3105	−0.0439	1.5171
	Ano_05	28.284797	113.013445	Hunan Province	28	5.40E-05	0.3101	−0.0489	1.5014
	Ano_06	35.094024	110.305033	Shanxi Province	28	5.86E-05	0.3096	−0.0256	1.5455
	Ano_08	38.890283	116.1103	Hebei Province	29	5.89E-05	0.3274	−0.0593	1.5500
	Ano_10	29.766441	115.603133	Jiangxi Province	28	5.78E-05	0.3064	−0.0266	1.5382
	Ano_11	32.303422	119.489938	Jiangsu Province	27	5.79E-05	0.3107	−0.0329	1.5403
	Ano_12	30.929789	120.602031	Jiangsu Province	30	5.71E-05	0.3013	−0.0206	1.5308
	Ano_13	30.925933	120.724113	Zhejiang Province	26	5.49E-05	0.3233	−0.0716	1.5106
	Ano_14	29.798101	113.439594	Hubei Province	27	5.67E-05	0.3125	−0.0439	1.5291
NFSF	Ano_17	30.145463	103.77089	Sichuan Province	30	5.48E-05	0.3151	−0.0628	1.5097
	Pearl_R	–	–	Pearl River	8	5.60E-05	0.3648	−0.0419	1.5452
	Yangtze_R	–	–	Yangtze River	9	5.77E-05	0.3578	−0.0481	1.5440
WP	Amur_R	–	–	Amur River	4	5.82E-05	0.4412	−0.0736	1.5516
	AN1				6	5.81E-05	0.1963	0.1532	1.3609
	AN2				9	3.90E-05	0.2515	−0.0313	1.4606
	AN3				18	4.98E-05	0.2656	−0.0453	1.4830
Genetic populations	AN4				15	5.16E-05	0.2580	−0.0246	1.4697
	AN5				206	5.08E-05	0.2983	−0.0301	1.5417
	AN6				12	4.68E-05	0.2362	0.0136	1.4303
	AN7				6	4.48E-05	0.2220	−0.0064	1.4109
	AN8				69	5.61E-05	0.2875	−0.0056	1.5221

Note: N, number of individuals in each population; π, Nucleotide diversity; H_O, Observed heterozygosity; F_{IS}, Inbreeding coefficient; AR, Allelic richness. — , unknown.

Table 2
Pairwise F_{ST} (lower triangle) and absolute genomic divergences D_{XY} (upper triangle) among genetic populations and WP calculated based on SNPs (159,140).

	AN1	AN2	AN3	AN4	AN5	AN6	AN7	AN8	Pearl_R	Yangtze_R	Amur_R
AN1		0.0305	0.0300	0.0303	0.0303	0.0306	0.0294	0.0304	0.0298	0.0298	0.0300
AN2	0.2387		0.0305	0.0307	0.0307	0.0310	0.0306	0.0307	0.0302	0.0302	0.0303
AN3	0.1997	0.1356		0.0304	0.0304	0.0307	0.0302	0.0305	0.0299	0.0299	0.0300
AN4	0.2134	0.1451	0.1222		0.0306	0.0309	0.0305	0.0307	0.0302	0.0302	0.0302
AN5	0.1251	0.0683	0.0530	0.0595		0.0309	0.0304	0.0307	0.0301	0.0301	0.0302
AN6	0.2465	0.1721	0.1458	0.1545	0.0770		0.0307	0.0307	0.0302	0.0303	0.0305
AN7	0.2685	0.1991	0.1673	0.1784	0.0931	0.2091		0.0305	0.0300	0.0300	0.0301
AN8	0.1425	0.0806	0.0677	0.0741	0.0136	0.0845	0.1095		0.0295	0.0299	0.0303
Pearl_R	0.1702	0.0946	0.0802	0.0888	0.0178	0.1069	0.1319	0.0104		0.0291	0.0296
Yangtze_R	0.1584	0.0874	0.0713	0.0802	0.0107	0.1009	0.1209	0.0144	0.0109		0.0296
Amur_R	0.1855	0.1007	0.0772	0.0879	0.0075	0.1220	0.1425	0.0243	0.0166	0.0115	

Note: Pearl_R, Pearl River; Yangtze_R, Yangtze River; Amur_R, Yangtze River.

2.4. Read trimming and variant calling

Raw reads were quality-trimmed using Trimmomatic and mapped to the *H. nobilis* reference genome from GeneBank (GCA_037950665) via Burrows-Wheeler Aligner (BWA) v0.7.15 (Li and Durbin, 2009). And then the mapping results were transformed into BAM format and organized using SAMtools v1.3.1 (Li et al., 2009). Duplicate reads were marked and removed using Picard v1.54 (https://broadinstitute.github.io/picard/). Variant calling (SNPs, indels) was performed using the Genome Analysis Toolkit (GATK) v4.0 (Franke and Crowgey, 2020), followed by hard filtering (QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5, and ReadPosRankSum < −8.0); the variant missing rate was less than 0.1, and the coverage depth ranged between 5 and 100; and the minor allele frequency (MAF) > 0.05; and the process of linkage disequilibrium pruning was executed utilizing PLINK v1.9 (Purcell et al., 2007). The parameters applied included a window size of 10 kb, a step size of one SNP, and an r^2 threshold of 0.5. The final dataset contained 362 individuals and 159,140 unlinked SNPs for population genetic structure analysis, and we got an output vcf file containing genotypes of individuals at each locus.

2.5. Population genetic structure and phylogeny

We tested population structure using two approaches, including the assignment test and discrimination analysis of principal components (DAPC), to ascertain the number of distinct populations (*K*) represented by our SNP dataset. We used the block relaxation approach assignment test in the program ADMIXTURE v1.3 (https://anaconda.org/bioconda/admixture) to estimate the ancestry of individuals, testing cluster values of $K = 1$ to $K = 20$. To determine the most appropriate value of *K*, we conducted a 10-fold cross-validation and selected the estimate with the lowest error, and the output was visualized in the R (v4.3.0) package pophelper (Francis, 2017). The DAPC analysis was performed using the R package Adegenet (Jombart, 2008) to determine the optimal *K* value for the 16 populations (see Additional file1: Table S2), and the principal component analysis (PCA) on the allele frequencies from the filtered SNP dataset was firstly executed and then used the resulting synthetic variables for a discriminant analysis (DA).

SplitsTree v6.3.40 (Huson and Bryant, 2024) was used to perform an unrooted Neighbor-Net analysis on the total SNP dataset (159,140). As a result, 362 individuals were used for the analysis with 1000 bootstrap replicates. Population clustering analyses revealed eight genetically

distinct populations of *H. nobilis*, which were subsequently used to construct a maximum likelihood tree (ML tree) illustrating their phylogenetic relationships. FASTQ data from *H. molitrix* ($n = 4$), *C. idella* ($n = 3$), *M. amblycephala* (n = 3), *O. sima* (n = 3), and *H. nobilis* ($n = 128$) were realigned to the reference *H. nobilis* genome (GCA_037950665). The 128 samples of *H. nobilis* encompassed all individuals from genetic populations AN1-AN4 and AN6-AN7, as well as AN5 (represented by three individuals from each population) and AN8 (similarly, three individuals from each population). The resulting vcf file was filtered using VCFtools (−-maf 0.05, −-max-missing 0.9, −-hwe 0.001, −-minmeanDP 10). The final dataset (vcf), comprising 141 individuals (see Additional file1: Table S3) and 46,512 unlinked SNPs, was converted into a phylip file using the vcf2phylip.py script (https://github.com/edgardomortiz/vcf2phylip) (Ortiz, 2019). The optimal nucleotide substitution model was automatically determined from a pool of candidates through a comparative analysis of the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). This process was facilitated by the built-in ModelFinder algorithm of IQ-TREE2 (Minh et al., 2020). ML tree search is performed based on the optimal model, while branch support is evaluated by Ultrafast Bootstrap (UFBoot) with 1000 replicates (Hoang et al., 2018). The inferred tree was rooted with *O. sima* as outgroup. And the final tree file was visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/), which included labeling for branch support values.

Furthermore, we executed a comparative analysis of linkage disequilibrium (LD) patterns within the eight genetic populations, as determined by population clustering. LD decay was quantified by computing correlation coefficients (r^2) for all SNP pairs within a 300 kb window, utilizing the PopLDdecay v3.40 (Zhang et al., 2019). The visualization of the average r^2 values among all genotype pairs was facilitated utilizing RSCRIPT.

2.6. Demographic history

The demographic history of *H. nobilis* were reconstructed using the Pairwise Sequentially Markovian Coalescent (PSMC 0.6.5-r67) (Li and Durbin, 2011) model to infer the effective population sizes (N_e) with SNP sites. We used the consensus autosomal sequences in FASTQ format for *H. nobilis* as the input for the PSMC modelling, and utilized BCFtools to identify SNPs, eliminate sites with a read depth that was either above twice or below one-third of the mean value, and procure a diploid consensus sequence. The PSMC executable was employed with the following parameters: —N25 -t15 -r5 -p "4 + 25 × 2 + 4 + 6", and the generation time of 5 years for *H. nobilis* and a substitution rate of 1.3×10^{-9} per site per year were set for the PSMC run (Allio et al., 2017; Wang et al., 2024). Additionally, we generated 100 bootstrap replicates to approximate the variance in the effective population size.

We also used a sequential Markovian coalescent (SMC++) to infer the changes of population sizes in history from the filtered SNP dataset (159,140 SNPs). The VCF files for each chromosome were converted to SMC input files using the vcf2smc module (https://github.com/popgenmethods/smcpp) with mask files to avoid artifactual and low-coverage sites. The obtained SMC input files for all of the chromosomes were then used to run the SMC++ estimate with options –knots 15 and time points between 1000 and 1000,000, using a mutation rate of 1.3×10^{-9} per nucleotide per year. After getting the output, it was used for the SMC++ plot command to obtain the plotting data using a generation time of 5 years. SMC++ is uniquely designed to simultaneously infer population size histories and divergence times across diverged populations (Terhorst et al., 2017). Building on this core capability, we leveraged the tool to estimate divergence times between our genetically distinct *H. nobilis* populations, specifically adopting its 'clean split' model for this analysis—an approach well-suited for scenarios where post-divergence gene flow is minimal. This is the first time that three key assumptions have been established for the 'clean split' model. These assumptions are all grounded in either prior genomic research on related taxa or empirical findings from our dataset: (1) A generation time of 5 years is consistent with that adopted in relevant population genomic studies of *H. nobilis* (Li et al., 2024a, 2024b); (2) A mutation rate (μ) ranging from 1×10^{-9} to 2×10^{-9} per site per year is widely recognized as appropriate for population genetic analyses of cyprinid fishes (Wang et al., 2024; Li et al., 2024a, 2024b); (3) No gene flow between target genetic populations (AN1–AN4, AN5) is supported by genetic differentiation results (Fig. 1), justifying use of the 'clean split' framework (no post-divergence gene flow).

2.7. Nucleotide diversity and fixation index analysis

Basic population genetic statistics like observed heterozygosity (H_O), allelic richness (A_R) and inbreeding coefficient (F_{IS}) for each population or genetically distinct populations, derived from ADMIXTURE and DAPC results, were estimated using HIERFSTAT (https://cran.r-project.org/web/packages/hierfstat/index.html) in the R package. Nucleotide diversity (π) and fixation index (F_{ST}) were computed utilizing VCFtools, employing a sliding window of 10 kb and a step size of 1 kb (parameters: –window-pi 10,000 –window-pi-step 1000; −-window-fst 10,000 –window-fst-step 1000), ensuring consistency with the window parameters applied in selection signal detection. Furthermore, we calculated the absolute D_{XY} genomic divergence for each locus among various populations, as well as between genetic populations. To account for missing data, variant and invariant sites from the VCF files were amalgamated using BCFtools. Subsequently, the all-sites VCF was employed to compute the D_{XY} in pixy v.1.2.6 (Korunes and Samuk, 2021).

2.8. Detection of selective signatures

Upon identifying eight genetic populations of *H. nobilis* within NFSF, we employed the detection of selective sweeps between these genetic populations and wild populations to reveal divergent selection. It should be noted that the wild populations integrated specimens from the Pearl River, Yangtze River, and Amur River. Furthermore, a comprehensive genome-wide scan was performed to identify selective sweeps occurring during the domestication and artificial breeding processes. This was accomplished by contrasting allele frequencies of *H. nobilis* in wild populations with those in seven NFSF populations that solely contained AN5 and AN8. In the pursuit of identifying directional selective sweeps, F_{ST} and θ_π were calculated using VCFtools with a sliding window of 10 kb and a step size of 1 kb. The F_{ST} (genetic differentiation index) analysis, a measure of population differentiation, indicates the degree of divergence between populations, a larger value signifies a higher level of population differentiation and stronger selection pressure. The θ_π (nucleotide diversity) analysis, which focuses on population diversity, is commonly used to examine base polymorphism. And the high regional selection pressure typically results in low polymorphism. Utilizing the screened high-quality SNP information, we employed the θ_π ratio and the top 5 % F_{ST} values as thresholds for analyzing population selection signals. This approach targets genomic regions with high population divergence (as measured by F_{ST}) and divergent nucleotide diversity (θπ) between NFSF-reared genetic populations and wild populations (WP)—a pattern indicative of divergent selection. For this analysis, F_{ST} (genetic differentiation index) and θπ (nucleotide diversity) were calculated using VCFtools v0.1.16 (Danecek et al., 2011), with a 10 kb sliding window and 1 kb step size. Negative FST values—artifacts of pairwise population comparisons—were set to 0 to prevent invalid divergence estimates. For threshold determination:

