



Range-wide population genetics of the tapertail anchovy *Coilia nasus* based on exon-capture data

Huirui Huang^{1,2} · Junlong Jiang^{1,2} · Fangyuan Cheng^{1,2} · Kishor Kumar Sarker^{1,2} · Jinkoo Kim³ · Chenhong Li^{1,2} 

Received: 6 August 2021 / Revised: 13 April 2022 / Accepted: 17 May 2022

© The Author(s), under exclusive licence to Senckenberg Gesellschaft für Naturforschung 2022

Abstract

The tapertail anchovy (*Coilia nasus*) is an economically important species, mainly distributed along the coast of the northwestern Pacific and associated freshwater bodies, including the Qiantang River, the Yangtze River, the Liaohe River, and the Yalu River of China, and further eastward to Korean Peninsula and Ariake Sound of Japan. There have been many studies on population genetics of *C. nasus*, but those were either focused on a few populations or used limited number of loci. The results are still controversial and the range-wide population structure of *C. nasus* is not resolved. This study is aimed to estimate genetic differences among populations from Japan, Korea and the major drainages of China using thousands of loci collected by exon-capture method. The reconstructed maximum likelihood tree, Network, and STRUCTURE analysis confirmed that the population from Lake Dongting should be considered as a separate species, *C. brachygnathus*, whereas the other populations were mixed together except that fish collected from the Shuangtaizi River of Liaohe drainage were grouped with the fish from Dongting. The AMOVA revealed that genetic variation was 6.36% among populations and 93.61% among individuals within population, when *C. brachygnathus* was excluded from the analysis. Pairwise F_{ST} showed that genetic difference between *C. brachygnathus* and *C. nasus* was high (0.0889–0.7350) and difference among the other populations were low (0.0086–0.3127) but significant, suggesting imperfect natal homing of migratory *C. nasus*. According to our results, an integrated management strategy should be taken jointly by the countries of this region to protect the valuable fisheries resource of *C. nasus*.

Keywords Migratory anchovies · Landlocked population · Genetic diversity · Population structure · The northwestern Pacific · Fisheries management

Introduction

There are a series of marginal seas in the northwestern Pacific with complex ocean currents. The population structure of marine organisms of this region may have been shaped by environment factors and historical events during the Pleistocene glacial and interglacial cycles (Hewitt 2000). Species with different life histories may have responded differently to the marine environment and show different patterns in population structure. For example, the spottedtail goby (*Synechogobius ommaturus*) have restricted gene flow between populations probably due to their benthic lifestyle and limited dispersal

capability (Song et al. 2010). The gizzard shad (*Konosirus punctatus*) has two distinct lineages that appear to be closely related to the separation of the seas during the Pleistocene glacial cycles (Gwak et al. 2015), whereas the Japanese anchovy (*Engraulis japonicus*), a pelagic species exhibits no significant population structure (Liu et al. 2006).

The tapertail anchovy (*Coilia nasus*) is a small to medium-sized migratory fish distributed in the coastal and estuarine areas of the northwestern Pacific, including the Qiantang River, the Yangtze River, the Liaohe River and the Yalu River of China, and further eastward to the Korean Peninsula and Ariake Sound of Japan (Yang et al. 2011;

Communicated by S. E. Lluch-Cota

✉ Chenhong Li
chli@shou.edu.cn

¹ Shanghai Universities Key Laboratory of Marine Animal Taxonomy and Evolution, Shanghai Ocean University, Shanghai 201306, China

² Engineering Research Center of Environmental DNA and Ecological Water Health Assessment, Shanghai Ocean University, Shanghai 201306, China

³ Department of Marine Biology, Pukyong National University, 45 Yongso-ro, Nam-Gu, Busan 48513, Korea

Whitehead et al. 1988). The spawning season of *C. nasus* is from May to August. At this time, adults migrate to river estuaries and further upstream to spawn (Liu et al. 2018), and some even settle in freshwater lakes (Cheng et al. 2019; Li et al. 2015; Yuan et al. 1980; Wang et al. 2017; Yang et al. 2010; Jiang et al. 2018). Different populations of *C. nasus* may show different life history strategies for spawning and use different habitats, such that some breed in the estuary area and some migrate far upstream into freshwater lakes and streams to spawn (Cheng et al. 2019; Liu et al. 2018). It has been shown that the sticklebacks (*Gasterosteus aculeatus*) with different life histories often have high genetic differentiation, because they can adapt to the new environment quickly after being isolated by freshwater, resulting in reproductive isolation (Kristjansson et al. 2002; Lavin and Mcphail 1993). So, have different life histories of *C. nasus* led to high genetic differentiation among populations in different river systems? Furthermore, anadromous fish such as *Oncorhynchus keta* and *Salmo gairdneri* tend to spawn at parental habitat (Tallman and Healey 1994; Lindsey et al. 1959). Therefore, *C. nasus* may also have natal homing behavior. If so, how would it affect the genetic structure among its populations from different drainages along the coast of the northwestern Pacific? These questions have not been fully resolved yet.

