scientific reports

Check for updates

OPEN

Genetic diversity analysis of four *Parabramis pekinensis* populations based on microsatellite markers

Yuting Liu✉, Chunmei Yan✉, Haibo Li, Xiangqin Jin, Peng Liu, Weiqiang Chen, Huiji Liu, Ying Zhang & Jianan Liu

To investigate the germplasm differences between northern and southern populations of *Parabramis pekinensis* and analyze their genetic diversity and population structure, this study employed 10 developed microsatellite markers to analyze 133 individuals collected from the Yangtze River Basin (Jiangsu JS, Hunan HN) and the Amur River Basin (Jilin JL, Heilongjiang HLJ). The HLJ population is considered a wild stock, while the other three groups are classified as cultured populations. The analysis revealed that four *Parabramis pekinensis* populations exhibited high levels of genetic diversity, with polymorphic information content (PIC) ranging from 0.5130 to 0.6142. A total of 249 alleles were detected, among which the HN population showed the highest genetic diversity (*Na* = 8.0 ± 4.59). The analysis of molecular variance (AMOVA) revealed significant genetic differentiation among populations, with 34% of gene variation attributed to inter-population differences (Fst = 0.337). The UPGMA clustering and STRUCTURE analysis (optimal K = 3) grouped the JL and HLJ populations into a single cluster, which subsequently merged with the JS population to form a larger clade. In contrast, the HN population was distinctly separated as an independent branch. The study demonstrates that the HN population of *P. pekinensis* should be prioritized for conservation as a distinct genetic unit, with the enhancement of habitat connectivity among northern populations could effectively mitigate genetic diversity loss. These findings provide a scientific foundation for the preservation of *P. pekinensis* germplasm resources.

Keywords *Parabramis pekinensis*, Genetic diversity, Genetic structure

Parabramis pekinensis is classified within the order Cypriniformes, family Cyprinidae, subfamily Cultrinae, and genus Parabramis. This species is distributed extensively, with a primary habitat in rivers and lakes within China's major river systems, including the Amur River basin, the Yangtze, and the Pearl Rivers. Additional occurrences have been reported in North Korea and Russia. *P. pekinensis* is highly regarded for its tender and delicate flesh, exquisite flavour, and high nutritional value, thus classifying it as a premium freshwater fish with significant market potential. The genetic diversity of *P. pekinensis* has been subject to a substantial decline as a consequence of habitat fragmentation. As a pelagic-egg spawning species, its reproductive capacity is significantly diminished by the development of hydraulic infrastructure, thereby posing a substantial threat to the viability of its population. Furthermore, practices such as lake reclamation for farmland, overfishing, unsustainable artificial domestication, and unscientific management have collectively led to a drastic decline in the germplasm resources of *P. pekinensis*, causing severe erosion of its genetic diversity.

In recent years, research on *P. pekinensis* has focused on the mechanisms of gonadal development[1], external morphological characteristics[2,3], dynamics of trophic niches[4], and structure of gut microbial communities[5]. However, systematic investigations into its germplasm resources remain relatively insufficient, particularly regarding the evaluation of genetic diversity and characterization of the genetic structure among different regional populations of *P. pekinensis*, for which there is currently a paucity of relevant reports.

Microsatellite markers are defined by three characteristics: codominant inheritance, high mutation rates and locus-specific variations. The capable to discern subtle population differentiation with a high degree of precision has led to the extensive utilisation of these methods in the evaluation of genetic structure and diversity in a multitude of fish species[6–10]. In recent years, microsatellite markers have been successfully developed and used to evaluate the genetic diversity of various cyprinid species[11–14]. However, there are currently no published microsatellite markers specifically available for the *P. pekinensis*.

Jilin Fisheries Research Institute, Changchun 130033, China. ✉email: 943071009@qq.com; 410977247@qq.com

Materials and methods

Ethic statement

All sampled *Parabramis pekinensis* were neither endangered nor protected species. According to Chinese regulations, no special permits are required for catching *Parabramis pekinensis* in natural waters. This study was approved by the Institutional Animal Care and Use Committee of Jilin Fisheries Research Institute (Approval No. JFRIAEC-2023-015), and all procedures were conducted in strict compliance with the ARRIVE guidelines. All experiments were performed in accordance with national laws and institutional guidelines for animal care and use established by Jilin Fisheries Research Institute.

Collection of experimental materials

A total of 133 *P. pekinensis* individuals were collected from wild samples in Fuyuan County, Heilongjiang Province (48°21′18.062″ N, 134°16′32.007″ E) (HLJ, n = 14); cultured samples in Jilin City, Jilin Province (43°57′2.200″ N,126°41′23.087″ E) (HLJ, n = 14) (JL, n = 52); cultured samples in Jingjiang County, Jiangsu Province (31°55′1.878″ N, 120°8′47.480″ E) (JS, n = 22), and cultured samples in the Hunan Academy of Aquatic Sciences (28° 16′ 14.833″ N, 113°6′19.459″ E) (HN, n = 45). The average overall length was 9.78 cm. The average body length was 7.92 cm, the average height was 2.55 cm and the average body mass was 8.07 g. The HLJ and JL populations were from the Heilongjiang River system and the HN and JS populations were from the Yangtze River system. The collected fin tissues were preserved by immersion in 95% ethanol. The DNA was extracted using the QIAGEN DNeasy Blood & Tissue Kit according to the instructions. Its integrity was verified by 1% agarose gel electrophoresis, and its concentration and purity were detected using the NanoDrop 2000 Micro UV–Vis Spectrophotometer. The DNA was diluted to 50 ng/μL, and then stored at − 20 °C for storage.

Microsatellite primer design and screening

Using the publicly available reference genome GCA_018812025.1 of *Megalobrama amblycephala* from the NCBI database. SSR Hunter was used to search for microsatellite regions. Sequences with more than eight trinucleotide repeats were selected as candidate sequences. Primers were designed using Primer Premier 5.0. The primers were synthesized by Beijing Le8 Technology Co. The PCR reaction system contained a total volume of 25 μL, including 2 μL (50 ng) of DNA template, 0.5 μL (20 μM) of primers, 0.5 μL of dNTPs, 2.5 μL of 10 × PCR buffer, 0.5 μL of Taq DNA polymerase (Takara) and 18.5 μL of ddH_2O. Amplification was performed using a GeneAmp PCR System 9600 manufactured by PerkinElmer, USA. The PCR reaction conditions were as follows: pre-denaturation at 95 °C for 5 min; 35 cycles, each consisting of 30 s at 95 °C, 30 s at 56 °C and 30 s at 72 °C and a final extension at 72 °C for 10 min. Forty-eight pairs of microsatellite primers were obtained.