(1) F_{ST} values were *Z*-transformed (Z-F_{ST}) via custom R scripts (v4.3.0) to normalize variation across the genome. The transformation formula was:

$$Z - F_{ST} = \frac{F_{ST} - \mu F_{ST}}{\sigma F_{ST}}$$

Where μF_{ST} = mean FST, σF_{ST} = standard deviation of F_{ST}. The top 5

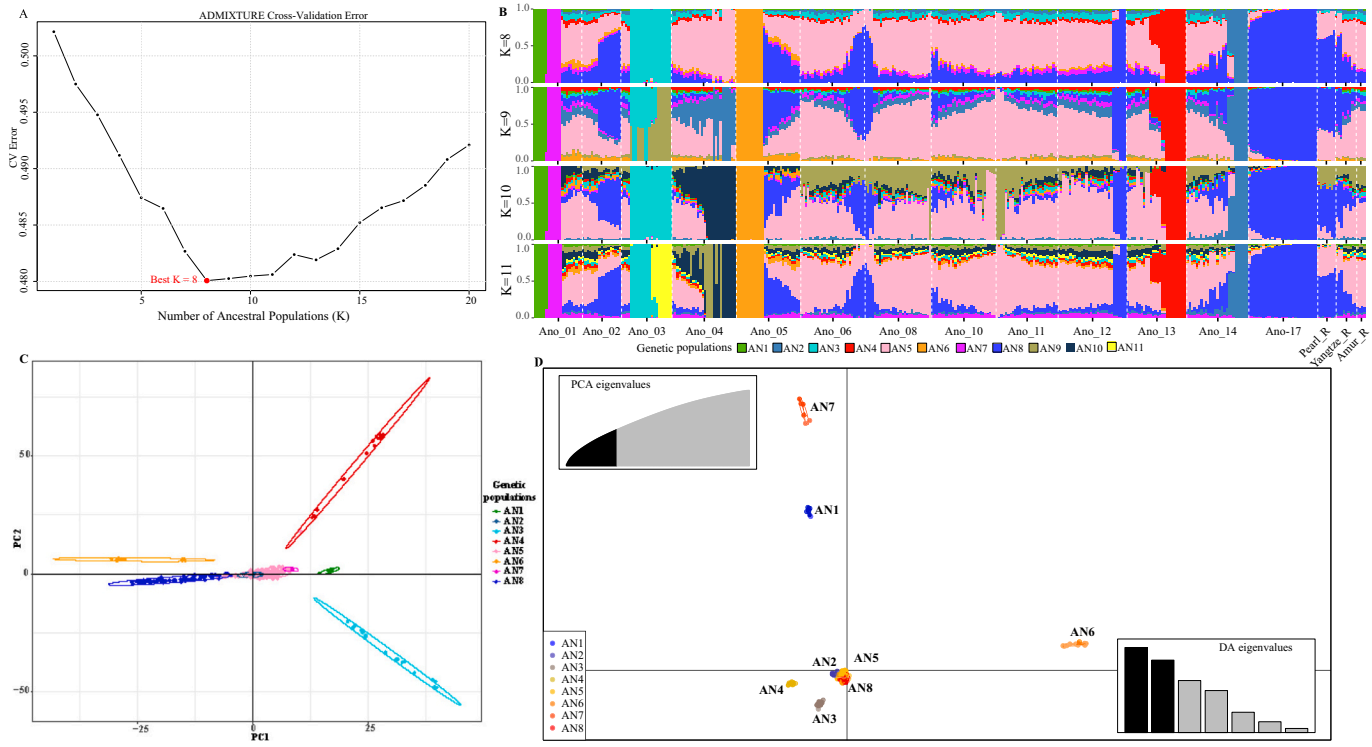


Fig. 1. Population genetic structure of the *H. nobilis* based on genome-wide SNPs (159,140). A, Cross-validation errors in admixture analysis; B, population structure inferred from ADMIXTURE, colors in each fragment represent the proportion of K = 8–11 ancestral populations assigned to individual genomes; C, population structure estimated by principal component analysis, and the PCA plots provide insights into the genetic variance among genetic populations; D, population cluster memberships from DAPC results for K = 8, clusters and genetic populations are shown by different colors and inertia ellipses.

% of Z- F_{ST} values were defined as high-divergence regions, consistent with commonly used thresholds for selection scans.

(2) The $\log_2$ratio of $\theta\pi = \log_2(\pi_{WP}/\pi_{AN})$ was computed to compare nucleotide diversity between wild populations (WP) and each NFSF-reared genetic population (AN). Regions with the top 5 % of $\log_{10}$ ratios (higher diversity in WP, suggestive of purifying selection in AN) and the bottom 5 % of $\log_{10}$ ratios (higher diversity in AN, suggestive of positive selection in AN) were classified as diversity-divergent regions. High-confidence selective regions were defined as the intersection of the top 5 % Z-F_{ST} regions and the extreme $\log_{10}(\theta\pi$ ratio) regions (top/bottom 5 %).

This dual-criteria approach reduces false positives from single-metric artifacts (e.g., high F_{ST} driven by genetic drift rather than selection) and ensures signals reflect both divergence (F_{ST}) and functional diversity ($\theta\pi$) changes. To control Type I errors (false positives from genome-wide window tests) and validate candidate gene list robustness, we applied the Benjamini-Hochberg (BH) procedure (FDR = 0.05) via R's p.adjust function. *P*-values (from *Z*-F_{ST} and $\log_{10}(\theta\pi$ ratio) statistics, calculated via 1000 permutation tests to assess window significance) were corrected, and only regions with adjusted $P < 0.05$ were retained for downstream candidate gene extraction. These results were visualized in scatter plots created using an R package. The genes located within these identified regions were extracted, thus suggesting their potential role in selection processes. Genes overlapping high-confidence selection regions (post-intersection and BH correction) were extracted via BEDTools v2.30.0 (Quinlan and Hall, 2010) by matching coordinates to the *H. nobilis* reference genome (GCA_037950665). Extracted genes were annotated with eggNOG-mapper v2 (Cantalapiedra et al., 2021) for functional descriptions, orthology groups, and protein domains. Candidate gene functional enrichment was analyzed using clusterProfiler v4.16.0 (Wu et al., 2021a, 2021b; https://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) for Gene Ontology (GO) terms (biological process [BP], molecular function [MF], cellular component [CC]) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. BH correction (FDR = 0.05) was applied to enrichment results to control multiple comparisons. Scatter plots of Z-F_{ST} vs. $\log_2(\theta\pi$-WP/$\theta\pi$-AN) were generated with ggplot2 v3.4.4 to visualize candidate regions (red) and non-selected regions (gray), allowing direct assessment of selection signal distribution and threshold stringency.

3. Results

3.1. Comparisons of morphometric traits

We conducted statistical analysis on ten morphometric indicators (BL/TL, BH/TL, HL/TL, SL/TL, SL/HL, ED/HL, ID/HL, ID/TL, CPL/TL, CPH/CPL) across genetic populations. The results showed there existing significant differences in seven morphometric characteristics (BH/TL, CPL/TL, ED/HL, HL/TL, ID/HL, ID/TL, and SL/HL) among most genetic populations ($P < 0.01$), with detailed data visualized in Additional file 2 (Fig. S8).

3.2. Summary of large-scale whole-genome resequencing and variation calling

Whole-genome resequencing (WGR) was performed on 341 individuals from thirteen populations of NFSF (Table 1), generating 9.27 Tb of raw base data with an average depth of 31.84 × per individual. Following quality filtering, 9.18 Tb of clean data were retained (see Additional file1: Table S2). The clean reads were then aligned to the *H. nobilis* reference genome (Jian et al., 2021), achieving mapping rates between 98.71 % and 99.77 %. The average depth of effectively mapped reads was 31×, with a range of 14.84× to 59.89× (see Additional file1: Table S2). Finally, 159,140 high-quality unlinked SNPs from 362 individuals (341 NFSF individuals +21 wild individuals) were obtained and used for subsequent population genetic analyses.

3.3. Genomic differentiation among H. nobilis from NFSF

Building on the 159,140 high-quality unlinked SNPs filtered in the preceding variation-calling workflow, we employed four complementary population genetic approaches to resolve the population structure of *H. nobilis*. And results across all methods were highly congruent, collectively validating the existence of eight genetically populations (AN1–AN8).

First, ancestry inference via ADMIXTURE identified K = 8 as the optimal number of genetic clusters (lowest cross-validation error; Fig. 1A), with this clustering pattern robust across 13 NFSF populations and 3 wild populations (Fig. 1B). Among the eight clusters, six (AN1–AN4, AN6–AN7) were exclusive to NFSF populations, while two (AN5, AN8) were detected in both NFSF and wild populations (e.g., Pearl_R, Yangtze_R, Amur_R). Notably, Ano_01 (one NFSF population) harbored two distinct clusters (AN1, AN7), whereas Ano_03, Ano_05, Ano_13, and Ano_14 each contained a single unique cluster (AN3, AN6, AN4, AN2, respectively); nearly all 16 total populations (13 NFSF +3 WP) included the admixed clusters AN5 and AN8—an observation consistent across all downstream analyses.

Second, principal component analysis (PCA) corroborated this structure: PC1 and PC2 (collectively explaining 18.7 % of genetic variance) effectively separated five of the eight clusters (AN1, AN3, AN4, AN6, AN8), with only minor overlap among AN2, AN5, and AN8 (Fig. 1C)—a pattern matching ADMIXTURE's admixed cluster signals.

Third, discriminant analysis of principal components (DAPC) further validated the eight clusters. After retaining 40 principal components (selected via lowest cross-validation RMSE), DAPC clearly delineated five isolated clusters (AN1, AN3, AN4, AN6, AN7) and three admixed clusters (AN2, AN5, AN8; Fig. 1D)—exactly mirroring the ADMIXTURE and PCA results.

Fourth, phylogenetic analyses (IQ-TREE maximum likelihood tree and SplitsTree Neighbor-Net network) reinforced the eight-cluster model. As shown in the IQ-TREE (Fig. 2A), six genetic populations (AN1–AN4, AN6, AN7) formed independent clades with ultrafast bootstrap $\geq$64 % (up to 100 %). While AN5 and AN8 formed paraphyletic groups, consistent with their admixed genetic background and widespread occurrence in both NFSF and wild populations. The SplitsTree network similarly grouped individuals into eight clusters, with no conflicting topological signals relative to ADMIXTURE, PCA, or DAPC (Fig. 2B).

To further confirm the reliability of this differentiation, we calculated genetic divergence metrics (F_{ST} and D_{XY}). The F_{ST} values between wild populations and AN5 (0.007–0.018) or AN8 (0.010–0.024) were significantly lower than those between wild populations and the other six NFSF-exclusive clusters (0.071–0.186; Table 2)—aligning with the clustering results that AN5/AN8 are shared with wild populations. Additionally, pairwise D_{XY} values among all eight clusters were consistently low (0.0295–0.0310; Table 2), indicating a recent common ancestry—yet the clear clustering across methods confirms these are genetically distinct units. The consistency of results across ADMIXTURE, PCA, DAPC, and phylogenetic analyses eliminates potential biases from single-method artifacts, providing robust evidence for eight genetically distinct populations of *H. nobilis* and their differential distribution between NFSF and wild habitats.

3.4. Comparison of LD and genomic diversity

The characterization of linkage disequilibrium (LD) patterns provides valuable insights into the historical forces shaping population genetics, including genetic drift, selection, and demographic processes (Slatkin, 2008). Among the eight genetic populations and three wild populations of *H. nobilis*, genetic populations AN1, AN2, AN4, AN6, AN7, and the wild Amur_R exhibited slow LD decay and high LD levels. In contrast, genetic populations AN3, AN5, AN8, and the remaining two wild populations (Yangtze_R, Pearl_R) showed rapid LD decay and low LD—with this pattern being most pronounced in AN5 and AN8 (Fig. 3). Since rapid LD decay and low LD typically correlate with minimal selective pressure and effective gene flow, AN5 and AN8 have likely experienced little selective pressure, thereby retaining relatively high genetic diversity.

For groups with slow LD decay (AN1, AN7, AN4, AN6, AN2, Amur_R), we quantified LD intensity: AN1 had the highest LD levels, followed by AN7. Since high LD often stems from population bottlenecks (reducing genetic variation and increasing haplotype retention), this highlights a significant bottleneck effect as the key driver of slow LD decay in these two—supporting our initial LD pattern observations in the larger group. Conversely, AN5 and AN8 had the shortest LD decay distances.

Three wild populations exhibited mean nucleotide diversities (π) ranging from 5.60×10^{-5} to 5.82×10^{-5} (Table 1, see Additional file2: Fig. S1), significantly higher than those observed in genetic populations from NFSF, which ranged from 3.89×10^{-5} to 5.16×10^{-5} ($P < 0.05$; Table 1, see Additional file2: Fig. S1). Notably, this observation held true for all except two genetic populations (AN5 and AN8), which exhibited diversities comparable to those of the wild populations. Nevertheless, the average nucleotide diversity showed limited variation among the thirteen populations from NFSF, with values ranging from 5.37×10^{-5} to 5.89×10^{-5} (Table 1, see Additional file2: Fig. S2). Thus, most of genetic populations in NFSF may have lost substantial genetic diversity

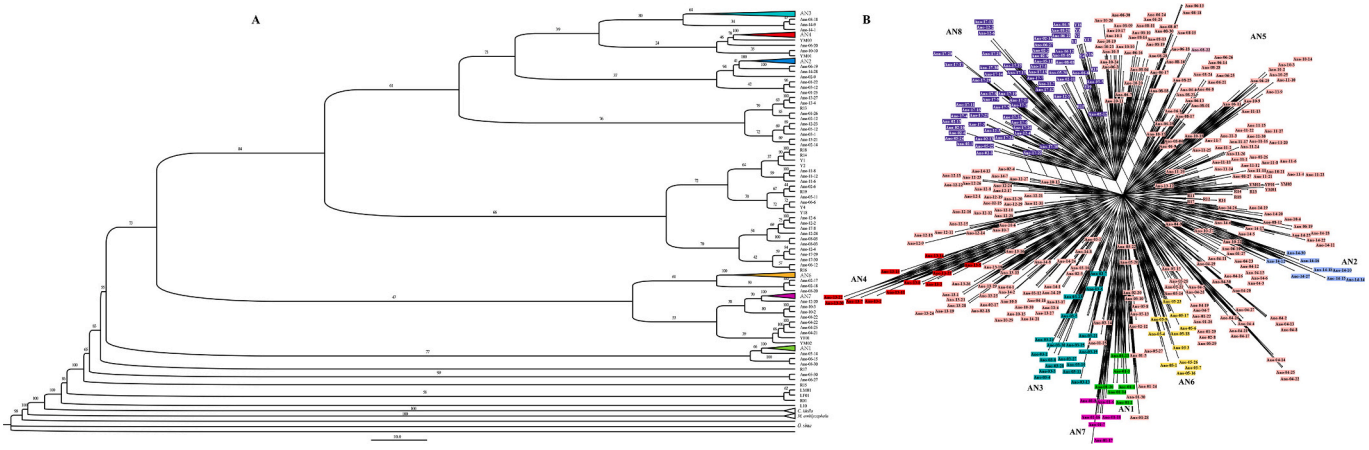


Fig. 2. Phylogenetic tree, the colors of the clades indicated different genetic populations. A, the maximum likelihood (ML) tree was reconstructed using IQ-TREE with 1000 bootstrap replications; B, unrooted phylogenetic network conducted in SplitsTree6 using the NeighborNet algorithm, and based on the 159,140 SNPs, illustrating genetic relationship among the 362 analyzed *H. nobilis*.