Three species/subspecies of *C. nasus* species complex have been reported based on morphological characters in the Yangtze River drainage, including the migratory *C. nasus*, landlocked *C. nasus taihuensis* and freshwater *C. brachygnathus* (Yuan et al. 1976), but subsequent analysis using mitochondrial loci or AFLP data suggested that all three were genetically indistinguishable (Yang et al. 2010). However, more recent studies applying microsatellite markers or genome-wide SNP data suggested that *C. nasus* and *C. brachygnathus* are two distinct species (Cheng et al. 2019; Xue et al. 2019; Xuan et al. 2021). Genetic relationships among populations other than *C. brachygnathus* also are controversial. Analyses based on phylogenetic reconstruction using mitochondrial loci suggested that fish from different populations were mixed together and no monophyletic clades were found related to any localities (Yang et al. 2010; Sun et al. 2019; Yang et al. 2018), but comparison of genetic distance or allele frequencies suggested that there was significant genetic divergence between the Japanese population and the continental populations (Gao et al. 2014; Sun et al. 2019; Xue et al. 2020; Yang et al. 2018; Yang et al. 2011). The samples taken from the Huaihe River drainage also were found to be different from other samples in one study using microsatellite data (Xue et al. 2019), but no difference was found in another study using more samples and exon-capture data (Cheng et al. 2019). In summary, the current consensus is that *C. brachygnathus* and *C. nasus* are two distinct species and *C. nasus* and *C. nasus taihuensis* are

indistinguishable, but the genetic relationships of populations of *C. nasus* from different countries and drainages are still controversial.

Mitochondrial loci often are the first choice for the population study of fishes, because they are easy to amplify, but microsatellite loci are preferred than mitochondrial markers, because mitochondrial loci are maternally inherited and evolve slower than microsatellite loci. Nonetheless, genome-wide single nucleotide polymorphisms (SNPs) data can be more powerful due to their high numbers (Lee and Edwards 2008). Exon-capture can be used to obtain data from thousands of loci at a low cost and it is easy to identify their orthology across species (Jiang et al. 2019). Exon-capture data are not as “neutral” as anonymous markers such as those from restriction site associated DNA sequencing (RAD Sequencing), so scrutiny should be applied when analyzing exon-capture data for studying genetic patterns among populations.

Coilia nasus is an economically important fish. Understanding genetic structure among different populations is crucial for sustainable use and management of this valuable resource. Previous studies on the genetic structure of populations of *C. nasus* have generated inconsistent results, probably due to incomplete sampling of the populations and use of inadequate loci used. In this study, we collected sequence data from thousands of nuclear loci by applying a cross-species exon-capture method (Li et al. 2013) on range-wide population samples to assess 1) genetic relationships of *C. nasus* among populations from China, Japan and the Korean Peninsula, and population structure of *C. nasus* of the major river systems in China; and 2) genetic diversity of all populations for a better management of *C. nasus*.

Materials and methods

Sample collection and DNA extraction

A total of 145 individuals were collected. *Coilia brachygnathus* often is diagnosed by its short supermaxilla that does not reach to the edge of the operculum, whereas the supermaxilla of the migratory *C. nasus* does (Whitehead et al. 1988). The short supermaxilla, however, is also seen in the landlocked samples of Lake Tai and later found in some samples collected from the main channel of the Yangtze River as well (Tang et al. 2007). Therefore, we tentatively assigned all fish to the names according to their locality: “*C. brachygnathus*” for fish from Lake Dongting, “*C. nasus taihuensis*” for fish from other freshwater lakes and “*C. nasus*” for the rest (Table 1). The fish were collected from coast and associated freshwater systems of Korea (KOR), Japan (JAP) and 16 sampling sites in China (Fig. 1), including the Liaohe River (LH), the Yellow Sea (YS), the Yalu River