Three samples were selected from each population to screen the polymorphic primers. PCR reaction system was 15 μL, including 0.8 μL of TP-M13 (5 mM), 1 μL of primers (5 mM), 7.5 μL of PCR Mix, 1.5 μL of DNA template, and ddH2O was added to 15 μL. Amplification was carried out by TP-M13-SSR fluorescence labeling PCR. The PCR reaction program was as follows: pre-denaturation at 94 °C for 2 min; 5 cycles, (each cycle denaturation at 94 °C for 30 s, annealing temperature starting from 60℃, lowering 1 °C to 55 °C for each cycle, extension at 72 °C for 30 s); 30 standard cycles (30 s at 94 °C, 30 s at 55 °C, 30 s at 72 °C); 10 labeling cycles (30 s at 94 °C, 30 s at 53 °C, 30 s at 72 °C), and final extension at 60 °C for 30 min to ensure product integrity. After the PCR reaction, the amplification products were detected by agarose gel electrophoresis, and those with the target bands were selected for detection on the 3730xl Genetic Analyzer. The results of electrophoresis detection were analyzed with Genemapper 4.0 and to Size results, and then the results of each pair of primers were analyzed, if the number of alleles was greater than or equal to three, it indicated that the primer was polymorphic, and it could be used as molecular markers for genetic diversity analysis at a later stage.

Finally, 10 pairs of microsatellite primers with good polymorphism were selected and detailed sequence information is listed in Table 1(Accession number: BankIt2968220 CCB-2 PV765790, BankIt2968220 CCB-3 PV765791, BankIt2968220 CCB-5 PV765792, BankIt2968220 CCB-6 PV765793, BankIt2969170 CCB-7 PV771622, BankIt2969170 CCB-10 PV771623, BankIt2969170 CCB-11 PV771624, BankIt2969170 CCB-13 PV771625, BankIt2969170 CCB-14 PV771626 and BankIt2969170 CCB-15 PV771627). The collected samples were subjected to PCR amplification, the PCR products were subjected to agarose gel electrophoresis, and the genotypes of each SSR locus of each sample were detected using a 3730xl Genetic Analyzer (Applied Biosystems, USA).

Data analysis

We used POPGENE 3.2[15] to analyse the genotypes of different loci across populations to calculate the following parameters for each microsatellite locus: the number of alleles (*Na*), the effective number of alleles (*Ne*), the observed heterozygosity (*Ho*), the expected heterozygosity (*He*), Shannon's information index (*I*), Nei's standard genetic distance, the inbreeding coefficient (*Fis*), the fixation index (*Fst*), the gene flow (*Nm*) and the Hardy–Weinberg equilibrium test (*p-value*), We also constructed phylogenetic trees among populations based on Nei's genetic distance. Polymorphic information content (*PIC*) was calculated using Cervus 3.0.7[16]. Molecular variance analysis (AMOVA) was conducted using Arlequin version 3.5.2.2[17]. The GenAlEx 6.5 Excel add-in[18] was employed for chi-square tests of Hardy–Weinberg equilibrium, as well as principal coordinates analysis (PCoA) based on pairwise genetic distances among populations and individuals.

Population genetic structure was analyzed using Structure Selector[19], which determined the optimal K value based on two criteria: Mean lnP(K)[20] and ΔK[21]. The integrated CLUMPAK program was then used to generate the final graphical results.

Locus	Peimer sequences (5′-3′)	Size(bp)	Repeat motif	*Na*	*Ne*	*Ho*	*He*	*I*	*PIC*
CCB-2	F: AGAGCTGCTCGTGATGATTAA R: GAGGAACGCTTGAGTTTGG	352	(CAT)10	4.5000	2.5269	0.4588	0.5850	1.0639	0.5187
CCB-3	F: AAAGGACTGAGGCTTCAAACA R: TCTCCGTCTTCCTCTGGTTC	323	(TGGA)7	5.0000	2.4189	0.5092	0.5930	1.0930	0.5446
CCB-5	F: AGGTTCGTTCAGAGGCAGTG R: CGTGTCGAGCGTTTCCAG	217	(TGA)8	4.0000	2.3773	0.6172	0.5861	1.0349	0.5180
CCB-6	F: TAATTGTATCCTGTAAAGCGTCCTT R: TTGGGTTCATCATCTCACCTTC	336	(TAT)10	5.5000	2.7994	0.6243	0.6535	1.2633	0.6091
CCB-7	F: CGGAAAGAACGTCATCAGC R: AAATGTTACCCATCCCAAACT	196	(AAT)7	3.0000	1.5035	0.3534	0.3117	0.5540	0.2801
CCB-10	F: CAGCTTCCTTCGGTTCTGC R: GTCCAAATCCCATCATTAGTGTAT	253	(TGT)8	4.7500	2.8353	0.5980	0.6469	1.2031	0.5722
CCB-11	F: TTGCCATGTTATGATATGTTGC R: GGGTCGGTGGATGGATAGA	164	(TCTA)16	15.2500	9.7019	0.7721	0.9070	2.4210	0.8794
CCB-13	F: AAGGGCAGGCAACAAACA R: CAGGGAAGAAGATGAAGAGGTG	282	(TCG)5 (TCC)9	4.5000	1.6997	0.3555	0.3557	0.7153	0.4346
CCB-14	F: ATGCTTCCCAGCCTGTACTTT R: CTGGTGCTACGCCTGATGT	248	(TCA)8	4.0000	1.7432	0.3646	0.4270	0.8005	0.3744
CCB-15	F: ACAGACCTGCTTTACATCATCC R: TTCCCATTTCGGCCATTT	339	(ATCT)12	11.7500	7.2112	0.7986	0.8580	2.1093	0.8000
Mean				6.2250	3.4817	0.5412	0.5924	1.2258	0.5531

Table 1. Information on 10 pairs of microsatellite primers and genetic diversity parameters of 133 *Parabramis pekinensis*. Na, Number of alleles; Ne, Number of effective alleles; Ho, Observed heterozygosity; He, Expected heterozygosity; I, Shannon's information index; PIC, Polymorphism information content.