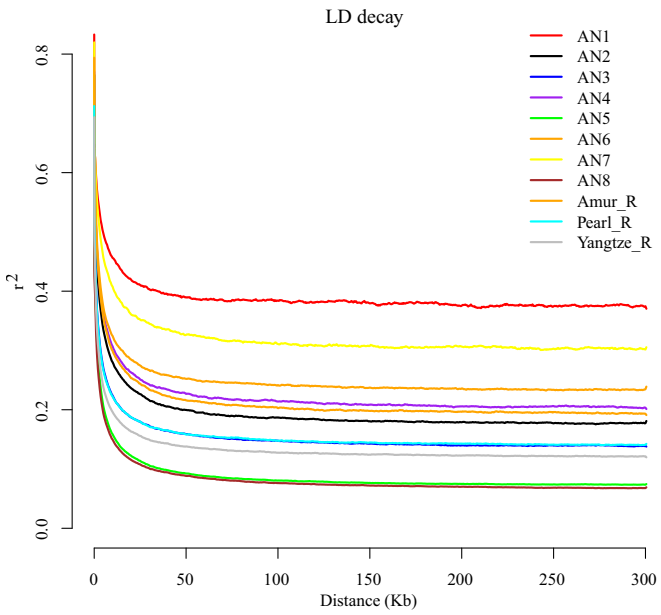


Fig. 3. LD decay across different genetic populations and wild populations.

due to inbreeding or founder effects. This observation was further supported by notably low values of H_O and AR when contrasted with those documented in wild populations (Table 1, see Additional file2: Fig. S3 and 4). In particular, the elevated F_{IS} was detected in AN1, while very low and negative values of F_{IS} were detected in the other genetic populations and wild populations (Table 1, see Additional file2: Fig. S5). Nevertheless, each NFSF population and wild population exhibited a negative F_{IS} value. The mean π value across the NFSF populations ranged from 5.37 × 10–5 to 5.89 × 10^{-5} (Table 1). The mean values of H_O in each NFSF population were significantly lower than those in the three wild populations, while no noticeable difference was observed in the mean value of AR.

3.5. Demographic history for H. nobilis

In order to gain a deeper understanding of the ancestral temporal dynamics associated with *H. nobilis*, we utilized the PSMC method to estimate the ancestral effective population sizes of this species. Additionally, we employed SMC++ to infer demographic history from genome-wide SNPs. This allowed us to compare the temporal dynamics of effective population sizes for each strain and wild population. Assuming a per-generation mutation rate of 0.7 × 10^{-8} and generation time of 5 for *H. nobilis*, the PSMC result suggested that *H. nobilis* underwent a rapid expansion, reaching its initial population peak (mean N_e: ~14.8 × 10^4) around 2.0 million years ago (Mya). Subsequently, it experienced a swift decline until 0.6–0.8 Mya, maintaining a low population level to the present day (Fig. 4A).

The inferred demographic history using SMC++ displayed inconsistent trends in effective population size changes over time for different genetic populations and wild populations (Fig. 4B). Nevertheless, a consistent decrease in effective population size over the past 12,000 years was observed for the majority of genetic populations (with the exception of AN2) and wild populations (excluding Yangtze_R). The effective population size of AN2 suffered from a rapid decline from 0.085 Mya to about 870 years ago, and then reached another peak approximately 60 years ago. Conversely, the effective population size of Yangtze_R exhibited a steady decrease until 4100 years ago when a slight expansion occurred. This subsequent growth has remained stable to present day.

Based on LD decay patterns, the inference suggested that AN5 and AN8 were indicative of the ancestral origins of the remaining six genetic populations. Therefore, SMC++ was used to estimate divergence times between AN5 and other genetic populations. Fortunately, the 'clean split' model showed no gene flow occurred after the four genetic populations (AN1-AN4) split. SMC++ estimates using the 'SMC++ split' model showed that the divergence times between AN1, AN2, AN3, AN4 (individually) and AN5 were 4055 years ago, 6757 years ago, 43,648 years ago, and 43,367 years ago, respectively (Fig. 5).

3.6. Selection signals and candidate genes

Given the extremely low F_{ST} value (0.0136) between AN5 and AN8 (Table 2), the two genetic populations were amalgamated into one single genetic population, denoted as AN5_8. Meanwhile, the three wild populations (Pearl_R, Yangtze_R, Amur_R) were pooled into a single group (WP) to facilitate subsequent selection signal detection. We employed a comprehensive method that integrates F_{ST} outliers and θ_π (nucleotide diversity) values to identify candidate genes responsible for the divergence between WP and the seven genetic populations (AN1, AN2, AN3, AN4, AN6, AN7, AN5/8). Genome-wide annotated SNPs were used to calculate π ratios and F_{ST} between wild population (WP) and these seven genetic populations, and the selection genes were identified using a 5 % threshold for *Z*-F_{ST} and $\log_{10}$ ($\theta_{\pi\text{-WP}}/\theta_{\pi\text{-AN}}$) ratios (see Additional file2: Fig. S6).

In total, 3865 genes exhibited signatures of selection across three or more comparisons between the WP and any given strain. However, only 30 genes were consistently selected across all seven comparisons

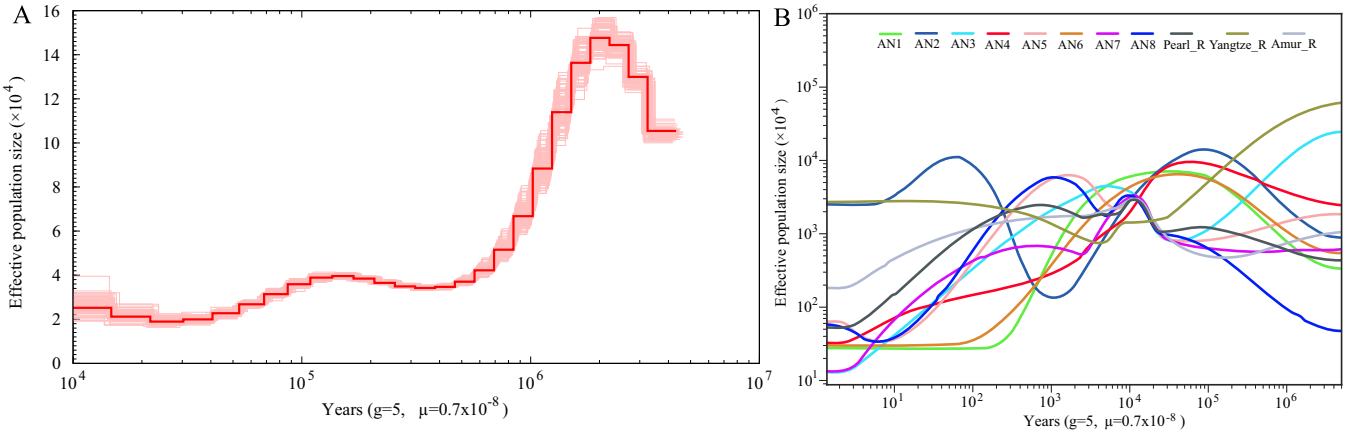


Fig. 4. Historical effective population size (N_e) estimated from PSMC and SMC ++ analysis for *H. nobilis* based on genome-wide SNPs. PSMC curves for one individual (Ano_01–8) using samtools pipeline, SMC++ curves of *H. nobilis* for all individuals in each strain and wild population.

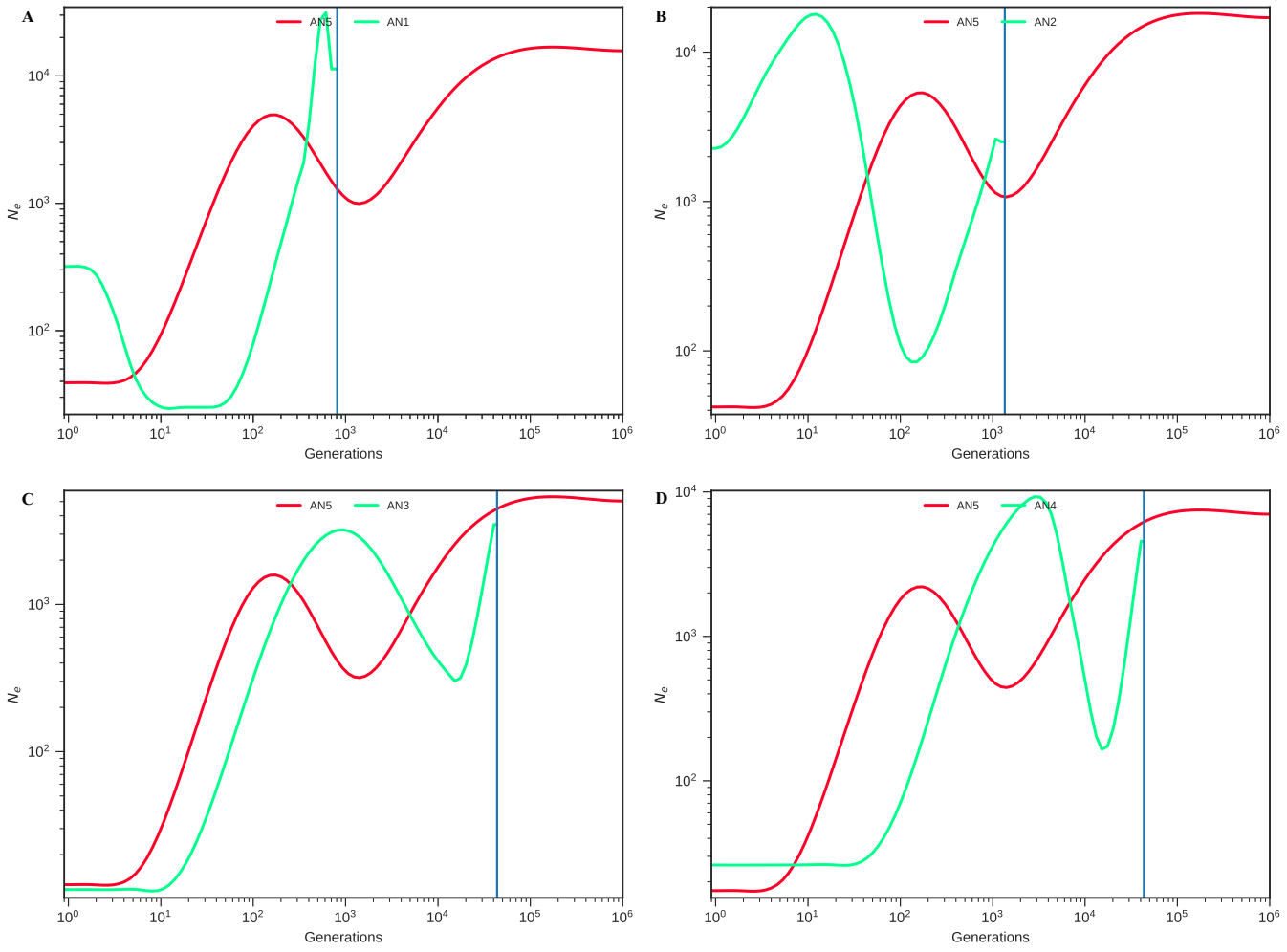


Fig. 5. Split time between AN5 and other four genetic populations estimated by SMC++ using the 'SMC++ split' model. A, the model estimated a divergence time of 4055 years ago for AN1 and AN5; B, 6757 years ago for AN2 and AN5; C, 43,648 years ago for AN3 and AN5; D, 43,367 years ago for AN4 and AN5.

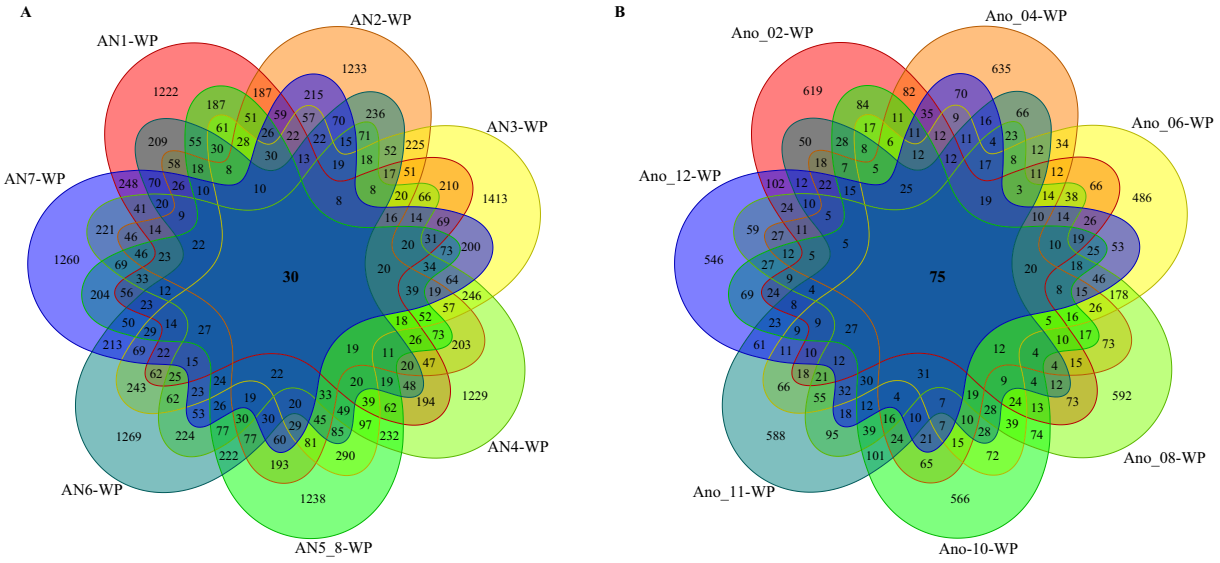


Fig. 6. Venn diagram that shows the genes identified under selection across both genetic populations and WP comparisons (A), and both seven NFSF populations and WP (B), highlighting shared and unique genetic targets of selection.

(Fig. 6A), suggesting these genes play a crucial role in the divergence of genetic populations. Notably, among the 30 genes identified under

selective sweep were *arap1* and *vti1a* (see Additional file1: Table S4), which play roles in endocytosis and autophagy, as well as the regulation of stem cell pluripotency, respectively. Specifically, *pkp4* stands out as a master regulator of cell adhesion and Wnt signaling. It exerts a significant influence on development, homeostasis, and disease progression by orchestrating β-catenin stability, Rho GTPase activity, and dynamics of cell connections. Additionally, genes such as Sarcoglycan, zeta (*sgcz*) underscore the adaptive responses of *H. nobilis* to cardiovascular disease.