Table 1 Sampling information of the 145 individuals of *Colia nasus*

Region	Abb.	Habitat	Site	Sampling location	Sample ID	No. of samples	Date of collection	Latitude	Longitude
Korea	KOR	Coastal	1	Hongdo, Heuksanmyeon, Sinan-gun, Jeollanam, Korea	CL603, CL606, CL607, CL614, CL616, CL617, CL624	10	2008	34°25'20.52"N	126°07'46.11"E
Japan	JAP	Semi-closed Bay	2	Ariake Sea, Japan	CL1308	2	2009	32°47'37.65" N	130°28'48.22"E
Liaoh River	LH	Liaoh River water System	3	Shuangtaizi River, Panjin, Liaoning	CL848	10	2016/12/14	40°57'7.70"N	121°49'49.60"E
Yellow Sea	YS	Coastal	4	Yingkou, Liaoning	CL1276	9	2016/6/12	34°40'56.04" N	122°18'10.60"E
Yalu River	YL	Yalu River water system	5	Dandong, Liaoning	CL1164	14	2016/11/13	39°59'39.57"N	124°25'28.39"E
Lake Hongze	HZ	Huaihe River water System	6	Sihong, Suqian, Jiangsu	CL506	10	2016/7/22	33°25'36.89" N	118°13'18.24"E
Lake Gaoyou	GY	Huaihe River water System	7	Gaoyou, Jiangsu	CL410	8	2013/11/12	32°46'3.40" N	119°24'49.30"E
Luan	LA	Huaihe River water System	8	Lu'an, Anhui	CL1256, CL1240	9	2013/4/10	31°45'11.70"N	116°28'58.54"E
Lake Chao	CH	Yangtze River water System	9	Chaohu, Anhui	CL1235, CL1236, CL507,	19	2016/9/6	31°33'22.13" N	117°21'24.99"E
Lake Tai	TH	Yangtze River water System	10	Wuxi, Jiangsu,	CL540, CL519, CL63, CL66, CL68	21	2016/11/6	31°16'57.72" N	120°23'35.70"E
			11	Suzhou, Jiangsu,			2016/11/3	31°28'28.81" N	120°12'51.31"E
			12	Guangfu, Suzhou, Jiangsu			2013/11/22	31°11'56.24" N	120°26'50.28"E
Estuarial Yangtze River	SH	Yangtze River water System	13	Chongming, Shanghai,		12	2013/4/1	31°36'19.71" N	121°23'42.86"E
			14	Luchaogang, Shanghai,			2013/11/23	30°50'30.32" N	121°50'58.48"E
			15	Jinjiang, Jiangshu			2012	31°58'10.70" N	120°19'18.60"E
							2013		
Qiantang River	QT	Qiantang River water System	16	Hangzhou, Zhejiang	CL452, CL1257, CL1237,	4	2016/8/20	29°47'38.47"N	119°40'46.45"E
Lake Dongting	DT	Yangtze River water System	17	Yueyang, Hunan		17	2013/4/22	29°17'2.60"N	113°02'36.19"E
			18	Yuanjiang, Hunan,			2016/9/8	28°51'15.66"N	112°21'53.93"E
							2016/11/4		

(YL), Lake Hongze (HZ), Lake Gaoyou (GY), Luan (LA), Lake Chao (CH), Lake Tai (TH), mouth of the Yangtze River (SH), the Qiantang River (QT) and Lake Dongting (DT). The sampling sites in China included five river drainages: the Qiantang River (QT), the Yangtze River (CH, TH, SH, DT), the Huaihe River (HZ, GY, LA), the Liaohe River (LH) and the Yalu River (YL). The sampling map was edited from source file downloaded from the National Platform for Common Geospatial Information Services of China (<https://www.tianditu.gov.cn/>). Fin or muscle tissue were taken and preserved in 95% alcohol, then stored in a refrigerator at 4 °C.

DNA was extracted by using Ezup column animal genomic DNA extraction kit (SANGON, Shanghai, China, # Cat: B518251). Integrity of extracted DNA was checked with agarose gel electrophoresis. Concentration of DNA was measured using a nanodrop 3300 ultramicro fluorescence spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA).

Molecular markers, library construction, gene enrichment, and high-throughput sequencing

A suite of 4,434 target loci were developed for ray-finned fishes by Jiang et al. (2019). MYbaits RNA probes targeting the 4,434 exon of clupeiform fishes were synthesized at Arbor Biosciences (Ann Arbor, USA; cat#: ClupiformsV2 MYbaits-1). These sets of probes were successfully used in