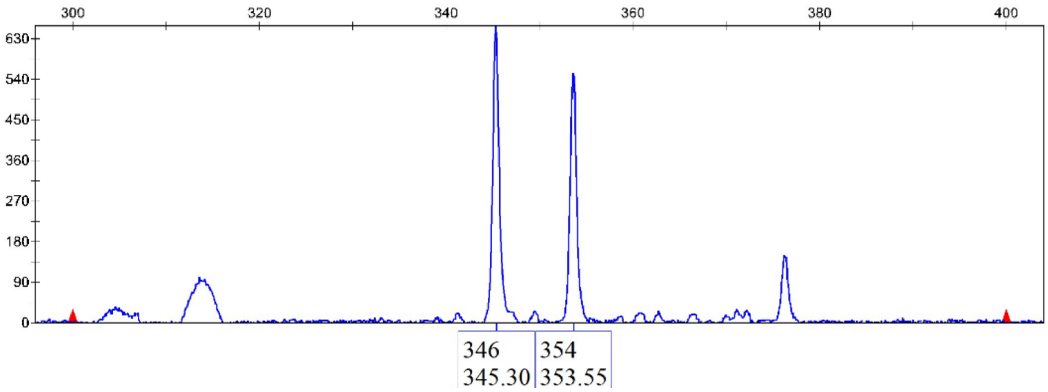

Fig. 1. Allelic peak profiles of microsatellite loci in representative *P. pekinensis* individuals (allele sizes in base pairs, bp).

Results and analysis

SSR genotyping results

Genomic DNA was extracted from 133 *P. pekinensis* tissue samples. This was followed by PCR amplification using 10 microsatellite primer pairs. The amplification products were separated by capillary electrophoresis, and genotypes were analyzed using GeneMapper software. Figure 1 displays the allele peak profiles of selected loci in representative individuals.

Genetic diversity

We evaluated the genetic diversity of four *P. pekinensis* populations using 10 microsatellite primer pairs. We analyzed the number of alleles (*Na*), the effective number of alleles (*Ne*), the observed heterozygosity (*Ho*), the expected heterozygosity (*He*), the Shannon's index (*I*), Nei's standard genetic distance, and polymorphic information content (*PIC*) (Table 2). A total of 249 alleles (*Na*) were detected, with an average of 6.225 per locus. The *He* values ranged from 0.0714 to 0.9656, and *I* values from 0.1541 to 2.7502. Nei's genetic distance ranged from 0.0689 to 0.9311. The average *I* values for all populations exceeded 1, indicating high genetic diversity. PIC values calculated using specialized software showed a maximum of 0.9268, a minimum of 0.0665, and an average of 0.5554 ($PIC > 0.5$). Notably, two microsatellite loci (*CCB-7*, *CCB-13*) exhibited *PIC* values below 0.3, whereas the remaining loci showed *PIC* values above 0.3, demonstrating that most markers were highly polymorphic.

Hardy–Weinberg equilibrium (*HWE*) tests were performed on the 10 microsatellite loci across the four populations. Among these, 27 loci conformed to *HWE* ($p > 0.05$), while 13 loci significantly deviated ($p < 0.05$).

locus	Group	*Na*	*Ne*	*Ho*	*He*	*Neil*	*I*	*PIC*	*H-W*
CCB-2	JS	3	1.6026	0.4545	0.3848	0.3760	0.6889	0.3436	0.640387
	HLJ	4	2.4968	0.5000	0.6217	0.5995	1.0762	0.5322	0.095086
	JL	4	2.5642	0.4808	0.6159	0.6100	1.0897	0.5397	0.000000
	HN	7	3.4439	0.4000	0.7176	0.7096	1.4017	0.6593	0.000463
CCB-3	JS	5	2.6814	0.5000	0.6416	0.6271	1.1983	0.5703	0.041650
	HLJ	5	2.6133	0.5714	0.6402	0.6173	1.1696	0.5601	0.321467
	JL	5	1.9860	0.3654	0.5013	0.4965	0.9353	0.5240	0.761118
	HN	5	2.3950	0.6000	0.5890	0.5825	1.0689	0.5241	0.074940
CCB-5	JS	3	2.1851	0.5909	0.5550	0.5424	0.8465	0.4367	0.514696
	HLJ	4	2.6849	0.7143	0.6508	0.6276	1.1596	0.5753	0.312993
	JL	4	2.2356	0.5192	0.5581	0.5527	1.0031	0.5182	0.909737
	HN	5	2.4036	0.6444	0.5905	0.5840	1.1306	0.5419	0.677564
CCB-6	JS	4	2.4506	0.5909	0.6057	0.5919	1.0180	0.5102	0.914532
	HLJ	5	3.1870	0.7143	0.7116	0.6862	1.3239	0.6385	0.956695
	JL	5	2.9504	0.4808	0.6675	0.6611	1.2415	0.6748	0.004183
	HN	8	3.6096	0.7111	0.7311	0.7230	1.4709	0.6127	0.041631
CCB-7	JS	3	1.4556	0.3636	0.3203	0.3130	0.5944	0.2894	0.817377
	HLJ	2	1.0740	0.0714	0.0714	0.0689	0.1541	0.0665	1.000000
	JL	3	1.6119	0.4231	0.3833	0.3796	0.6031	0.3424	0.830701
	HN	4	1.8724	0.5556	0.4712	0.4659	0.8642	0.4221	0.552825
CCB-10	JS	5	2.4757	0.6364	0.6099	0.5961	1.1720	0.5561	0.871692
	HLJ	5	2.5621	0.5000	0.6323	0.6097	1.1967	0.5662	0.675124
	JL	3	2.6839	0.5000	0.6335	0.6274	1.0330	0.4910	0.032024
	HN	6	3.6193	0.7556	0.7318	0.7327	1.4105	0.6756	0.010656
CCB-11	JS	10	6.9640	0.8182	0.8763	0.8564	2.0704	0.8402	0.099290
	HLJ	17	14.5158	0.7143	0.9656	0.9311	2.7502	0.9268	0.000844
	JL	15	7.9529	0.7115	0.8827	0.8743	2.3524	0.8641	0.035092
	HN	19	9.3750	0.8444	0.9034	0.8933	2.5111	0.8847	0.102786
CCB-13	JS	3	1.2040	0.0000	0.1734	0.1694	0.3676	0.5582	0.000280
	HLJ	5	2.4198	0.6429	0.6085	0.5867	1.1786	0.5524	0.782182
	JL	3	1.1450	0.1346	0.1279	0.1267	0.2742	0.1638	0.972602
	HN	7	2.0301	0.6444	0.5131	0.5074	1.0408	0.4642	0.983644
CCB-14	JS	3	1.6491	0.3182	0.4027	0.3936	0.6809	0.3440	0.429024
	HLJ	3	1.5556	0.2857	0.3704	0.3571	0.6560	0.3254	0.301902
	JL	4	1.7060	0.3654	0.4128	0.4135	0.7594	0.3372	0.373461
	HN	6	2.0621	0.4889	0.5208	0.5151	1.1056	0.4911	0.630885
CCB-15	JS	6	3.6667	0.6818	0.7442	0.7273	1.4533	0.6810	0.154177
	HLJ	12	9.8000	0.7857	0.9312	0.8980	2.3759	0.8892	0.000145
	JL	16	7.1629	0.8846	0.8687	0.8604	2.3202	0.8486	0.000026
	HN	13	8.2150	0.8444	0.8881	0.8783	2.2867	0.8665	0.000555
Mean	JS	4.5000	2.6335	0.4955	0.5314	0.5193	1.0090	0.5130	–
	HLJ	6.2000	4.2912	0.5500	0.6204	0.5982	1.3041	0.5638	–
	JL	6.2000	3.1999	0.4865	0.5657	0.5602	1.1612	0.5304	–
	HN	8.0000	3.9026	0.6489	0.6657	0.6583	1.4291	0.6142	–