In examining seven populations (Ano-02, Ano-04, Ano-06, Ano-08, Ano-10, Ano-11, Ano-12) from NFSF, we identified the same genetic populations (AN5/8) present in wild populations. Our objective was to elucidate the selection signatures influencing the adaptation of *H. nobilis* throughout domestication and improvement stages.

A genome-wide scan for selective sweeps using the 5 % thresholds for Z-F_{ST} and $\log_{10}(\theta_{\pi\text{-WP}}/\theta_{\pi\text{-AN}})$ ratios (see Additional file2: Fig. S7) identified 619 candidate genes unique to WP/Ano-02, 635 to WP/Ano-04, 486 to WP/Ano-06, 592 to WP/Ano-08, 566 to WP/ Ano-10, 588 to WP/ Ano-11, and 546 to WP/ Ano-12. Furthermore, we identified 75 genes consistently under selection across all these populations (Fig. 6B). Notably, functional annotation was applied to the 75 overlapping genes and a considerable proportion of these genes were revealed to be potentially involved in the domestication process. GO enrichment analysis and KEGG pathway annotation of these 75 genes highlighted crucial functions in eukaryotic translation, immune response, fatty acid beta-oxidation, spermatogenic, proteasome complex, protein homeostasis playing a pivotal role in reproduction, stress response, and immunity (Fig. 7A-B). Among the selected genes, several are implicated in growth, early development, and reproduction, as well as immune and nervous system responses (see Additional file1: Table S5). Notably, genes such as Eukaryotic translation initiation factor 4 gamma (*eif4g*), and ATPase phospholipid transporting 9B (*atp9b*), and especially fibroblast growth factor (*fgf12*) play crucial roles in the positive regulation of cell growth and proliferation, processes that are indispensable for growth and early development (Fig. 8). Interestingly, several genes are intricately associated with the regulation of penetrating pressure, the preservation of acid-base equilibrium, reproductive development, and adaptability to environmental fluctuations, such as gene solute carrier family 26 (*slc26a6*) and gene ADP-ribosyl cyclase (*cd38*) (Fig. 8). Other genes of interest that were identified were the mechanistic target of rapamycin (serine threonine kinase) (*mTOR*), mitogen-activated protein kinase (*mapk*), Interleukin 21 receptor (*il21r*), Interleukin 4 receptor (*il4r*) and ADP-ribosyl cyclase (*cd38*), all of which are pivotal in bridging adaptive and innate immunity (Fig. 8).

4. Discussion

The role of NFSF in preserving fish genetic resources remains inconclusive. In this study, we utilized population genomics analyses to examine and compare the genetic diversity, population structure, diversification, and selective signals of *H. nobilis* in wild populations versus those cultured in NFSF or cultured farms. Our results highlighted the importance of NFSF in maintaining genetically distinct populations of *H. nobilis*, suggesting that they possessed the same level of genetic diversity as natural populations despite potential founder effects, bottlenecks, and inbreeding. Additionally, we identified key genes associated with development and disease resistance that contribute to the sustainability of these genetically distinct populations of *H. nobilis*. We also evaluated the selective signatures in seven *H. nobilis* populations cultured in NFSF relative to their wild counterparts.

4.1. Implications for H. nobilis genetic populations identification

Genetic differentiation is a crucial indicator of the degree of genetic isolation among populations, while gene flow is universally recognized as a counterbalance to genetic differentiation (Broyles et al., 2023; Slatkin, 1987). Extensive population genetic studies over several decades have consistently confirmed that populations with higher dispersal capabilities exhibit lower levels of genetic differentiation (Benestan et al., 2021; Schiebelhut and Dawson, 2018). The species *H. nobilis* possesses the potential for rapid colonization of river systems, owing to its ability for swift long-distance swimming and its drifting-egg spawning strategy (Coulter et al., 2013; Jian et al., 2021). As a result, it is plausible that its population structure and differentiation have been significantly impacted by frequent gene flow. In line with our predictions, the population genetic analysis showed that the majority of populations of *H. nobilis* from the Yangtze River system exhibited low

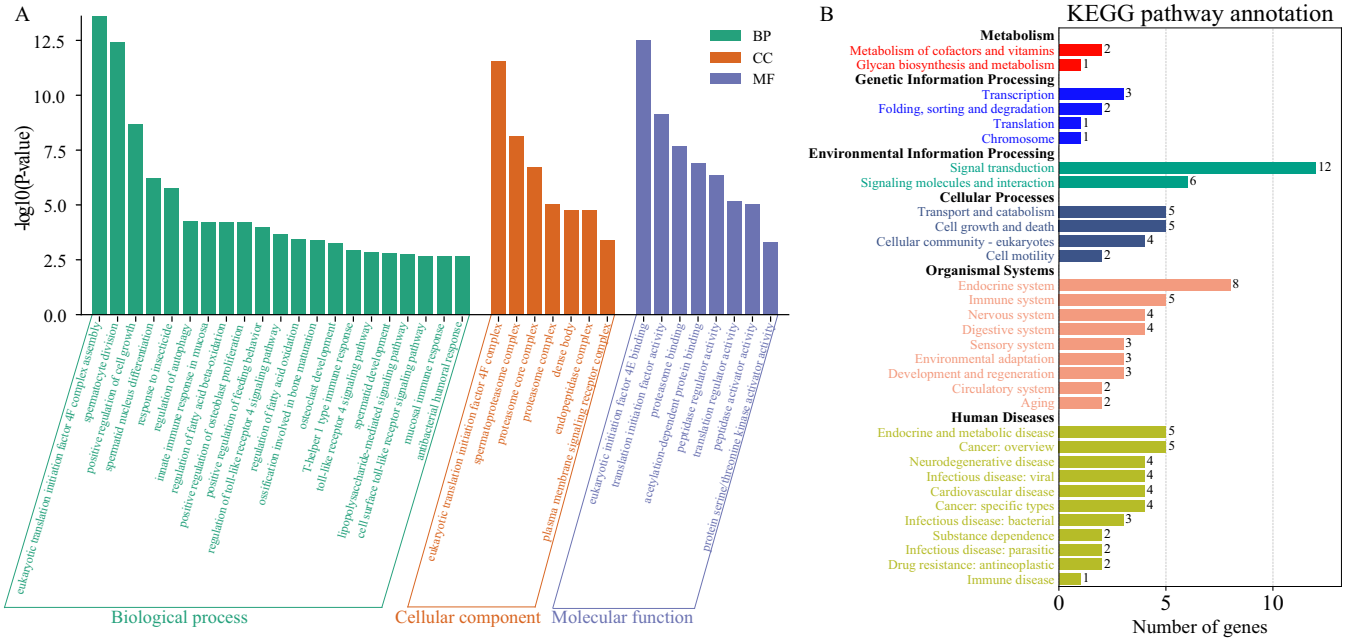


Fig. 7. Functional annotation of 75 overlapping genes among seven NSFS populations and WP. A, GO enrichmental analysis for overlapping genes; B, KEGG pathway enrichment statistics for overlapping genes.

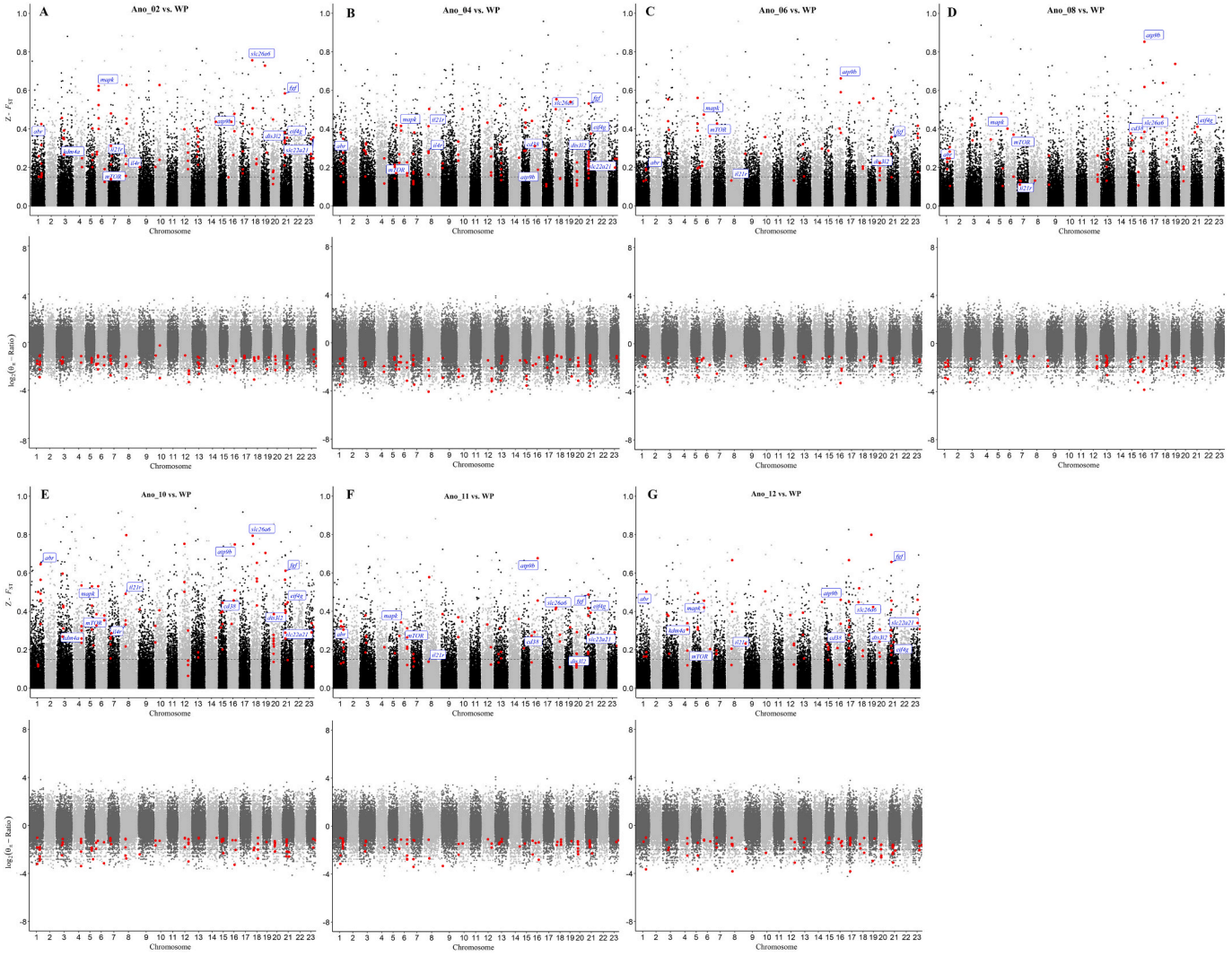


Fig. 8. Genomic signatures of selection detected in domestication and improvement. The genomic signatures of selection were detected by integration of nucleotide diversity (θ_π) and F_{ST} between seven NFSF populations (A, Ano_02; B, Ano_04; C, Ano_06; D, Ano_08; E, Ano_10; F, Ano_11; G, Ano_12) and WP. Red dots indicated loci under selective sweep. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

levels of genetic differentiation (Li et al., 2024a, 2024b). Additionally, our findings indicated no significant genetic differentiation among populations in different river systems, including Pearl River, Yangtze River and Amur River (Table 2). As anticipated, the populations of *H. molitrix* and *C. idella*, which share the same ecological and reproductive habits as *H. nobilis*, exhibited low genetic differentiation (F_{ST}) regardless of their locations within the upstream and downstream stretches of the Yangtze River or other river systems (Chen et al., 2009; Fang et al., 2021; Yang et al., 2022). However, attributing low levels of genetic differentiation entirely to anthropogenic translocations and stock enhancement may not be appropriate for these species. In contrast, despite extensive stock enhancement efforts for numerous economically important fish species in recent decades across various river systems in China, many of these species, including *Culter alburnus*, *Carassius auratus* and *Neosalanx taihuensis*, exhibit significant genetic differentiation and even distinct phylogeographical patterns (Fang et al., 2022; Gao et al., 2012; Gu et al., 2022; Qi et al., 2013). Their evolutionary history has been confirmed to not be the result of recent human-mediated translocations. These species lack the capabilities for long-distance swimming and drifting-egg spawning. As a result, the barrier to gene flow between distinct river systems is likely a prerequisite for their population divergence. The ADMIXTURE clustering analysis identified eight distinct genetic populations (AN1-AN8), six of which were genetically unique to the *H. nobilis* populations found in NFSF. These genetic relationships and population structures were further substantiated through the use of IQ-tree, Splitstree and PCA (Fig. 1C, Fig. 2). It is significant to note that both genetic populations AN5 and AN8 were not only detected in natural populations but were also consistently identified within each NFSF population (Fig. 1B). However, the phylogenetic analysis revealed that the six additional genetic populations each formed separate branches, yet they were all encompassed within the genetic populations AN5 and AN8 (Fig. 2A-B). The widespread distribution of AN5 and AN8 may be largely attributed to their ancestral lineage attributes, natural gene flow, and high adaptive potential. This view is supported by phylogenetic evidence of their ancestral status, by the connectivity maintained through the species' migratory and reproductive ecology, and by their conserved wild genetic diversity coupled with adaptive genes that enable persistence in multiple environments.