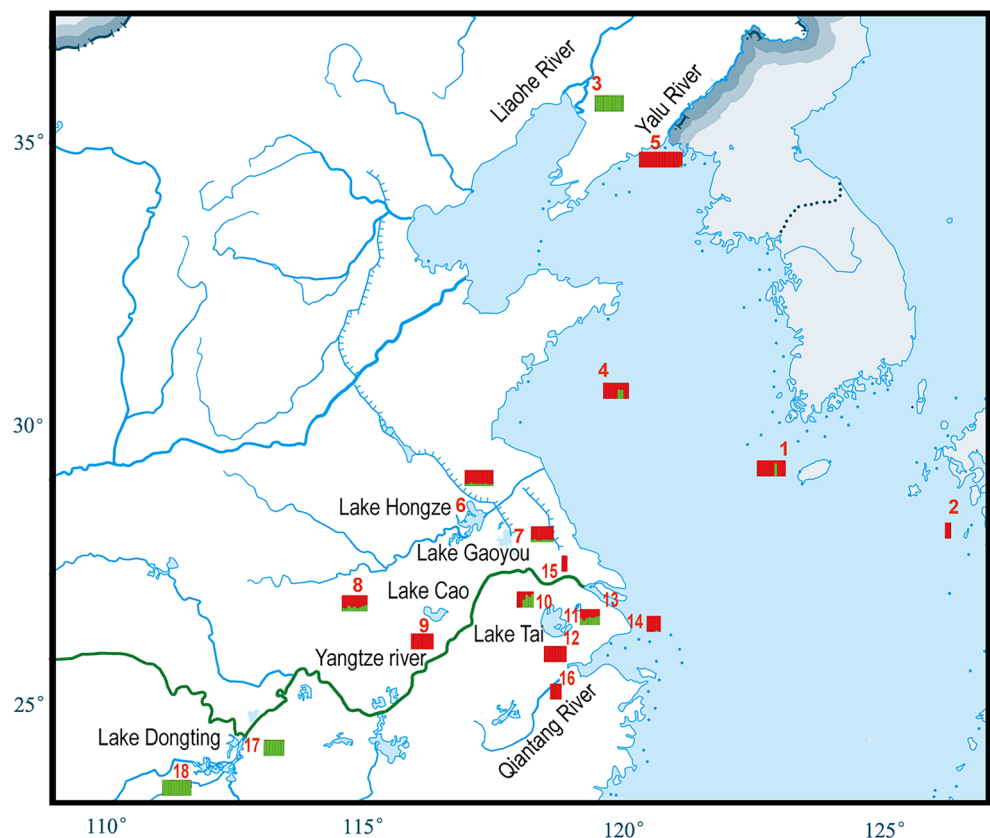
investigating freshwater invasion of *C. nasus* in the Yangtze River (Cheng et al. 2019).

The extracted DNA was sheared to around 250 bp using Covaris M220 (Woburn, Massachusetts, USA). DNA library was constructed according to the methods of Li et al. (2013). Gene enrichment of 4434 single-copy nuclear genes was carried following the protocol of Cheng et al. (2019). The enriched libraries were sequenced on an Illumina HiSeq 2500 platform by GENWIZ biotechnology company (Suzhou, China).

Read assembly and SNP calling

The Assexon pipeline (Yuan et al. 2019) was employed to process the reads. The sequence reads were separated to each individual according to 8 bp barcode sequence. The sequencing adaptor and reads with low quality bases (Qscore < 20) were trimmed by using a custom Perl script (trim_adaptor.pl). Removal of PCR repeats and sequence assembling followed the Assexon pipeline (Yuan et al. 2019). The 4434 loci of *Danio rerio* were used as the reference for read assembling (Jiang et al. 2019). Sequences with poor alignment results were filtered out and the reference sequences were removed for re-alignment using a custom Perl script (Yuan et al. 2019). The loci with poor enrichment result (only one sample sequence) were excluded from further analysis.

Fig. 1 Sampling localities: KOR (1), JAP (2), LH (3), YS (4), YL (5), HZ (6), GY (7), LA (8), CH (9), TH (10, 11, 12), SH (13, 14, 15), QT (16), DT (17, 18). The details of the abbreviation of the region name and the sample site number can be found in Table 1. The colored histogram show the allelic composition of each individual sampled at each site estimated from STRUCTURE analysis. Red/dark indicates alleles from *Coilia nasus*, and green/light shows alleles from *C. brachygnathus*. Map source: National Platform for Common Geospatial Information Services



Consensus sequences were made from the alignment of each locus using a custom Perl script (consensus.pl) (Cheng et al. 2019). BWA v0.7.16 (Li and Durbin 2009) were used to index the consensus sequence. The trimmed reads were aligned to the consensus sequence using BWA v0.7.16. BAM format conversion and PCR repeat deduplication were completed using Picard (<http://broadinstitute.github.io/picard/>). The default parameters of GATK pipeline were used for SNP calling (Depristo et al. 2011, Van der Auwera et al. 2013). The obtained VCF file was converted into NEXUS file and STRUCTURE file by a custom Perl script (vcftosnps.pl) (Cheng et al. 2019). In our previous study on freshwater invasion of *Coilia* in the Yangtze River, we found 124 outlier exons between the anadromous fish (SH) and the resident fish of DT and PY and 22 outliers between the anadromous fish (SH) and the resident fish of CH and eastern TH applying F-DIST analysis (Cheng et al. 2019), so we analyzed the data both excluding and including these adaptive loci for comparison.