Table 2. Parameters of genetic diversity analyzed by ten loci in populations of *P. pekinensis*.

Genetic structure

The analysis of genetic differentiation coefficients (Fst) and gene flow (Nm) between pairwise populations of four *Parabramis pekinensis* populations (Table 3) revealed that Nm ranged from 1.3428 to 3.3530. Nm was greater than 1 between populations, with the highest gene flow observed between the HLJ and JL populations (3.3530), and the lowest between the HN and JS populations (1.3428). The genetic differentiation coefficients (Fst) were less than 0.15 for other population pairs except for the HN-JS and HN-JL population pairs.

The genetic similarity coefficients and genetic distances between the four *P. pekinensis* populations ranged from 0.4194 to 0.8336 and 0.1819 to 0.9567, respectively (Table 4). The HLJ and JL populations exhibited the highest genetic similarity (0.8336) and the smallest genetic distance (0.1819), while the HN and JS populations showed the lowest genetic similarity (0.3841) and the largest genetic distance (0.9567). Analysis of Molecular Variance (AMOVA) revealed that 34% of the genetic variation originated from differences between populations,

Populations	JS	HLJ	JL	HN
JS		1.9566	3.1982	1.1467
HLJ	0.1133		3.3530	1.8074
JL	0.0725	0.0694		1.3428
HN	0.1790	0.1215	0.1570	

Tab.3. Fst (diagonal below) and gene flow (diagonal above) in four populations of *P. pekinensis*.

Populations	JS	HLJ	JL	HN
JS		0.6756	0.7975	0.3841
HLJ	0.3921		0.8336	0.5285
JL	0.2263	0.1819		0.4194
HN	0.9567	0.6378	0.8689	

Table 4. Nei's genetic identity (diagonal above) and genetic distance (diagonal below) in four populations of *P. pekinensis*.

Source of variance	Degree of freedom	Sum of square of mean deviation	Estimated Variance	Percentage of total variance	Fixation index	*P*-value
Among populations	3	123.923	1.292	34	0.337	0.001
Within populations	129	328.227	2.544	66		
Total	132	455.150	3.836	100		

Table 5. The results of analysis of molecular variance of *Parabramis pekinensis* populations.

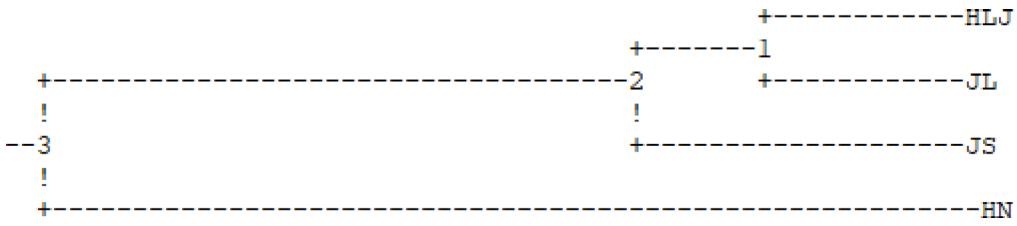

Fig. 2. UPGMA phylogenetic tree of four *P. pekinensis* populations constructed based on Nei's genetic distance.

while 66% stemmed from variation within populations (Table 5). These results suggest significant genetic differentiation between the four *P. pekinensis* populations, to varying extents.

The phylogenetic tree constructed using UPGMA method and based on Nei's genetic distance (Fig. 2) showed that the four *P. pekinensis* populations share a common ancestral origin and clustered into two major groups. Specifically, the JL and HLJ populations first grouped together, then clustered with the JS population, while the HN population formed a distinct group. The results of Principal Coordinates Analysis (PCoA) are shown in Fig. 3, no significant genetic clustering was observed.

Genetic clustering analysis

The optimal genetic clustering number (K = 3) was determined by Structure Harvester analysis, revealing that the four *P. pekinensis* populations were derived from three distinct ancestral gene pools (Fig. 4). Additionally, based on Q-value statistics of the population probability distribution of the 133 samples, shows that these populations can be categorized into three taxa: the Jiangsu population, the Heilongjiang and Jilin populations, and the Hunan population (Fig. 5).