LD patterns provide valuable insights into evolutionary history, encompassing events such as bottlenecks and admixture, and even can be utilized to deduce potential ancestral populations (Deinum et al., 2015; Ding et al., 2025; Theunert et al., 2012). Post-'Out of Africa', non-African populations had bottlenecks, leading to higher LD than Africans (Haddrill et al., 2005; Xu et al., 2007). Xu et al. (2007) noted that Africans (e.g., the Yoruba) have lower LD (larger N_e, longer random mating), while non-Africans (e.g., Han, Europeans) have higher LD

(severe bottlenecks, e.g., founder effects). Haddrill et al. (2005) confirmed this via LD decay tests, noting non-Africans retain longer haplotypes (high-LD hallmark) due to reduced post-bottleneck recombination. Bottlenecks linking to higher LD are common in animals. Consuegra et al. (2005a, 2005b) showed small peripheral Atlantic salmon had higher LD than northern core groups (glacial bottlenecks reducing N_e). Ciezarek et al. (2022) noted farmed tilapia had higher LD than Lake Victoria wild populations (domestication bottlenecks, e.g., small broodstock). In rodents, Deinum et al. (2015) showed via comparative whole-genome data that bottlenecks regulate LD decay—wild *Rattus norvegicus* had slower LD decay than wild house mice (*Mus musculus castaneus*). Our *H. nobilis* genomic data fit the well-documented "bottleneck-high LD" evolutionary pattern. Specifically, among the 8 NFSF-identified genetic populations, AN1 and AN7 stand out with slower LD decay and higher overall LD levels (Fig. 3) than other NFSF populations (e.g., AN5 and AN8, which show the opposite trend). This distinct LD profile in AN1 and AN7 links to severe population bottlenecks—a key driver of elevated LD in natural and captive populations. Consistently, SMC++ analysis revealed a rapid decline in N_e for these two populations, directly confirming reduced genetic variation post-bottleneck (Fig. 4). In contrast, wild *H. nobilis* (e.g., Yangtze_R, Pearl_R) and ancestral-like AN5/AN8 show faster LD decay and lower LD (Fig. 3, likely larger N_e, minimal pressure). This pattern of rapid LD decay and low LD in AN5 and AN8 aligns with the typical traits of ancestral populations—specifically, the accumulation of more recombination events over time, which accelerates the breakdown of LD (Barton and Otto, 2005; Ortiz-Barrientos et al., 2016)—thereby supporting the hypothesis that AN5 and AN8 may serve as the ancestral lineages of the other six genetic populations of *H. nobilis*. This inference is further corroborated by our phylogenetic analyses: as illustrated in Fig. 2A and B, all six unique genetic populations exclusive to National Fish Stock Farms (NFSF) are nested within the clades corresponding to AN5 and AN8.

Furthermore, the majority of the genetic populations displayed a consistent, albeit varied, decrease in their effective population size (N_e) over the last 12,000 years (Fig. 5). Particularly, the fluctuations of N_e since 100 kya were nearly identical between AN5 and AN8, despite their very different N_e values prior to that time. SMC++ analysis estimated that the four genetic populations (AN1-AN4) underwent divergence from AN5 approximately 4055, 6757, 43,648, and 43,367 years ago, respectively (Fig. 5). This observation was largely consistent with the pattern of linkage disequilibrium decay, which remained high over long distances in recently divergent genetic populations but declined more swiftly in older genetic populations (Fig. 3). The level of LD serves as a "molecular clock" for evolutionary time. High LD indicates recent evolutionary events, such as domestication or bottlenecks, while low LD suggests long-term effects of random mating and recombination (Deinum et al., 2015; Serba et al., 2019).

While the 'SMC++ split' model successfully estimated the divergence times of AN1–AN4 from AN5, it failed for AN6/AN7, probably due to residual post-divergence gene flow (violating 'clean split' assumptions). Notably, all analyzable populations diverged prior to *H. nobilis* domestication. This observation underscores that natural lineage divergence, not artificial selection or breeding, is the main force driving the formation of *H. nobilis*' genetic structure. As the first attempt to apply the 'SMC++ split' model for inferring genetic divergence among different genetic populations of *H. nobilis*, this study directly fills a methodological gap in cyprinid population genetic analyses, particularly for taxa with germplasm resources conserved in captive breeding systems (e.g., NFSF). However, a key limitation of the current work lies in its exclusive reliance on genomic SNPs, with no employment of other molecular markers (such as mitochondrial genome markers); this omission precludes cross-validation of our conclusions through multiple complementary methodological approaches. Notably, this limitation further underscores the necessity of integrating diverse molecular markers—including mitochondrial DNA and microsatellites—with the 'SMC++ split' model in future research. Doing so would not only cross-verify the robustness of conclusions regarding *H. nobilis*' divergence times and genetic structure but also strengthen the reliability of inferences about its population divergence processes and genetic structure formation, thereby addressing the methodological gap identified in this study comprehensively.

4.2. Implications for preserving of *H. nobilis* in NFSF

Populations of smaller size typically exhibit lower levels of genetic variation (Bergl et al., 2008; Consuegra et al., 2005a, 2005b). Sustaining large-scale breeding populations for fish species within the NFSF poses a substantial challenge. In domesticated populations, low genetic diversity was anticipated due to the lack of gene flow from external populations and selective breeding practices (Zhang et al., 2016). Furthermore, these cultured populations might be at risk of experiencing a diminished genetic diversity, which could potentially arise from inbreeding or the bottleneck effect (Bradshaw et al., 2007; Chen et al., 2025; Moravciková et al., 2023). It was intriguing to observe that the population cultured in NFSF exhibited almost the same level of genetic diversity as observed in wild populations, particularly in Ano-08 and Ano-06, as evidenced by an elevated values of π and AR (Table 1, see Additional file2: Fig. S2). Nevertheless, it was noteworthy that markedly low levels of nucleotide diversity were detected in the genetically distinct populations of *H. nobilis* when compared to their wild counterparts, with the exception of genetic population AN5 and AN8 (Table 1, see Additional file2: Fig. S1). The lowest average values of π were found in AN1, followed by AN7. This coincided with the elevated LD levels and decreased LD decay distance detected in AN1 and AN7. Generally, a slow LD decay typically signifies low nucleotide diversity (π), whereas higher genetic diversity is indicative of a more rapid LD decay (Luo et al., 2012; Muñoz-Pérez et al., 2022). The demographic history, as inferred via SMC++, coupled with the observed LD decay, substantiated a rapid reduction in the effective population size for AN1 and AN7. This provided compelling evidence for the existence of either a bottleneck event or inbreeding (Deinum et al., 2015; Hu et al., 2023; Xie et al., 2025). Notably, the highest and positive value of F_{IS} was observed in AN1, with extremely lower and negative values detected in other genetic populations (Table 1). Notably, wild samples in this study primarily rely on published datasets, as current research on wild *H. nobilis* germplasm using whole-genome resequencing remains extremely limited—few studies have reported high-depth genomic data for this species' wild populations, thus limiting our ability to access additional wild samples. This resulting sampling imbalance, we acknowledge, may introduce biases—for instance, potentially leading to overestimated genetic diversity in wild populations (e.g., $\pi = 5.82 \times 10^{-5}$ for Amur_R). Accordingly, we suggest future studies expand the wild sample size to ≥30 individuals per population to validate our conclusions. In contrast, the larger NFSF sample size ($n = 341$) was specifically designed to capture genetic variation across breeding units from different river basins, a design critical to ensuring the accuracy of genetic population identification and subsequent captive population structure analyses.

In light of the preceding analysis, it is crucial to prioritize the preservation of genetic populations AN1 and AN7, both originating from the same population (Ano-01) of NFSF. In spite of the low genetic diversity of these six distinct genetic populations identified in NFSF, it served as a vital reservoir for maintaining the genetic diversity of *H. nobilis*. The two dominant genetic populations, AN5 and AN8, within the NFSF populations, showed no indications of having undergone bottlenecks or inbreeding depression. The established NFSF ex situ population of *H. nobilis* displayed different genetically unique reservoirs, demonstrating >90 % allelic richness retention compared to wild populations (Table 1). To mitigate genetic erosion in *H. nobilis*, three key measures serve as the cornerstone for maintaining the genetic viability of the species' germplasm recovery efforts: protecting extant gene pools, increasing genetic connectivity, and expanding breeding cohorts.

Despite these efforts, its unique gene pools remain threatened. Thus, the proposed genetic population-specific conservation units (e.g., priority populations AN1/AN7, which exhibit unique genetic signatures and high LD from bottlenecks) serve a foundational role in facilitating genetic connectivity—not as isolated repositories but as "genetic reservoirs". During controlled crossbreeding for connectivity enhancement, these units supply high-quality, distinct genetic material to prevent inbreeding depression in other populations.

To mitigate unique gene pool loss, we designed targeted measures: (1) Crossbreeding parents are selected by $F_{ST} < 0.05$, ensuring only genetically close populations are paired to minimize unique allele dilution and over-hybridization-induced allelic loss; (2) Highly unique populations (e.g., AN1/AN7) keep a non-hybridized "core subgroup" ($\geq$30 individuals) to preserve original gene pools, with a separate subgroup for connectivity; (3) Expanding breeding cohorts to $\geq$50 individuals per genetic population reduces rare allele loss from genetic drift. These three synergistic measures rely on preserving distinct populations for effective genetic connectivity, while targeted controls (selection thresholds, core subgroups) protect unique gene pools. Together, they ensure the genetic viability of *H. nobilis* germplasm recovery.

4.3. Divergence of selection signal within these genetic populations of H. nobilis

Our findings indicated that the genetically unique populations of *H. nobilis* in the NFSF originated primarily from divergent selection, not anthropogenic domestication. Preliminary evidence could be based on the divergence time between these genetic populations and AN5, which significantly far predated the period of human-induced domestication in China (Fig. 5). Existing literature indicated that the tradition of artificially breeding the Four Domesticated Fish in China (Jian et al., 2021), including *H. nobilis*, can be traced back to the Tang Dynasty (618–907 CE). During this period, the propagation of *H. nobilis* was largely dependent on natural rivers as their primary source of sustenance. This dependence emerged because the fish could not reproduce autonomously when cultured in ponds. Instead, they required large, fast-flowing rivers with appropriate hydrological conditions for successful reproduction (DeGrandchamp et al., 2007; Long et al., 2014). Consequently, the impact of artificial pond farming on their genetic differentiation was negligible. The significant turning point towards achieving full artificial reproduction and facilitating large-scale farming of *H. nobilis* occurred in the 1950s, merely 70 years ago.

The rapid long-distance swimming ability and drifting-egg spawning of *H. nobilis* enable it to rapidly colonize river systems (Coulter et al., 2013; Degrandchamp et al., 2008; Jian et al., 2021). Nevertheless, the capability of fish to sustain populations in novel environments is fundamentally a comprehensive indicator of their resistance to stress, encompassing molecular, physiological, behavioural, and ecosystem levels of adaptation. Our investigation into the genomic dynamics of *H. nobilis* lends empirical support to the hypothesis, underscoring the pivotal role of divergent selection for stress resistance in promoting differentiation among various genetic populations. Interestingly, our results identified eight genetic populations, primarily collected from the Yangtze River system and have preserved in different NFSF for nearly seventy years. This discovery underscores the prevalence of divergent selection likely occurring in natural populations of *H. nobilis*. It was further corroborated when four ancestors of this species were identified in the Yangtze River system (Li et al., 2024a, 2024b). *H. nobilis* primarily feeds on zooplankton, as well as phytoplankton, leading to the regular accumulation of microcystin (Chen et al., 2007). Hence, *H. nobilis* demonstrated a remarkable resilience to microcystin, as it can withstand the enduring stress of increased microcystin concentrations, without showing noticeable damage to the liver and kidneys (Zhang et al., 2012). Nevertheless, it induced metabolic disorders in the liver and dysfunction in the kidneys of the *H. nobilis* (Qiu et al., 2009). Consequently, *H. nobilis* residing for extended periods in eutrophic waters, such as lakes, usually exhibited enhanced liver antioxidant capacity due to the stress induced by microcystin toxins, compared to those in oligotrophic waters, such as rivers and reservoirs (Sun et al., 2013). These ecological and biological characteristics might have enhanced the divergence of the *H. nobilis* through divergent selection. The analyzed genetic populations displayed selection signatures suggestive of positive selection pressure on several key genes, such as *arap1*, *vti1a*, *pkp4*, sgcz, *rstk* (see Additional file1: Table S4). Particularly notable is the gene Immunoglobulin V-set domain (see Additional file1: Table S4), which was proposed to augment antigen-binding specificity via genetic rearrangement in the catfish *Ictalurus punctatus* (Bengtén et al., 2006).

These findings align with our understanding that both development and disease resistance are vital for survival and reproduction, and thus promote the persistence and divergence of genetic populations. Although there was a paucity of direct research examining the functions of these genes in fish, their potential roles could be inferred from the conserved functions of these genes in vertebrates. It was postulated that the majority of these genes confer survival advantages by facilitating responses to environmental stressors, immune defense, osmotic regulation, and developmental control (Gan et al., 2023; Mandiyan et al., 1999; Sun et al., 2024; Zhang et al., 2020a, 2020b), which is crucial for the successful colonization of new environments by *H. nobilis*. In future research, it is imperative to integrate gene editing techniques with multi-omics technologies for a comprehensive analysis of its functionalities. This integration will offer novel targets for disease prevention and management in aquaculture, while also providing theoretical foundations for the cultivation of stress-resistant varieties.

4.4. Unveiling genomic signatures of selection under domestication and improvement

The historical data indicated that NFSF had been in operation for the preservation and breeding of *H. nobilis* for a minimum of 20 years, and potentially for as long as 70 years. For instance, Ano-02, established in the 1950s, is still active. As a result, sustained conservation and breeding initiatives could lead to detectable indications of selective pressure on the genome of *H. nobilis* within the NFSF. We investigated selection signals in seven NFSF populations, in which only contain genetic populations AN5 and AN8, to search their genomes and highlight genetic variants for the development of productive breeds. In line with our predictions, the results unveiled a multitude of genes potentially implicated in biological processes encompassing early development, growth, immunity, and environmental adaptation (Fig. 7), which may be susceptible to pressures from domestication and direct selection.

The growth rate and disease resistance are widely recognized as key factors for potential genetic enhancement in aquatic production (Gjedrem and Rye, 2018). These factors serve as critical economic indicators for evaluating germplasm quality, and can be documented when identifying prospective breeding individuals (Gjedrem and Robinson, 2014). Most genetic enhancement projects in fish have focused on traits related to development and growth (Abaho et al., 2025; Cleveland et al., 2020; Gjedrem and Rye, 2018; Sun et al., 2022)—a trend that includes *H. nobilis*, which has been listed as a new aquatic product (GS-01-003-2022) by the annual conference of the National Appraisal Committee of Aquatic Protospecies and Improved Varieties in China. In the current study, substantial morphological variations ($P < 0.01$) were observed across most populations, encompassing morphometric traits such as BH/TL, CPL/TL, ED/HL, HL/TL, ID/HL, ID/TL, and SL/HL (see Additional file2: Fig. S8). These differences highlight the nuanced morphological attributes that distinguish one population from another.