Genetic diversity and population analysis

A VCF format file containing all SNP sites was used as the input file of DNAsp v5.10 (Librado and Rozas 2009) to calculate the nucleotide diversity (π). A phylogenetic tree was constructed using the maximum likelihood method in RAxML and multipoint data (Stamatakis 2014). A GTRGAMMA model was set for the ML analysis with 1000 bootstrap replicates. Arlequin v3.5 (Excoffier and Lischer 2010) was used to calculate pairwise fixation index (F_{ST}) (Weir and Cockerham 1984) and Analysis of Molecular Variance (AMOVA) (Excoffier et al. 1992) based on SNP data. Default parameters were used to calculate pairwise F_{ST} where the iteration parameter of AMOVA was set to 1000. A total of 934 SNP sites with less than 4% missing data were used for network analysis with median-joining (MJ) calculation (Bandelt et al. 1999). VCF file was filtered using Vcftools 0.1.15 (<https://vcftools.github.io/>). Weight of characters was set as 10, and weights of single nucleotide mutation transversions and transitions was set as 1:1, and epsilon = 0, frequency > 1 and MJ square option as inactive.

Genetic structures of the populations were investigated using STRUCTURE v2.3.4 (Pritchard et al. 2000) with the SNP data. The parameters were set as: length of burnin period, 500; number of MCMC reps after burnin, 500000; number of genetic cluster, K value: 1 to 3 and each run was repeated for three times. The best K value was retrieved from the online platform “STRUCTURE HARVESTER” (Earl and Vonholdt 2012).

Mantel test with 1000 permutation was used to test the null hypothesis of isolation by distance between populations of *C. nasus* distributed in different drainages (Mantel 1967; Smouse et al. 1986). The geographic distance between

sampling locations was calculated along the rivers and coastlines using ArcGIS v10.2 (ESRI 2011). The genetic distance was represented by F_{ST} values.

Results

Read assembly and SNP calling

Each fish was enriched to an average of 1,920 loci with a maximum of 2,714 and a minimum of 981 loci. A total of 3,761 loci were obtained for all fish and 2,565 loci remained after filtering. Through GATK mutation detection, we obtained 9,572 SNP sites, with an average of 3.73 SNP sites per locus. We found 95 adaptive loci in the resulting data. The results excluding these adaptive loci were almost the same as those including these loci. Hereafter, we only report the results from analyses excluding the adaptive loci.

Genetic diversity and population structure

The nucleotide diversity (π) of different populations ranged from 0.00011 (LH) to 0.0063 (KOR), with the LH population having the lowest nucleotide diversity and the Korea population having the highest (Table 2). AMOVA analysis of all populations showed that 49.55% of the molecular variation existed between *C. nasus* and *C. brachygnathus* ($p < 0.01$), 4.54% among populations within *C. nasus* and *C. brachygnathus* ($p < 0.01$) (Table 3). After removing

Table 2 Nucleotide diversity (π) of each population

Population	Theta (per site) from π
LH	0.00106
DT	0.00111
JAP	0.00161
YS	0.00222
CH	0.00232
TH	0.00298
SH	0.00311
YL	0.00314
GY	0.00335
QT	0.00415
LA	0.00463
HZ	0.00467
KOR	0.00631

Populations: Korea (KOR), Japan (JAP), Liaohe River (LH), the Yellow Sea (YS), the Yalu River (YL), Lake Hongze (HZ), Lake Gaoyou (GY), Luan (LA), Lake Chao (CH), Lake Tai (TH), mouth of the Yangtze River (SH), the Qiantang River (QT) and Lake Dongting (DT)

Table 3 Results of analysis of molecular variance (AMOVA). All populations included

Source of variation	Sum of squares	Variance components	Percentage of variation	<i>P</i> -value
Among groups (species)	682.999	14.791	49.55	0
Among populations within groups	309.611	1.356	4.54	0.00098
Within populations	1808.894	13.704	45.91	0.0088
Total	2801.503	29.851		