Discussion

Genetic diversity analysis

Genetic diversity is defined as the degree of genetic variation among individuals or populations within a species. It is an important resource for species to cope with environmental change and human interference. Higher genetic diversity has been demonstrated to provide populations with richer adaptive variation, thus enhancing their survival and evolutionary potential in dynamic environments[22–24]. In this study, we obtained discrete

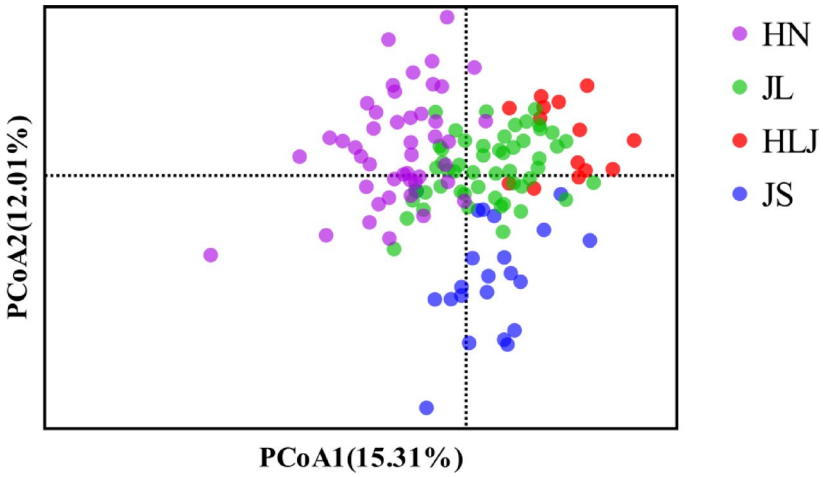

Fig. 3. Principal coordinate analysis (PCoA) plot of *P. pekinensis* individuals based on genetic distance.

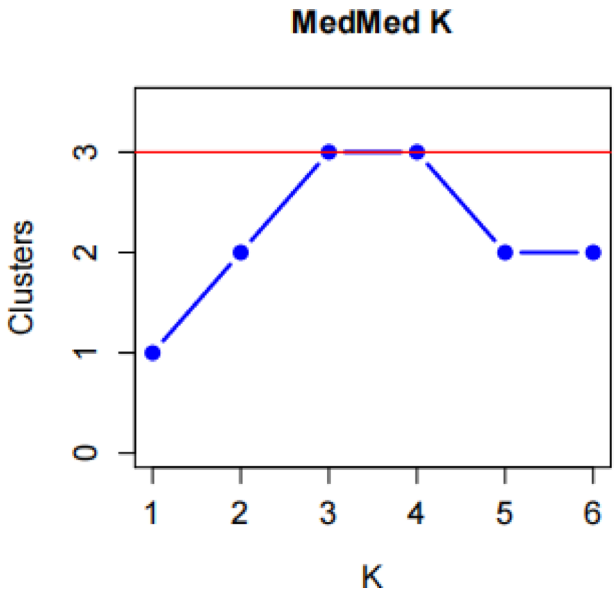

Fig. 4. Relationship between the rational cluster K and estimated value ΔK.

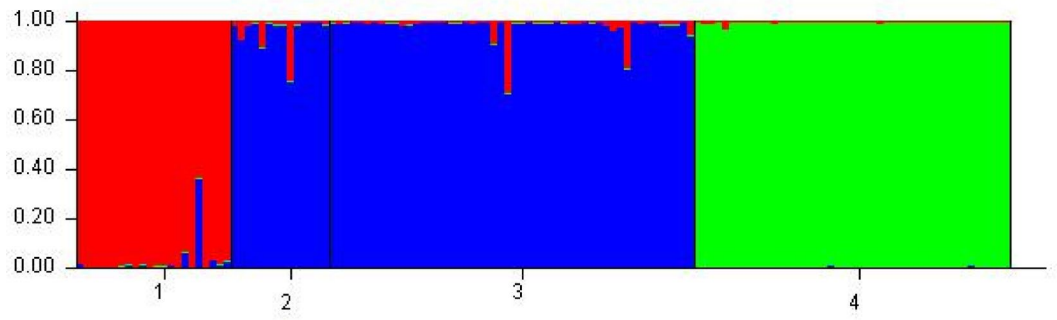

Fig. 5. Genetic structure analysis of individuals from four *P. pekinensis* populations under K = 3 hypothesis. Each vertical bar represents an individual, with different colors indicating the proportional contribution of each K genetic cluster to the individual's genotype: 1. JS population, 2. HLJ population, 3. JL population, and 4. HN population.

allelic data for 133 *Parabramis pekinensis* accessions from 4 populations based on analysis using 10 microsatellite markers. The average polymorphic information content (*PIC*) for each population ranged from 0.5130 to 0.6142. This was higher than that reported for other carp species in the Yangtze River basin, such as the bighead carp (*Hypophthalmichthys nobilis*)[25,26] and grass carp (*Ctenopharyngodon idella*)[27]. This suggests that the *P. pekinensis* population may have greater adaptive potential, which is highly valuable for breeding and resource conservation purposes. We evaluated the genetic diversity of *P. pekinensis* in terms of allelic richness (*Na*), heterozygosity (*He*), and allelic diversity (*I*). In the present study, the mean genetic parameters detected at each microsatellite locus ($Na = 6.2250$, $He = 0.5924$, $I = 1.2258$) indicated that the genetic diversity level of *P. pekinensis* was slightly lower than that of blunt snout bream (*Megalobrama amblycephala*) ($Na = 10.000$, $He = 0.734$, $I = 1.755$)[28], a closely related species from the genus Megalobrama. Furthermore, compared to other aquaculture fish species, the genetic diversity of *P. pekinensis* was higher than that of largemouth bass (*Micropterus salmoides*) ($Na = 6.571$, $He = 0.484$, $I = 0.988$)[29] but lower than that of spotted sea bass (*Lateolabrax maculatus*) ($Na = 28.7273$, $He = 0.8248$, $I = 2.3202$)[30].