A multitude of candidate genes potentially associated with development and growth were identified in this study (see Additional file2: Table S5, Fig. 8). Gene *mTOR* (chromosome 6, CM075473; Fig. 8) is key for NFSF's *H. nobilis* superior growth breeding (captive viability). The mTOR pathway is evolutionarily conserved in growth/development (Valet et al., 2022; Zhu and Thompson, 2019). As a serine/threonine

kinase, it regulates protein synthesis, energy metabolism, cell proliferation and autophagy. These processes underpin vertebrate growth (Ben-Sahra and Manning, 2017; Liu and Sabatini, 2020; Nie et al., 2018; Yu et al., 2008). In fish, *mTOR* activation enhances protein synthesis and growth (Wu et al., 2021). Notably, fast-growing triploid crucian carp show elevated *mTOR* pathway activity in metabolically active tissues (Huang et al., 2023). Within the nutrient-rich NFSF environment, this mechanism likely enables *H. nobilis* to maximize growth efficiency by rapidly converting consistent feed into biomass. This contrasts with wild populations, where metabolic flexibility for coping with food scarcity is prioritized over rapid growth. Consequently, *mTOR* activation may directly underpin the enhanced growth phenotypes observed in certain NFSF populations—including distinct morphological traits—making it a high-priority candidate marker for selecting growth-optimized strains (e.g., morphological differences in body height/total length; Additional file2: Fig. S8). This aligns with the study's overarching goal of advancing genetic improvement in NFSF germplasm.

Complementing *mTOR*'s role, *eif4g* reinforces the study's focus on dissecting genetic mechanisms underlying NFSF-adapted growth. As a core component of the eIF4F complex, *eif4g* integrates mTORC1 signaling with mRNA structural features to prioritize the translation of growth-related genes under nutrient-rich conditions (Gentry et al., 2025; Truitt et al., 2015). This function is particularly relevant to NFSF's captive environment, where controlled feeding regimes are designed to maximize growth efficiency: *eif4g*'s selection signal suggests it may facilitate *H. nobilis*' adaptation to these conditions, ensuring consistent growth performance in captive populations. However, wild populations rely on more generalized translation machinery to adapt to fluctuating food quality. For germplasm management, this makes *eif4g* a complementary marker to *mTOR*—helping identify strains that not only carry growth-driving variants but also thrive under NFSF's breeding conditions, thus supporting both genetic improvement and conservation goals. Additionally, genes including *kdm4a*, a*tp9b*, and *fgf12* address the study's objective of preserving viable, functionally diverse germplasm in NFSF—by safeguarding early developmental potential, a prerequisite for maintaining sustainable captive populations. For *kdm4a*, functional studies in zebrafish show broad expression during embryonic development, with CRISPR/Cas9-mediated knockout reducing embryo survival by 26 h post-fertilization (Han et al., 2023). This conserved role suggests kdm4a may be critical for ensuring high embryonic viability in NFSF-bred *H. nobilis*—a key consideration for avoiding germplasm loss due to developmental failure. Similarly, *atp9b* (involved in cellular lipid transport) and *fgf12* (a regulator of cell proliferation and tissue morphogenesis) contribute to early growth and development, further supporting the study's goal of preserving functional genetic diversity: their selection signals indicate NFSF populations have retained genes essential for robust development, validating the farms' role in germplasm conservation. In NFSF, where juvenile survival is high (low predation and stable environmental conditions), enhanced early development supported by *fgf12* improves population recruitment efficiency—an advantage less critical in the wild, where juvenile mortality is driven by predation and environmental stress rather than developmental rate.

Disease resistance breeding underpins sustainable aquaculture, and for the ecologically and economically vital *H. nobilis*, no formal programs yet exist—this study identifies genomic targets via selection scans of NFSF populations, focusing on candidates' potential roles (based on cross-species conservation) to inform future immune-focused breeding efforts. The *mapk* gene exhibited significant selection signals in multiple NFSF populations (e.g., Ano_02, Ano_08; Fig. 8, Table S5), aligning with the study's goal of uncovering genes that could enhance *H. nobilis*' defense against aquaculture pathogens. In other fish species, MAPK pathways are well-documented for their roles in mediating pathogen resistance and maintaining immune homeostasis (Su et al., 2021; Yang et al., 2022): for instance, inhibition of the Trxlp-ASK1-Erk/p38-MAPK axis enhanced *Edwardsiella piscicida* pathogenesis in zebrafish larvae (Yang et al., 2019), and MAPK/Erk signaling supported adaptive immune responses during bacterial infection in Nile tilapia (Wei et al., 2020). Its selection in NFSF populations could reflect adaptive responses to pathogen exposure in captive environments (e.g., higher density leading to increased disease transmission), making *mapk* a potential target for breeding programs aiming to boost bacterial resistance. Notably, while direct functional validation in *H. nobilis* is lacking, the gene's presence in selection regions and conserved immune roles across fish highlight its utility as a preliminary marker for selecting strains with improved bacterial defense. The *il21r* was also identified with consistent selection signals in NFSF-reared *H. nobilis* (Table S5). In recent studies, *il21r* functions as a key immune signaling receptor, regulating both innate and adaptive immune responses (Yu et al., 2020; Yuan et al., 2023). A critical example is *C. idella*, where intraperitoneal injection of IL21 (the ligand for *il21r*) improved survival in individuals infected with *Aeromonas hydrophila* and reduced pathogen-induced tissue damage in the head kidney and intestine (Zhang et al., 2020). For *H. nobilis*, which faces similar threats from *Aeromonas hydrophila* in aquaculture settings, *il21r*'s potential to modulate anti-bacterial immunity could be transformative for breeding. In NFSF's high-density setting, these alleles improve resistance to crowding-induced pathogens by boosting antibody production and macrophage activation. In contrast, wild populations face lower pathogen exposure (sparse distribution) and rely more on broad-spectrum immune responses rather than targeted pathogen resistance.

Additionally, gene *il4r* showed robust selection signals in NFSF *H. nobilis* populations (Fig. 8, Table S5), offering another potential target for breeding strains with enhanced immune competence. Fish *il4r* shares structural homology with its mammalian counterpart and is hypothesized to regulate diverse immune processes in cyprinids, including Th2-type immune responses (critical for parasitic defense), macrophage polarization, and general anti-infection defense (Jiang et al., 2019; Wang et al., 2011). This broad functional scope is particularly relevant for *H. nobilis*, which faces threats from both bacterial pathogens (e.g., *Aeromonas*) and parasites (e.g., gill flukes) in captive settings. As with *mapk* and *il21r*, however, its specific role in *H. nobilis* immunity remains unvalidated; cross-species conservation provides a foundation for its potential utility, but future assays (e.g., in vitro immune challenge experiments) are needed to confirm its function.

Notably, the associations between the aforementioned candidate genes and *H. nobilis*' growth, development, or disease resistance are primarily inferred from functional annotations and studies in homologous species (e.g., zebrafish, Nile tilapia). Further functional validation—such as gene editing (e.g., CRISPR/Cas9-mediated knockout/overexpression), in vitro functional assays, or phenotype-genotype association analyses in larger *H. nobilis* populations—is necessary to confirm their actual roles in these traits and strengthen the robustness of our inferences.

5. Conclusions

Our population genomics revealed the existence of several distinct genetic populations of *H. nobilis* within the NFSF, which play a crucial role in preserving the genetic diversity of *H. nobilis* germplasm resources. The divergence among these genetic populations was not influenced by selection in domestication or improvement programs, but rather by environmental stressors and the immune defense experiencing divergent selection. This further highlights the critical role of NFSF in safeguarding *H. nobilis* germplasm resources and maintaining a comprehensive collection of genes responsible for environmental adaptability and stress resistance. However, numerous candidate genes were identified as being correlated with adaptive processes in captivity and economically significant traits. Intriguingly, the enrichment of several candidate genes correlated with production traits, growth, and disease resistance, suggesting potential significance for genetic improvement. More importantly, NFSF populations serve a key role in

preserving *H. nobilis*' adaptive genetic diversity—e.g., AN5/AN8 (with wild-like variation)—and unique genetic clusters (e.g., AN1, AN7), offering practical references for fish ex situ conservation. Notably, our wild sampling is limited to 3 river basins and small sample sizes, failing to fully capture natural genetic variation. Future work should integrate more wild population data to refine the "NFSF + wild" comprehensive conservation framework.

CRediT authorship contribution statement

Qianhong Gu: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Dengqiang Wang:** Writing – review & editing, Resources, Methodology, Formal analysis, Data curation. **Weitong Xu:** Resources, Investigation, Formal analysis, Data curation. **Ziyi Wang:** Writing – original draft, Methodology. **Liang Guo:** Methodology. **Hang Sha:** Data curation. **Hongwei Liang:** Resources. **Chuangju Li:** Resources. **Xinbin Duan:** Writing – original draft, Validation, Supervision, Resources, Project administration, Funding acquisition.

Ethics statement

The research and samplings were conducted in accordance with the Norwegian Animal Research Authority regulations and approved by the Hunan Normal University's College of Life Sciences Animal Ethics Committee (No. 2024–825).

Funding

The study was financially supported by the National Natural Science Foundation of China, Ministry of Water Resources of the People's Republic of China and China Three Gorges Corporation-Yangtze River Water Science Research Joint Fund Project (U2240214), the Central Public-interest Scientific Institution Basal Research Fund, CAFS (NO.2023TD09), and the 2025 Open Research Projects of Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River Key Laboratory of Sichuan Province (NJTCSC25–1).

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.

Acknowledgments

We thank the High-performance Computing Platform of Yuelushan Laboratory Aquatic Variety Breeding Center for providing computing resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aquaculture.2025.743606.

Data availability

The raw sequence data reported in this paper (PRJNA1278812) have been deposited on the NCBI Short Read Archive (SRA, SRR34018512-SRR34018528, SRR34653533-SRR34653561, SRR34698258-SRR34698284, SRR34715377-SRR34715398, SRR34716366-SRR34716392, SRR34970984- SRR34971011, SRR34971045-SRR34971131, SRR34971142-SRR34971167, SRR34971178-SRR34971207, SRR35733573-SRR35733593), other supporting data of this study are available from co-authors Dr. Qianhong Gu (gqh@hunnu.edu.cn) upon reasonable request.

References

Abaho, I., Kwikiriza, G., Atukwatse, F., Izaara, A.A., Ekwangu, J., Baguma, S.D., Kubiriba, J., Kasozi, N., 2025. Selective breeding for genetic improvement of *Nile tilapia* (*Oreochromis niloticus* Linnaeus, 1758) in Uganda: current status, challenges, and future perspectives. Animals 15 (2). https://doi.org/10.3390/ani15020142.

Allio, R., Donega, S., Galtier, N., Nabholz, B., 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. 34 (11), 2762–2772. https://doi.org/10.1093/molbev/msx197.

Andrews, K.R., Seaborn, T., Egan, J.P., Fagnan, M.W., New, D.D., Chen, Z.Q., Hohenlohe, P.A., Waits, L.P., Caudill, C.C., Narum, S.R., 2023. Whole genome resequencing identifies local adaptation associated with environmental variation for redband trout. Mol. Ecol. 32 (4), 800–818. https://doi.org/10.1111/mec.16810.

Bartley, D.M., Bell, J.D., 2008. Restocking, stock enhancement, and sea ranching: arenas of progress. Rev. Fish. Sci. 16 (1–3), 357–365. https://doi.org/10.1080/10641260701678058.

Barton, N.H., Otto, S.P., 2005. Evolution of recombination due to random drift. Genetics 169 (4), 2353–2370. https://doi.org/10.1534/genetics.104.032821.

Bell, J.D., Leber, K.M., Blankenship, H.L., Loneragan, N.R., Masuda, R., 2008. A new era for restocking, stock enhancement and sea ranching of coastal fisheries resources. Rev. Fish. Sci. 16 (1–3), 1–9. https://doi.org/10.1080/10641260701776951.

Benestan, L., Fietz, K., Loiseau, N., Guerin, P.E., Trofimenko, E., Rühs, S., Schmidt, C., Rath, W., Biastoch, A., Pérez-Ruzafa, A., Baixauli, P., Forcada, A., Arcas, E., Lenfant, P., Mallol, S., Goñi, R., Velez, L., Höppner, M., Kininmonth, S., Mouillot, D., Puebla, O., Manel, S., 2021. Restricted dispersal in a sea of gene flow. P Roy Soc B-Biol Sci. 288 (1951). https://doi.org/10.1098/rspb.2021.0458.

Bengtén, E., Quiniou, S., Hikima, J., Waldbieser, G., Warr, G.W., Miller, N.W., Wilson, M., 2006. Structure of the catfish IGH locus:: analysis of the region including the single functional *IGHM* gene. Immunogenetics 58 (10), 831–844. https://doi.org/10.1007/s00251-006-0139-9.

Ben-Sahra, I., Manning, B.D., 2017. mTORC1 signaling and the metabolic control of cell growth. Curr. Opin. Cell Biol. 45, 72–82. https://doi.org/10.1016/j.ceb.2017.02.012.

Bergl, R.A., Bradleys, B.J., Nsubuga, A., Vigilant, L., 2008. Effects of habitat fragmentation, population size and demographic history on genetic diversity: the Cross River gorilla in a comparative context. Am. J. Primatol. 70 (9), 848–859. https://doi.org/10.1002/ajp.20559.

Bradshaw, C.J.A., Isagi, Y., Kaneko, S., Brook, B.W., Bowman, D., Frankham, R., 2007. Low genetic diversity in the bottlenecked population of endangered non-native banteng in northern Australia. Mol. Ecol. 16 (14), 2998–3008. https://doi.org/10.1111/j.1365-294X.2007.03365.x.

Broyles, G.G., Myers, B.M., Friedman, N.R., Gawin, D.F., Mohd-Taib, F.S., Sahlan, P.G. M., Seneviratne, S.S., de Silva, N.C.G., Lekamlage, T.T.M., Hund, A.K., Scordato, E.S. C., 2023. Evolutionarily labile dispersal behavior and discontinuous habitats enhance population differentiation in island versus continentally distributed swallows. Evolution 77 (12), 2656–2671. https://doi.org/10.1093/evolut/qpad179.

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: functional annotation, Orthology assignments, and domain prediction at the metagenomic scale. Mol. Biol. Evol. 38 (12), 5825–5829. https://doi.org/10.1093/molbev/msab293.

Chen, J., Xie, P., Zhang, D.W., Lei, H.H., 2007. In situ studies on the distribution patterns and dynamics of microcystins in a biomanipulation fish -: bighead carp (*Aristichthys nobilis*). Environ. Pollut. 147 (1), 150–157. https://doi.org/10.1016/j.envpol.2006.08.015.