C. brachygnathus from the AMOVA analysis, the genetic variation was 6.39% among populations ($p < 0.01$) and the remaining variation was among individuals within population (Table 4). The F_{ST} value between *C. nasus* and *C. brachygnathus* were high ($F_{ST} = 0.2870\text{--}0.7350$, $P < 0.00064$ after Bonferroni correction) (Table 5). After removing *C. brachygnathus*, the genetic differentiation within population was low but significant ($F_{ST} = 0.0086\text{--}0.3127$, $P < 0.00064$). The genetic differentiation between DT and LH was low and not significant ($F_{ST} = 0.0889$, $P = 0.00901$); whereas the F_{ST} value between DT and other groups ($0.5108\text{--}0.7350$), and between LH and other groups ($0.2870\text{--}0.6554$) were high ($P < 0.00064$) (Table 5). The paired F_{ST} values among 8 populations in five Chinese water systems (the Qiantang River, the Yangtze River, the Huaihe River, the Liaohe River and the Yalu River) were lower ($0.0086\text{--}0.2649$) and some values were significant ($P < 0.00064$), but the values are far less than the F_{ST} value between *C. nasus* and *C. brachygnathus* (Table 5).

The maximum likelihood tree from RAxML analysis showed that the 145 individuals were mainly divided into two clades. One clade included QT, HZ, TH, CH, LA, YL, GY, SH, YS, JAP, and KOR samples and the other included LH and DT, with the exception that two individuals of *C. nasus taihuensis* were clustered into the clade of *C. brachygnathus* (Fig. 2). The samples from Korea, Japan, and most of the Chinese water were gathered together with no monophyletic groups formed according to locality (Fig. 2). The result of STRUCTURE analysis also showed that two groups ($K = 2$) best represented the population variation (Fig. 3). LH and DT samples formed one group, and the rest of samples were clustered as the other group, which confirmed that *C. nasus* from China, Japan, and Korea maintained substantial gene flow among them. Mantel test showed that there

was a moderate correlation between geographical distance and genetic structure ($r = 0.358746$, $P = 0.027$). Some individuals from TH, HZ, GY, and LA were mixed in between of the two groups suggesting their hybrid origin, which is consistent with the previous study (Fig. 3) (Cheng et al. 2019). Pairwise F_{ST} showed that genetic difference between LH and DT and other populations was high ($0.2870\text{--}0.7350$, $P < 0.00064$) and difference among the other populations were low but significant ($0.0086\text{--}0.3127$, $P < 0.00064$) (Table 5). Network analysis also revealed the same pattern as the results of the RAxML tree and the Structure analysis. The results of the Network analysis showed that the populations of LH and DT and other populations were separated (Fig. 4) and two samples of TH were nested in *C. brachygnathus*.

Discussion

New population of *Coilia brachygnathus* found in Liaohe River

Coilia brachygnathus was described from fish found in Lake Dongting based on morphological characters, such as the rear extension of supermaxilla being shorter than that of *C. nasus* (Yuan et al. 1980). However, individuals with shorter supermaxilla also have been found in Lake Tai and main tributaries of the Yangtze River (Ni and Wu 2006; Tang et al. 2007), so the morphological differences are not always reliable for taxonomy of *Coilia*. Subsequent molecular analyses suggested that *C. brachygnathus* and *C. nasus* are the same species and indistinguishable genetically (Tang et al. 2007; Yang et al. 2010). However, these early studies used only mitochondrial loci (Tang et al. 2007), so the data may not have had adequate power to differentiate recent diverged species.

Table 4 Results of analysis of molecular variance (AMOVA). Analysis with population of Lake Dongting and the Liaohe River excluded

Source of variation	Sum of squares	Variance components	Percentage of variation	<i>P</i> -value
Among populations (within <i>C. nasus</i>)	491.483	1.957	6.39	0.00684
Within populations	3069.610	28.688	93.61	
Total	3561.093	30.645		

Table 5 Pairwise F_{ST} values between populations (below diagonal) and their p-value (above diagonal). The asterisk indicates significance at level of $p < 0.00064$ after Bonferroni correction