In our study, significant genetic diversity differences were detected among the four geographically distinct populations of *P. pekinensis*. In biostatistics and sampling surveys, a sample size of 30 is conventionally used as the threshold to distinguish between large and small samples. Although the sample sizes varied across populations in this study, the HLJ and JS populations had fewer than 30 individuals. Nevertheless, these samples were sufficient to broadly reflect the genetic resource status of their respective regional groups. Due to the influence of sample size on the number of alleles (*Na*), expected heterozygosity (*He*) is more commonly used to measure genetic diversity within populations[31]. When *He* values range between 0.500 and 0.800, the population can be considered to possess relatively high genetic diversity[32]. In conclusion, our findings demonstrate that the HN population displayed the highest level of genetic variability ($Na = 8.0000$, $He = 0.6657$, $I = 1.4291$), followed sequentially by the HLJ and JL populations, with the JS population exhibiting the lowest diversity. The complex hydrological conditions characteristic of the middle and lower Yangtze River basin likely generate heterogeneous selective pressures that maintain multiple adaptive alleles through balancing selection[33,34], potentially accounting for the elevated genetic diversity observed in the HN population. The HLJ population ($Na = 6.2000$, $He = 0.6204$, $I = 1.3041$) and JL population ($Na = 6.2000$, $He = 0.5657$, $I = 1.1612$) exhibited similarly high levels of genetic diversity. This genetic pattern may be attributed to the relatively intact aquatic ecosystem and lower intensity of anthropogenic disturbances characteristic of the Heilongjiang River basin. In contrast, the JS population exhibited significantly lower genetic diversity parameters ($Na = 4.5000$, $He = 0.5314$, $I = 1.0090$). As a major aquaculture region[35], the cultivation of *P. pekinensis* in Jiangsu Province may have long relied on a limited number of high-quality broodstocks for artificial propagation. This intensive artificial selection pressure likely contributed to the reduced genetic diversity observed in the Jiangsu population, a phenomenon consistent with the findings of Zhang et al.[36] regarding the congeneric *M. amblycephala*. Genomic analyses further revealed that farmed populations of *M. amblycephala* in Jiangsu experienced significant declines in key genetic diversity indices, including allele number *(Na)* and heterozygosity (*He*), following seven generations of artificial selection.

It is particularly important to emphasize that the neutral theory of molecular evolution provides an incomplete explanation for genetic diversity patterns. This theory originated from the flawed "molecular clock hypothesis", which fundamentally misinterpreted the phenomenon of genetic equidistance. Recent empirical studies have conclusively refuted the neutral theory[37]. A recent seminal study has demonstrated that the human mitochondrial DNA genome is predominantly functional rather than neutral, providing direct empirical evidence that falsifies the neutral theory of molecular evolution[38]. While short tandem repeats (STRs) were originally postulated to be evolutionarily neutral, emerging evidence demonstrates their functional capacity to bind transcription factors[25]. This paradigm shift has contributed to the development of novel theoretical frameworks, most notably the Maximum Genetic Diversity Theory, which is progressively superseding the neutral theory of molecular evolution[39].

The Hardy–Weinberg equilibrium conditions are: infinitely large groups; random mating; there are no mutations; there is no choice; there is no migration; there is no genetic drift (random fluctuations in gene frequencies within a small population)[40]. In practical applications, populations rarely meet all ideal assumptions. The deviation of genotype distribution from Hardy–Weinberg equilibrium is typically measured using P-values, where lower *P*-values indicate greater deviations. In this study, there may be insufficient sample size, wrong typing, or biased sample selection[41], and only 27 of the 40 polymorphic loci met the H-W equilibrium.

Genetic structure analysis

Maintaining high genetic diversity within biological populations plays a crucial role in their environmental adaptability. Specifically, parameters such as the fixation index (FST) can effectively reflect the level of genetic differentiation among populations, and these differences are often closely associated with ecological adaptation processes[42,43]. The molecular variance analysis (AMOVA) in this study revealed that 34% of the genetic variation occurred among populations (fixation index $FST = 0.337$, $P < 0.001$), while the remaining 66% of variation was distributed within populations (Table 5). This differentiation pattern aligns well with the ecological characteristics of habitat fragmentation in freshwater fishes, likely resulting from dispersal limitations and hydrological connectivity constraints[44]. Current research demonstrates significant variation in FST values (0.0725–0.1790) among the four geographically distinct *P. pekinensis* populations, indicating varying degrees of genetic differentiation between these groups. The HN population exhibited relatively high pairwise FST values (0.1570–0.1790) compared to other populations, which may be attributed to its geographical location. The Yangtze River likely serves as a natural barrier that restricts gene flow between the southern HN population and northern populations (JS, JL, HLJ). A similar phenomenon was reported by Wang et al.[30] in Large yellow croaker (*Larimichthys crocea*) populations from the South China Sea, where the Qingdao (QD) population, though geographically located in northern waters, genetically clustered with southern populations due to restricted gene

flow caused by natural barriers formed by the Yellow Sea and East China Sea. The JS, JL, and HLJ populations exhibited low genetic differentiation but strong genetic admixture, as evidenced by their interpopulation FST values (0.0694–0.1133), gene flow estimates (1.9566–3.3530), and Nei's genetic similarity coefficients (0.6756–0.8336). Notably, the JL and HLJ populations showed the highest Nei's genetic similarity (0.8336) and smallest genetic distance (0.1819), likely attributable to their hydrological connectivity—the JL population inhabits Songhua Lake, a tributary of the Heilongjiang River system, which probably facilitates gene exchange between these populations.

The UPGMA phylogenetic tree based on Nei's genetic distance (Fig. 1) clustered all *P. pekinensis* specimens into three major clades: one comprising the northeastern populations (JL and HLJ), one consisting of the JS population, and one containing the HN population. This clustering pattern is consistent with the population structure analysis (K = 3) (Figs. 4, 5). In conclusion, we recommend designating the HN population as a distinct Evolutionarily Significant Unit (ESU) for prioritized conservation, while simultaneously enhancing ecological connectivity among northern populations' habitats to maintain their genetic diversity levels. However, significant discrepancies were observed among the phylogenetic tree, population structure analysis, and PCA results. These inconsistencies may stem from the limited genetic differentiation and close genetic distances among individuals. Furthermore, methodological differences between phylogenetic reconstruction and principal component analysis could contribute to this divergence[45,46]. The underlying causes of these discordant patterns warrant further investigation.