Chen, Q., Wang, C.H., Lu, G.Q., Song, X., Xu, J.W., Yang, Q.L., Li, S.F., 2009. Analysis of genetic variation in grass carp (*Ctenopharyngodon idellus*) from native and colonized regions using ISSR markers. Biochem. Syst. Ecol. 37 (5), 549–555. https://doi.org/10.1016/j.bse.2009.08.002.

Chen, H.M., Zhao, H., Zhu, Q.Y., Yan, C., Liu, Y.Q., Si, S., Jamal, M.A., Xu, K.X., Jiao, D. L., Lv, M.J., Wang, W., Zhao, H.Y., Chen, L., Wang, M.S., Wei, H.J., 2025. Genomic consequences of intensive inbreeding in miniature inbred pigs. BMC Genomics 26 (1). https://doi.org/10.1186/s12864-025-11333-2.

Ciezarek, A., Ford, A.G.P., Etherington, G.J., Kasozi, N., Malinsky, M., Mehta, T.K., Penso-Dolfin, L., Ngatunga, B.P., Shechonge, A., Tamatamah, R., Haerty, W., Di Palma, F., Genner, M.J., Turner, G.F., 2022. Whole genome resequencing data enables a targeted SNP panel for conservation and aquaculture of *Oreochromis* cichlid fishes. Aquaculture 548. https://doi.org/10.1016/j.aquaculture.2021.737637.

Cleveland, B.M., Gao, G., Leeds, T.D., 2020. Transcriptomic response to selective breeding for fast growth in rainbow trout (*Oncorhynchus mykiss*). Marine Biotechnol. 22 (4), 539–550. https://doi.org/10.1007/s10126-020-09974-3.

Consuegra, S., Megens, H.-J., Schaschl, H., Leon, K., Stet, R.J.M., Jordan, W.C., 2005a. Rapid evolution of the MH class I locus results in different allelic compositions in recently diverged populations of Atlantic Salmon. Mol. Biol. Evol. 22 (4), 1095–1106. https://doi.org/10.1093/molbev/msi096.

Consuegra, S., Verspoor, E., Knox, D., de Leániz, C.G., 2005b. Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. Conserv. Genet. 6 (5), 823–842. https://doi.org/10.1007/s10592-005-9042-4.

Coulter, A.A., Keller, D., Amberg, J.J., Bailey, E.J., Goforth, R.R., 2013. Phenotypic plasticity in the spawning traits of bigheaded carp (*Hypophthalmichthys* spp.) in novel ecosystems. Freshw. Biol. 58 (5), 1029–1037. https://doi.org/10.1111/fwb.12106.

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., McVean, G., Durbin, R., 2011. 1000 Genomes Project Analysis Group. The variant call format and VCFtools.

Bioinform. (Oxford, England) 27 (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
Danet, A., Giam, X., Olden, J.D., Comte, L., 2024. Past and recent anthropogenic pressures drive rapid changes in riverine fish communities. Nat. Ecol. Evol. 8 (3). https://doi.org/10.1038/s41559-023-02271-x.
DeGrandchamp, K.L., Garvey, J.E., Csoboth, L.A., 2007. Linking adult reproduction and larval density of invasive carp in a large river. T. Am. Fish. Soc. 136 (5), 1327–1334. https://doi.org/10.1577/t06-233.1.
Degrandchamp, K.L., Garvey, J.E., Colombo, R.E., 2008. Movement and habitat selection by invasive Asian carps in a large river. T. Am. Fish. Soc. 137 (1), 45–56. https://doi.org/10.1577/t06-116.1.
Deinum, E.E., Halligan, D.L., Ness, R.W., Zhang, Y.H., Cong, L., Zhang, J.X., Keightley, P. D., 2015. Recent evolution in *Rattus norvegicus* is shaped by declining effective population size. Mol. Biol. Evol. 32 (10), 2547–2558. https://doi.org/10.1093/molbev/msv126.
Ding, R.J., Yu, D., Yang, K., Wu, X.H., Liu, H.Z., 2025. 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. Ecol. Evol. 15 (1). https://doi.org/10.1002/ece3.70874.
Fang, D.A., Luo, Y.T., Xu, D.P., Yang, X.W., Wang, X.H., 2021. Relationship between genetic risk and stock enhancement of the silver carp (*Hypophthalmichthys molitrix*) in the Yangtze River. Fish. Res. 235. https://doi.org/10.1016/j.fishres.2020.105829.
Fang, D.A., He, M., Ren, Y.F., Luo, H., Zhou, Y.F., Jiang, S.L., You, Y., 2022. Assessment of genetic diversity of the Salangid, *Neosalanx taihuensis*, based on the mitochondrial COI gene in different Chinese River basins. Biology-Basel 11 (7). https://doi.org/10.3390/biology11070968.
Francis, R.M., 2017. Pophelper: an R package and web app to analyse and visualize population structure. Mol. Ecol. Resour. 17 (1), 27–32. https://doi.org/10.1111/1755-0998.12509.
Franke, K.R., Crowgey, E.L., 2020. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for genome analysis toolkit algorithms. Genom. Informa. 18 (1), e10. https://doi.org/10.5808/GI.2020.18.1.e10.
Gan, Z., Xu, X.L., Tang, S.S., Wen, Q.Q., Jin, Y., Lu, Y.S., 2023. Identification and functional characterization of protein kinase R (PKR) in amphibian *Xenopus tropicalis*. Dev. Comp. Immunol. 141. https://doi.org/10.1016/j.dci.2023.104648.
Gao, Y., Wang, S.Y., Luo, J., Murphy, R.W., Du, R., Wu, S.F., Zhu, C.L., Li, Y., Poyarkov, A.D., Nguyen, S.N., Luan, P.T., Zhang, Y.P., 2012. Quaternary palaeoenvironmental oscillations drove the evolution of the Eurasian *Carassius auratus* complex (Cypriniformes, Cyprinidae). J. Biogeogr. 39 (12), 2264–2278. https://doi.org/10.1111/j.1365-2699.2012.02755.x.
Gentry, R.C., Ide, N.A., Comunale, V.M., Hartwick, E.W., Kinz-Thompson, C.D., Gonzalez Jr., R.L., 2025. The mechanism of mRNA cap recognition. Nature 637 (8046). https://doi.org/10.1038/s41586-024-08304-0.
Gjedrem, T., Robinson, N., 2014. Advances by selective breeding for aquatic species: a review. Agric. Sci. 05 (12), 1152–1158. https://doi.org/10.4236/as.2014.512125.
Gjedrem, T., Rye, M., 2018. Selection response in fish and shellfish: a review. Rev. Aquac. 10 (1), 168–179. https://doi.org/10.1111/raq.12154.
Gu, Q.H., Wang, S., Zhong, H., Yuan, H., Yang, J.L., Yang, C.H., Huang, X.X., Xu, X.W., Wang, Y.D., Wei, Z.H., Wang, J., Liu, S.J., 2022. Phylogeographic relationships and the evolutionary history of the *Carassius auratus* complex with a newly born homodiploid raw fish (2*n*NCRC). BMC Genomics 23 (1). https://doi.org/10.1186/s12864-022-08468-x.
Haddrill, P.R., Thornton, K.R., Charlesworth, B., Andolfatto, P., 2005. Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. Genome Res. 15 (6), 790–799. https://doi.org/10.1101/gr.3541005.
Han, L., Luo, J., Qu, S., Shi, X., Zhang, J., Han, B., 2023. kdm4aa is required for reproduction and development of zebrafish. Aquaculture and Fisheries 8 (6), 713–719. https://doi.org/10.1016/j.aaf.2022.05.005.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35 (2), 518–522. https://doi.org/10.1093/molbev/msx281.
Hu, W.J., Hao, Z.Q., Du, P.Y., Di Vincenzo, F., Manzi, G., Cui, J.L., Fu, Y.X., Pan, Y.H., Li, H.P., 2023. Genomic inference of a severe human bottleneck during the early to middle Pleistocene transition. Science 381 (6661), 979–984. https://doi.org/10.1126/science.abq7487.
Huang, Z., Dai, L.Y., Peng, F.Y., Tang, L.W., Wang, X.J., Chen, J.Y., Liu, J.H., Fu, W., Peng, L.Y., Liu, W.B., Xiao, Y.M., 2023. mTOR signaling pathway regulates embryonic development and rapid growth of triploid crucian carp. Aquacult. Rep. 33. https://doi.org/10.1016/j.aqrep.2023.101860.
Huson, D.H., Bryant, D., 2024. The SplitsTree app: interactive analysis and visualization using phylogenetic trees and networks. Nat. Methods 21 (10), 1773–1774. https://doi.org/10.1038/s41592-024-02406-3.
Jian, J.B., Yang, L.D., Gan, X.N., Wu, B., Gao, L., Zeng, H.H., Wang, X.Z., Liang, Z.Q., Wang, Y., Fang, L.H., Li, J., Jiang, S.J., Du, K., Fu, B.D., Bai, M.Z., Chen, M., Fang, X. D., Liu, H.Z., He, S.P., 2021. Whole genome sequencing of silver carp (*Hypophthalmichthys molitrix*) and bighead carp (*Hypophthalmichthys nobilis*) provide novel insights into their evolution and speciation. Mol. Ecol. Resour. 21 (3), 912–923. https://doi.org/10.1111/1755-0998.13297.
Jiang, X.Y., Gao, J.D., Xue, Y.J., Qin, Y.T., Li, X., Sun, Z.S., Xie, H.X., Chang, M.X., Nie, P., Zou, J., Gao, Q., 2019. Identification and expression analysis of IL-4/13 receptors in grass carp *Ctenopharyngodon idella*. Fish Shellfish Immunol. 87, 254–264. https://doi.org/10.1016/j.fsi.2019.01.009.
Jombart, T., 2008. *Adegenet*: a R package for the multivariate analysis of genetic markers. Bioinformatics 24 (11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.
Korunes, K.L., Samuk, K., 2021. Pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol. Ecol. Resour. 21 (4), 1359–1368. https://doi.org/10.1111/1755-0998.13326.
Laikre, L., Schwartz, M.K., Waples, R.S., Ryman, N., Ge, M.W.G., 2010. Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. Trends Ecol. Evol. 25 (9), 520–529. https://doi.org/10.1016/j.tree.2010.06.013.
Larson, G., Burger, J., 2013. A population genetics view of animal domestication. Trends Genet. 29 (4), 197–205. https://doi.org/10.1016/j.tig.2013.01.003.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25 (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li, H., Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. Nature 475 (7357), 493–U484. https://doi.org/10.1038/nature10231.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data, P, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25 (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
Li, Y.Z., Sun, M., Yang, X.Y., Yang, M.L., Kleisner, K.M., Mills, K.E., Tang, Y., Du, F.Y., Qiu, Y.S., Ren, Y.P., Chen, Y., 2024a. Social-ecological vulnerability and risk of China's marine capture fisheries to climate change. Proc. Natl. Acad. Sci. U. S. A. 121 (1). https://doi.org/10.1073/pnas.2313773120.
Li, W.T., Yu, J.Y., Que, Y.F., Hu, X.K., Wang, E.Z., Liao, X.L., Zhu, B., 2024b. Population genetic investigation of *Hypophthalmichthys nobilis* in the Yangtze River basin based on RAD sequencing data. Biology-Basel 13 (10). https://doi.org/10.3390/biology13100837.
Liu, G.Y., Sabatini, D.M., 2020. mTOR at the nexus of nutrition, growth, ageing and disease. Nat. Rev. Mol. Cell Bio. 21 (4), 183–203. https://doi.org/10.1038/s41580-019-0199-y.
Liu, J., You, X.X., Xu, P., Zhuang, P., Zheng, Y.P., Zhang, K., Wang, M., Lv, Y.Y., Xu, G.C., Zhao, F., Wu, J.H., Fan, H.Y., Xu, J.A., Ruan, Z.Q., Bian, C., Liu, K., Xu, D.P., Chen, J. H., Xu, J.M., Shi, Q., 2018. Assessing the genetic diversity of the critically endangered Chinese sturgeon *Acipenser sinensis* using mitochondrial markers and genome-wide single-nucleotide polymorphisms from RAD-seq. Sci. China Life Sci. 61 (9), 1090–1098. https://doi.org/10.1007/s11427-017-9254-6.
Long, J.M., Liang, Y., Shoup, D.E., Dzialowski, A.R., Bidwell, J.R., 2014. GIS-based rapid-assessment of bighead carp *Hypophthalmichthys nobilis* (Richardson, 1845) suitability in reservoirs. Manag. Biol. Invasion. 5 (4), 363–370. https://doi.org/10.3391/mbi.2014.5.4.07.
Luo, N., Yu, X.Q., Liu, J.X., Jiang, Y.W., 2012. Nucleotide diversity and linkage disequilibrium in antioxidant genes of *Brachypodium distachyon*. Plant Sci. 197, 122–129. https://doi.org/10.1016/j.plantsci.2012.09.010.
Lv, S.C., Chen, L.L., Lamao, J.B., Jiumai, Z.X., Lemao, C.R., Ma, Y., 2024. Whole genome resequencing and its approach in population genomics. The Chinese Livestock and Poultry Breeding 20 (08), 46–52.
Mandiyan, V., Andreev, J., Schlessinger, J., Hubbard, S.R., 1999. Crystal structure of the ARF-GAP domain and ankyrin repeats of PYK2-associated protein beta. The EMBO J. 18 (24), 6890–6898. https://doi.org/10.1093/emboj/18.24.6890.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37 (5), 1530–1534. https://doi.org/10.1093/molbev/msaa015.
Moravciková, N., Kasarda, R., Zidek, R., McEwan, J.C., Brauning, R., Landete-Castillejos, T., Chonco, L., Ciberej, J., Pokorádi, J., 2023. Traces of human-mediated selection in the Gene Pool of Red Deer populations. Animals 13 (15). https://doi.org/10.3390/ani13152525.
Muñoz-Pérez, J.M., Cañas, G.P., López, L., Arias, T., 2022. Genome-wide diversity analysis to infer population structure and linkage disequilibrium among *Colombian coconut* germplasm. Sci. Rep-UK. 12 (1). https://doi.org/10.1038/s41598-022-07013-w.
Nie, X.G., Zheng, J.X., Ricupero, C.L., He, L., Jiao, K., Mao, J.J., 2018. mTOR acts as a pivotal signaling hub for neural crest cells during craniofacial development. PLoS Genet. 14 (7). https://doi.org/10.1371/journal.pgen.1007491.
Ortiz, E.M., 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. GitHub repository. https://doi.org/10.5281/zenodo.2540861. https://github.com/edgardomortiz/vcf2phylip.
Ortiz-Barrientos, D., Engelstädter, J., Rieseberg, L.H., 2016. Recombination rate evolution and the origin of species. Trends Ecol. Evol. 31 (3), 226–236. https://doi.org/10.1016/j.tree.2015.12.016.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575. https://doi.org/10.1086/519795.
Qi, P.Z., Guo, B.Y., Xie, C.X., Wu, C.W., Lu, S.M., Duan, Y.J., Zhou, X.J., 2013. Assessing the genetic diversity and population structure of *Culter alburnus* in China based on mitochondrial 16S rRNA and COI gene sequences. Biochem. Syst. Ecol. 50, 390–396. https://doi.org/10.1016/j.bse.2013.04.010.
Qiu, T., Xie, P., Guo, L., Zhang, D., 2009. Plasma biochemical responses of the planktivorous filter-feeding silver carp (*Hypophthalmichthys molitrix*) and bighead carp (*Aristichthys nobilis*) to prolonged toxic cyanobacterial blooms in natural waters. Environ. Toxicol. Phar. 27 (3), 350–356. https://doi.org/10.1016/j.etap.2008.12.001.