	QT	SH	TH	GY	HZ	CH	LA	YL	YS	JAP	KOR	LH	DT
QT		0	0	0	0	0.01802	0	0	0.69369	0.15315	0.28829	0	0
SH	0.1951*		0.07207	0	0	0.06306	0.28829	0	0	0.01802	0	0	0
TH	0.1878*	0.0758		0	0	0.25225	0.01802	0	0	0.02703	0	0	0
GY	0.2615*	0.1378*	0.0596*		0	0	0.17117	0	0	0.00901	0	0	0
HZ	0.2649*	0.1109*	0.0299*	0.0101*		0	0	0	0.09910	0	0.01802	0	0
CH	0.2190	0.1115	0.0880	0.0113*	0.0340*		0.01802	0	0	0.03604	0	0	0
LA	0.2150*	0.1196	0.0086	0.0464	0.0235*	0.0979		0	0.04505	0.08018	0	0	0
YL	0.2454*	0.0909*	0.0925*	0.1023*	0.0510*	0.0717*	0.0708*		0	0.02703	0.00901	0	0
YS	0.0205	0.1069*	0.1245*	0.1545*	0.1321	0.1631*	0.0974	0.0983*		0.09009	0.8018	0	0
JAP	0.2028	0.2454	0.2061	0.2687	0.3127*	0.2625	0.2424	0.2893	0.0857		0.10811	0.03604	0
KOR	-0.0857	0.0574*	0.0723*	0.0359*	0.0325	0.0766*	0.0264*	0.0254	-0.04070	-0.0999		0	0
LH	0.5346*	0.5525*	0.3819*	0.5873*	0.5872*	0.5816*	0.4856*	0.6139*	0.4381*	0.6554	0.2870*		0.00901
DT	0.6924*	0.6682*	0.5327*	0.6854*	0.6818*	0.6870*	0.6172*	0.7060*	0.6108*	0.7350*	0.5108*	0.0889	

Populations: Korea (KOR), Japan (JAP), Liaohe River (LH), the Yellow Sea (YS), the Yalu River (YL), Lake Hongze (HZ), Lake Gaoyou (GY), Luan (LA), Lake Chao (CH), Lake Tai (TH), mouth of the Yangtze River (SH), the Qiantang River (QT) and Lake Dongting (DT)

Yang et al. (2010) applied AFLP data to investigate the relationship between *C. brachygnathus* and *C. nasus* and concluded that they were the same species, but they mistakenly assigned *C. nasus* collected from Lake Tai as *C. brachygnathus* based on its short supermaxilla extension that was superficially similar to *C. brachygnathus*. Recent studies using more nuclear loci have unequivocally indicated that *C. brachygnathus* is a valid species genetically different from *C. nasus* (Yang et al. 2018; Xue et al. 2019; Xuan et al. 2021).

Our results showed that populations from the LH and DT were grouped together, whereas the other populations formed another cluster, except that two *C. nasus* from Lake Tai were grouped with fish from LH and DT (Figs. 2 and 4). The STRUCTURE analysis also showed that the best number of groups, $K = 2$, which divided the LH and DT populations from others (Fig. 3). AMOVA results showed that 49.55% of the molecular variation was mainly caused by the differences between *C. nasus* and *C. brachygnathus* ($P < 0.01$) (Table 3). Although the TH, HZ, GY and LA populations had frequent gene exchange with LH and DT populations (Figs. 2 and 3), they still reached high level of differentiation ($F_{ST}=0.3819-0.6854$, $P < 0.00064$) (Table 5). The results of STRUCTURE analysis provided additional evidence that TH, HZ, GY, and LA are diverged from the LH and DT group (Fig. 3). All of our results are consistent with the current consensus, that *C. nasus* and *C. brachygnathus* are two valid species. The two individuals of *C. nasus* of Lake Tai that clustered with *C. brachygnathus* were probably due to hybridization (Figs. 2 and 4).

A surprising finding of our results is that the samples from the LH are genetically similar to samples from DT (Figs. 2 and 3). The F_{ST} between LH and DT was much lower than that between them and the other groups (Table 5). The nuclear diversity of LH and DT populations was quite smaller than that of other populations, suggesting the small effective population size of *C. brachygnathus* comparing to *C. nasus* (Table 2). It was estimated that the first freshwater invasion of *C. nasus* occurred 4.07 million years ago and it gradually settled in Lakes in the middle reach of Yangtze River and differentiated into *C. brachygnathus* from its anadromous ancestor (Cheng et al. 2019). *Coilia brachygnathus* has been recorded nowhere other than the middle reach of the Yangtze River and its associated lakes. Zhang et al. (2020) collected “*C. nasus*” samples from the Dayang River, an upper reach of Liaohe River as well. They found that “*C. nasus*” of the Dayang River was different from that of the Yalu River and Lake Taihu, which was consistent with our results. Both the Shuangtaizi River and the Dayang River are located in the inland fresh water basin of the Liaohe River. Therefore, there may be a population of *C. brachygnathus* in the Liaohe River Basin undiscovered previously. Xue et al. (2019) showed that the Liaohhe population was not different from other populations of *C. nasus* in China which was different from our results. This may be related to their sampling site being different from ours. We collected samples from the Shuangtaizi River, an upstream section of Liaohe drainage. Why *C. brachygnathus* appears in Liaohe River basin and whether there are populations of *C. brachygnathus* in other regions should be investigated with more samples from different locations.