Conclusion

This study systematically analyzed the genetic diversity and population structure of four geographic populations (JS, HN, JL, HLJ) of *P. pekinensis* using 10 microsatellite markers. The results demonstrated that *P. pekinensis* populations collectively maintained relatively high genetic diversity (mean PIC = 0.5130–0.6142). Specifically, the HN population exhibited the highest genetic variation ($Na = 8.0000$, $He = 0.6657$, $I = 1.4291$), followed by HLJ and JL populations, while the JS population showed the lowest diversity levels. Both population structure analysis (K = 3) and UPGMA phylogenetic tree consistently identified three distinct clusters: northeastern populations (JL and HLJ) forming one branch, JS population as a separate branch, and HN population as an independently diverged lineage. The study revealed close genetic relationships between the JS population and northeastern populations (JL/HLJ), indicating substantial historical gene flow among these groups. The study recommends designating the HN population as a distinct Evolutionarily Significant Unit (ESU) for prioritized conservation due to its unique genetic divergence, while simultaneously enhancing habitat connectivity among northern populations (JS, JL, HLJ) to maintain their genetic diversity. These findings provide crucial scientific foundations for the conservation and management of *P. pekinensis* germplasm resources.

Data availability

The datasets used and analysed during the current study available from the corresponding author on reasonable request.

Received: 23 April 2025; Accepted: 4 July 2025
Published online: 23 July 2025

References

1. Gu, H., Feng, Y. & Yang, Z. Differential study of the *Parabramis pekinensis* intestinal microbiota according to different gonad development stages. *Fish. Sci.* **88**, 721–731 (2022).
2. Xu, W. et al. Effects of morphological traits on body weight and analysis of growth-related genes of *Parabramis pekinensis* at different ages. *BMC Zool.* **8**, 13. https://doi.org/10.1186/s40850-023-00174-9 (2023).
3. Tian, X., Gao, P., Xu, Y., Xia, W. & Jiang, Q. Reduction of biogenic amines accumulation with improved flavor of low-salt fermented bream (*Parabramis pekinensis*) by two-stage fermentation with different temperature. *Food Biosci.* **44**, 101438 (2021).
4. Luo, S. et al. Formation of adaptive trophic niches of Euryphagous fish species in response to off-seasonal water level regulation in Hongze lake. *Anim. Basel* **15**, 59. https://doi.org/10.3390/ani15010059 (2024).
5. Gu, H., Feng, Y., Zhang, Y., Yin, D. & Tang, W. Differential study of the *Parabramis pekinensis* intestinal microbiota according to different habitats and different parts of the intestine. *Ann. Microbiol.* **71**, 1–11 (2021).
6. Abbas, K., Zhou, X., Li, Y., Gao, Z. & Wang, W. Microsatellite diversity and population genetic structure of yellowcheek, *Elopichthys bambusa* (Cyprinidae) in the Yangtze River. *Biochem. Syst. Ecol.* **38**, 806–812 (2010).
7. Zhu, W. et al. Genetic diversity and population structure of bighead carp (*Hypophthalmichthys nobilis*) from the middle and lower reaches of the Yangtze River revealed using microsatellite markers. *Aquac. Rep.* **27**, 101377 (2022).
8. Zhong, L. et al. Genetic diversity and population structure of yellow catfish *Pelteobagrus fulvidraco* from five lakes in the middle and lower reaches of the Yangtze River, China, based on mitochondrial DNA control region. *Mitochondrial Dna J. Dna Mapp. Seq. Anal.* **24**, 552–558 (2013).
9. Nikolic, N., Fève, K., Chevalet, C., Høyheim, B. & Riquet, J. A set of 37 microsatellite DNA markers for genetic diversity and structure analysis of Atlantic salmon *Salmo salar* populations. *J. Fish Biol.* **74**, 458–466 (2009).
10. Pacheco-Almanzar, E., Ramírez-Saad, H., Velázquez-Aragón, J. A., Serrato, A. & Ibáez, A. L. Diversity and genetic structure of white mullet populations in the Gulf of Mexico analyzed by microsatellite markers. *Estuar. Coast. Shelf Sci.* **198**, 249–256 (2017).
11. Kim, K. R., Kwak, Y. H., Sung, M. S., Cho, S. J. & Bang, I. C. Population structure and genetic diversity of the endangered fish black shinner *Pseudopungtungia nigra* (Cyprinidae) in Korea: A wild and restoration population. *Sci. Rep.* **13**, 9692 (2023).
12. Du, R., Zhang, D., Wang, Y., Wang, W. & Gao, Z. Cross-species amplification of microsatellites in genera *Megalobrama* and *Parabramis*. *J. Genet.* **92**, e106-109 (2013).
13. Qianqian, Z., Jie, C., Xiayun, J. & Shuming, Z. Establishment of DNA fingerprinting and analysis on genetic structure of different *Parabramis* and *Megalobrama* populations with microsatellite. *J. Fish. China* **38**, 15–22 (2014).
14. Lin, L. I. et al. Establishment of microsatellite identification on there Culterinae fishes in Pearl River. *Guangdong Agric. Sci.* **41**, 102–101 (2014).
15. Yeh, F. C. & Boyle, T. J. B. Population genetic analysis of codominant and dominant markers and quantitative traits. *Belg. J. Bot.* **129**, 157 (1997).