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26 (6), 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Ruiz-Luna, A., Berlanga-Robles, C.A., Vera, J.M., García, C.R.R., 2023. Pacific white shrimp (*Penaeus vannamei*) stock enhancement with aquaculture recruits: a case study in the coastal lagoons of southern Sinaloa, Mexico (2019-2020). Lat. Am. J. Aquat. Res. 51 (5), 671–684. https://doi.org/10.3856/vol51-issue5-fulltext-3022.

Ryman, N., Laikre, L., 1991. Effects of supportive breeding on the genetically effective population size. Conserv. Biol. 5 (3), 325–329. https://doi.org/10.1111/j.1523-1739.1991.tb00144.x.

Schiebelhut, L.M., Dawson, M.N., 2018. Correlates of population genetic differentiation in marine and terrestrial environments. J. Biogeogr. 45 (11), 2427–2441. https://doi.org/10.1111/jbi.13437.

Serba, D.D., Muleta, K.T., St Amand, P., Bernardo, A., Bai, G.H., Perumal, R., Bashir, E., 2019. Genetic diversity, population structure, and linkage disequilibrium of pearl millet. Plant. Genome 12 (3). https://doi.org/10.3835/plantgenome2018.11.0091.

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science (New York, N.Y.) 236 (4803), 787–792. https://doi.org/10.1126/science.3576198.

Slatkin, M. 2008. Linkage disequilibrium–understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. 9(6), 477–485. doi:https://doi.org/10.1038/nrg2361.

Song, B., Ning, W.D., Wei, D., Jiang, M.Y., Zhu, K., Wang, X.W., Edwards, D., Odeny, D. A., Cheng, S.F., 2023. Plant genome resequencing and population genomics: current status and future prospects. Mol. Plant 16 (8), 1252–1268. https://doi.org/10.1016/j.molp.2023.07.009.

Su, Y.L., Jiang, Y.F., Chen, H.J., Ye, S., Zhou, W.H., Liu, H.P., Dong, J.D., Wei, J.G., Qin, Q.W., Sun, H.Y., 2021a. Molecular characterization, expression and function analysis of *Epinephelus coioides* MKK4 response to SGIV and *vibrio alginolyticus* infection. Dev. Comp. Immunol. 119. https://doi.org/10.1016/j.dci.2021.104020.

Su, G.H., Logez, M., Xu, J., Tao, S.L., Villéger, S., Brosse, S., 2021b. Human impacts on global freshwater fish biodiversity. Science 371 (6531), 835-+. https://doi.org/10.1126/science.abd3369.

Sun, H.J., Wang, W.Q., Geng, L.L., Chen, Y.F., Yang, Z., 2013. In situ studies on growth, oxidative stress responses, and gene expression of juvenile bighead carp (*Hypophthalmichthys nobilis*) to eutrophic lake water dominated by cyanobacterial blooms. Chemosphere 93 (2), 421–427. https://doi.org/10.1016/j.chemosphere.2013.05.022.

Sun, C.F., Dong, J.J., Li, W.H., Tian, Y.Y., Hu, J., Ye, X., 2022. Response to four generations of selection for growth performance traits in mandarin fish (*Siniperca chuatsi*). Aquaculture 548. https://doi.org/10.1016/j.aquaculture.2021.737590.

Sun, X.Y., Bao, N., Rui, C., Xue, Y.D., Fang, Q., Zheng, T.Y., Lin, Z.Y., Liu, X.M., Wang, X. B., 2024. Identification of large yellow croakers (*Larimichthys crocea*) scavenger receptor genes: involvement in immune response to pseudomonas plecoglossicida infection and hypoxia-exposure experiments. Fish Shellfish Immunol. 144. https://doi.org/10.1016/j.fsi.2023.109307.

Terhorst, J., Kamm, J.A., Song, Y.S., 2017. Robust and scalable inference of population history froth hundreds of unphased whole genomes. Nat. Genet. 49 (2), 303–309. https://doi.org/10.1038/ng.3748.

Theunert, C., Tang, K., Lachmann, M., Hu, S.L., Stoneking, M., 2012. Inferring the history of population size change from genome-wide SNP data. Mol. Biol. Evol. 29 (12), 3653–3667. https://doi.org/10.1093/molbev/mss175.

Truitt, M.L., Conn, C.S., Shi, Z., Pang, X.M., Tokuyasu, T., Coady, A.M., Seo, Y., Barna, M., Ruggero, D., 2015. Differential requirements for eIF4E dose in Normal development and Cancer. Cell 162 (1), 59–71. https://doi.org/10.1016/j.cell.2015.05.049.

Valet, M., Siggia, E.D., Brivanlou, A.H., 2022. Mechanical regulation of early vertebrate embryogenesis. Nat. Rev. Mol. Cell Bio. 23 (3), 169–184. https://doi.org/10.1038/s41580-021-00424-z.

Wang, T.H., Huang, W.S., Costa, M.M., Secombes, C.J., 2011. The gamma-chain cytokine/receptor system in fish: more ligands and receptors. Fish Shellfish Immunol. 31 (5), 673–687. https://doi.org/10.1016/j.fsi.2011.05.016.

Wang, J., Hu, Z.B., Liao, X.L., Wang, Z.Y., Li, W., Zhang, P.P., Cheng, H., Wang, Q., Bhat, J.A., Wang, H., Liu, B.A., Zhang, H.Y., Huang, F., Yu, D.Y., 2022a. Whole-genome resequencing reveals signature of local adaptation and divergence in wild soybean. Evol. Appl. 15 (11), 1820–1833. https://doi.org/10.1111/eva.13480.

Wang, L., Sun, F., Lee, M., Yue, G.H., 2022b. Whole-genome resequencing infers genomic basis of giant phenotype in Siamese fighting fish (*Betta splendens*). Zool. Res. 43 (1), 78–80. https://doi.org/10.24272/j.issn.2095-8137.2021.266.

Wang, Y., Zhang, X., Wang, J., Wang, C., Xiong, F., Qian, Y.T., Meng, M.H., Zhou, M., Chen, W.J., Ding, Z.F., Yu, D., Liu, Y., Chang, Y.M., He, S.P., Yang, L.D., 2024. Genomic insights into the seawater adaptation in Cyprinidae. BMC Biol. 22, 87. https://doi.org/10.1186/s12915-024-01885-2.=11004.

Wei, X.M., Zhang, Y., Li, C., Ai, K.T., Li, K., Li, H.Y., Yang, J.L., 2020. The evolutionarily conserved MAPK/Erk signaling promotes ancestral T-cell immunity in fish via c-Myc-mediated glycolysis. J. Biol. Chem. 295 (10), 3000–3016. https://doi.org/10.1074/jbc.RA119.012231.

Wei, T., van Treuren, R., Liu, X.J., Zhang, Z.W., Chen, J.J., Liu, Y., Dong, S.S., Sun, P.N., Yang, T., Lan, T.M., Wang, X.G., Xiong, Z.Q., Liu, Y.Q., Wei, J.P., Lu, H.R., Han, S.P., Chen, J.C., Ni, X.M., Wang, J., Yang, H.M., Xu, X., Kuang, H.H., van Hintum, T., Liu, X., Liu, H., 2021. Whole-genome resequencing of 445 *Lactuca* accessions reveals the domestication history of cultivated lettuce. Nat. Genet. 53 (5), 752-+. https://doi.org/10.1038/s41588-021-00831-0.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., Yu, G., 2021a. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. The. Innovation 2 (3), 100141. https://doi.org/10.1016/j.xinn.2021.100141.

Wu, L.H., Liang, H.L., Hamunjo, C.M.K., Ge, X.P., Ji, K., Yu, H., Huang, D.Y., Xu, H., Ren, M.C., 2021b. Culture salinity alters dietary protein requirement, whole body composition and nutrients metabolism related genes expression in juvenile genetically improved farmed Tilapia (GIFT) (*Oreochromis niloticus*). Aquaculture 531. https://doi.org/10.1016/j.aquaculture.2020.735961.

Xiang, T., Dong, X.H., Ju, T., Shi, L., Grenouillet, G., 2023. Anthropogenic activities and environmental filtering have reshaped freshwater fish biodiversity patterns in China over the past 120 years. J. Environ. Manage. 344. https://doi.org/10.1016/j.jenvman.2023.118374.

Xie, H.X., Xing, K.F., Zhou, J., Zhao, Y., Zhang, J., Rong, J., 2025. Single-nucleotide polymorphisms and copy number variations drive adaptive evolution to freezing stress in a subtropical evergreen broadleaved tree: Hexaploid wild *Camellia oleifera*. Plant Diversity 47 (2), 214–228. https://doi.org/10.1016/j.pld.2024.07.009.

Xu, S.H., Huang, W., Wang, H.F., He, Y.G., Wang, Y., Wang, Y., Qian, J., Xiong, M.M., Jin, L., 2007. Dissecting linkage disequilibrium in african-american genomes: roles of markers and individuals. Mol. Biol. Evol. 24 (9), 2049–2058. https://doi.org/10.1093/molbev/msm135.

Yang, D.H., Liu, X.H., Xu, W.T., Gu, Z.Y., Yang, C.T., Zhang, L.Z., Tan, J.C., Zheng, X., Wang, Z., Quan, S., Zhang, Y.X., Liu, Q., 2019. The *Edwardsiella piscicida* thioredoxin-like protein inhibits ASK1-MAPKs signaling cascades to promote pathogenesis during infection. PLoS Pathog. 15 (7). https://doi.org/10.1371/journal.ppat.1007917.

Yang, J.P., Li, Y.F., Zhu, S.L., Li, X.H., Li, J., Chen, W.T., 2022. Mitochondrial diversity and population structure of grass carp (*Ctenopharyngodon idella*) in the Pearl River after anthropogenic release. Knowl. Manag. Aquat. Ec. 423. https://doi.org/10.1051/kmae/2022012.

Yu, S.W., Shen, G.X., Khor, T.O., Kim, J.H., Kong, A.N.T., 2008. Curcumin inhibits Akt/mammalian target of rapamycin signaling through protein phosphatase-dependent mechanism. Mol. Cancer Ther. 7 (9), 2609–2620. https://doi.org/10.1158/1535-7163.Mct-07-2400.

Yu, C., Zhang, P., Li, X.P., Sun, L., 2020. Japanese flounder *Paralichthys olivaceus* interleukin 21 induces inflammatory response and plays a vital role in the immune defense against bacterial pathogen. Fish Shellfish Immunol. 98, 364–373. https://doi.org/10.1016/j.fsi.2020.01.043.

Yuan, G.L., Zhao, W.H., Zhang, Y.W., Jia, Z., Chen, K.Y., Wang, J.Y., Feng, H., Zou, J., 2023. The biological functions and intestinal inflammation regulation of IL-21 in grass carp (*Ctenopharyngodon idella*) during infection with *Aeromonas hydrophila*. Cells 12 (18). https://doi.org/10.3390/cells12182276.

Zhang, D.W., Yang, Q., Xie, P., Deng, X.W., Chen, J., Dai, M., 2012. The role of cysteine conjugation in the detoxification of microcystin-LR in liver of bighead carp (*Aristichthys nobilis*): a field and laboratory study. Ecotoxicology 21 (1), 244–252. https://doi.org/10.1007/s10646-011-0783-1.

Zhang, Y., Lu, Y.F., Yindee, M., Li, K.Y., Kuo, H.Y., Ju, Y.T., Ye, S.H., Faruque, M.O., Li, Q., Wang, Y.C., Cuong, V.C., Pham, L.D., Bouahom, B., Yang, B.Z., Liang, X.W., Cai, Z.H., Vankan, D., Manatchaiworakul, W., Kowlim, N., Duangchantrasiri, S., Wajjwalku, W., Colenbrander, B., Zhang, Y., Beerli, P., Lenstra, J.A., Barker, J.S.F., 2016. Strong and stable geographic differentiation of swamp buffalo maternal and paternal lineages indicates domestication in the China/Indochina border region. Mol. Ecol. 25 (7), 1530–1550. https://doi.org/10.1111/mec.13518.

Zhang, C., Dong, S.S., Xu, J.Y., He, W.M., Yang, T.L., 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics 35 (10), 1786–1788. https://doi.org/10.1093/bioinformatics/bty875.

Zhang, A.Y., Jian, X.Y., Wang, D., Ren, J.Q., Wang, X.Y., Zhou, H., 2020a. Characterization and bioactivity of grass carp (*Ctenopharyngodon idella*) interleukin-21: inducible production and involvement in inflammatory regulation. Fish Shellfish Immunol. 99, 19–26. https://doi.org/10.1016/j.fsi.2020.01.059.

Zhang, X., Liu, Z., Wu, S., Sun, M., Wei, J., Qin, Q., 2020b. Fish RIP1 mediates innate antiviral immune responses induced by SGIV and RGNNV infection. Front. Immunol. 11, 1718. https://doi.org/10.3389/fimmu.2020.01718.

Zhang, Y., Shen, R.J., Gu, X.H., Li, K.Y., Chen, H.H., He, H., Mao, Z.G., Johnson, R.K., 2023. Simultaneous increases of filter-feeding fish and bivalves are key for controlling cyanobacterial blooms in a shallow eutrophic lake. Water Res. 245. https://doi.org/10.1016/j.watres.2023.120579.

Zhu, J.J., Thompson, C.B., 2019. Metabolic regulation of cell growth and proliferation. Nat. Rev. Mol. Cell Bio. 20 (7), 436–450. https://doi.org/10.1038/s41580-019-0123-5.