16. Kalinowski, S. T. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099–1106 (2010).
17. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x (2010).
18. Smouse, P. E., Whitehead, M. R. & Peakall, R. An informational diversity framework, illustrated with sexually deceptive orchids in early stages of speciation. *Mol. Ecol. Resour.* **15**, 1375–1384. https://doi.org/10.1111/1755-0998.12422 (2015).
19. Li, Y. L. & Liu, J. X. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.* **18**, 176–177. https://doi.org/10.1111/1755-0998.12719 (2018).
20. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959. https://doi.org/10.1093/genetics/155.2.945 (2000).
21. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x (2005).
22. de Lafontaine, G., Napier, J. D., Petit, R. J. & Hu, F. S. Invoking adaptation to decipher the genetic legacy of past climate change. *Ecology* **99**, 1530–1546. https://doi.org/10.1002/ecy.2382 (2018).
23. Lima, J. S., Ballesteros-Mejia, L., Lima-Ribeiro, M. S. & Collevatti, R. G. Climatic changes can drive the loss of genetic diversity in a Neotropical savanna tree species. *Glob. Change Biol.* **23**, 4639–4650. https://doi.org/10.1111/gcb.13685 (2017).
24. Onogi, A., Shirai, K. & Amano, T. Investigation of genetic diversity and inbreeding in a Japanese native horse breed for suggestions on its conservation. *Anim. Sci. J.* **88**, 1902–1910. https://doi.org/10.1111/asj.12867 (2017).
25. Horton, C. A. et al. Short tandem repeats bind transcription factors to tune eukaryotic gene expression. *Science* **381**, 18 (2023).
26. Farrington, H. L., Edwards, C. E., Bartron, M. & Lance, R. F. Phylogeography and population genetics of introduced Silver Carp (*Hypophthalmichthys molitrix*) and Bighead Carp (*H. nobilis*) in North America. *Biol. Invasions* **19**, 1–23 (2017).
27. Li, D. et al. A multiplex microsatellite PCR method for evaluating genetic diversity in grass carp (*Ctenopharyngodon idellus*). *Aquac. Fish.* **3**, 238–245 (2018).
28. Dong, F. et al. Genetic diversity and population structure analysis of blunt snout bream (*Megalobrama amblycephala*) in the Yangtze River Basin: Implications for conservation and utilization. *Aquac. Rep.* **35**, 101925 (2024).
29. Du, J. et al. Genetic diversity analysis and DNA fingerprinting of different populations of largemouth bass (*Micropterus salmoides*) in China with fluorescence-labeled microsatellite markers. *BMC Genomics* **26**, 1–15 (2025).
30. Wang, W., Ma, C., Ouyang, L., Chen, W. & Ma, L. Genetic diversity and population structure analysis of *Lateolabrax maculatus* from Chinese coastal waters using polymorphic microsatellite markers. *Sci. Rep.* **11**, 15260 (2021).
31. Kalinowski, S. T. hp-rare 1.0: A computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* **5**, 187–189 (2010).
32. Takezaki, N. & Nei, M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389–399 (1996).
33. Shen, Y. et al. Important fish diversity maintenance status of the tributaries in a hotspot fish conservation area in the upper Yangtze River revealed by eDNA metabarcoding. *Sci. Rep.* **14**, 24128. https://doi.org/10.1038/s41598-024-75176-9 (2024).
34. Chen, H. J., Wang, D. Q., Duan, X. B., Chen, D. Q. & Li, Y. Genetic diversity of white bream, *Parabramis pekinensis* from the middle Yangtze river. *Chin. J. Ecol.* **35**, 2175 (2016).
35. Hu, M., Wang, C., Liu, Y., Zhang, X. & Jian, S. Fish species composition, distribution and community structure in the lower reaches of Ganjiang River, Jiangxi, China. *Sci. Rep.* **9**, 10100. https://doi.org/10.1038/s41598-019-46600-2 (2019).
36. Tang, S. J., Si-Fa, L. I., Cai, W. Q. & Zhao, Y. Microsatellite analysis of variation among wild, domesticated, and genetically improved populations of blunt snout bream (*Megalobrama amblycephala*). *Zool. Res.* **35**, 30–39 (2014).
37. Lynch, M., Wei, W., Ye, Z. & Pfrender, M. The genome-wide signature of short-term temporal selection. *Proc. Natl. Acad. Sci.* **121**, 10 (2024).
38. Lake, N. J. et al. Quantifying constraint in the human mitochondrial genome. *Nature* **635**, 390–397 (2024).
39. Huang, Y. M., Xia, M. Y. & Huang, S. Evolutionary process unveiled by the maximum genetic diversity hypothesis. *Hereditas* **35**, 599–606 (2013).
40. Hardy, G. H. Mendelian proportions in a mixed population. *Science* **28**, 49–50. https://doi.org/10.1126/science.28.706.49 (1908).
41. Bucklin, K. A. et al. Assessing genetic diversity of protected coho salmon (*Oncorhynchus kisutch*) populations in California. *Can. J. Fish. Aquat. Sci.* **64**, 30–42 (2007).
42. Wang, L., Meng, Z., Liu, X., Zhang, Y. & Lin, H. Genetic diversity and differentiation of the orange-spotted grouper (*Epinephelus coioides*) between and within cultured stocks and wild populations inferred from microsatellite DNA analysis. *Int. J. Mol. Sci.* **12**, 4378–4394. https://doi.org/10.3390/ijms12074378 (2011).
43. Wright, S. *Evolution and the genetics of populations. Vol. 1. Genetic and biométrie foundations*. (Evolution and the genetics of populations, 1968).
44. Knowles, L. L. et al. Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Hered. Int. J. Genet.* **119**, 287–294 (2017).
45. Wang, J. et al. Analysis of the genetic structure and diversity of upland cotton groups in different planting areas based on SNP markers. *Gene* **809**, 146042 (2022).
46. Liu, Y. et al. Geographic population genetic structure and diversity of Sophora moorcroftiana based on genotyping-by-sequencing (GBS). *PeerJ* **8**, e9609. https://doi.org/10.7717/peerj.9609 (2020).

Author contributions

Y.L .and C.Y .designed the research framework and wrote the main manuscript text. H.L.participated in the collection of Parabramis pekinensis specimens. J.X.participated in the collection of Parabramis pekinensis specimens. L.P. and C.W.participated in the cultivation of experimental fish stocks. L.H.and Z.Y. and L.J. participated in data recording.

Funding

The funding was provided by National Freshwater Aquatic Germplasm Resource Bank (Grant No. FGRC: 18537), Basic Scientific Research Fund Project of Provincial Public Welfare Research Institutes in Jilin Province (Grant No. JSCYJK202301), Jilin Provincial Science and Technology Development Program (Garnt No. 20220202064NC).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-10664-0.

Correspondence and requests for materials should be addressed to Y.L. or C.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.