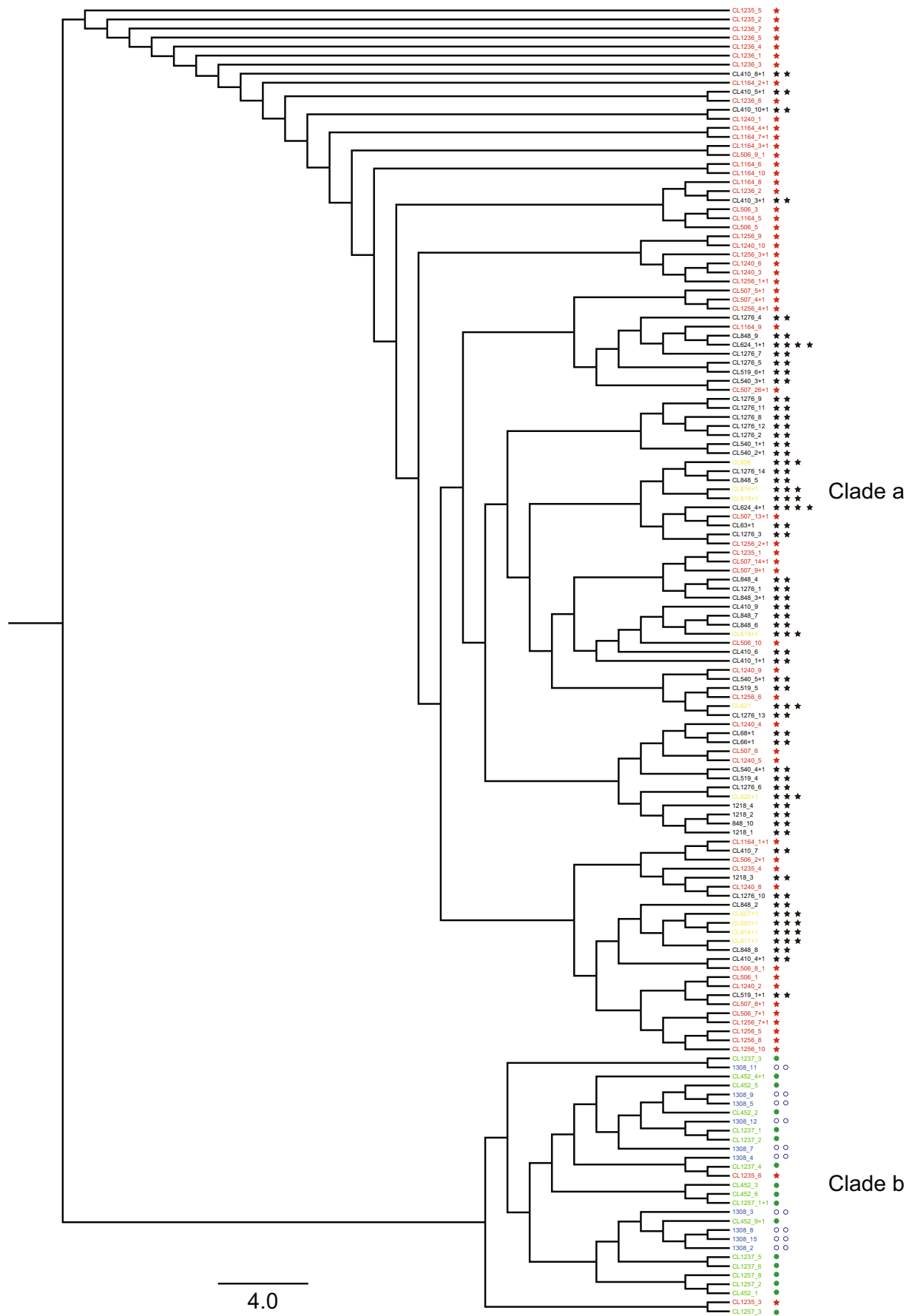


Fig. 2 A ML tree based on sequences concatenating 2,565 loci. Clade a, *Coilia nasus*; clade b, *C. brachygnathus*. One asterisk, freshwater *C. nasus*; two asterisks, migratory *C. nasus*; three asterisks, *C. nasus* from Korea, and four asterisks, Japanese samples. One dot, *C. brachygnathus* from Lake Dongting; two circles, *C. brachygnathus* from the Liaohe River

Genetic difference among populations of *Coilia nasus* from China, Korea, and Japan

As the northwestern Pacific is vast, there may be limited gene flow or significant genetic differences among different geographical populations in the Northwest Pacific. There are many factors that usually affect the population structure of marine fish including physical barriers, ocean currents, life history of fish, and juvenile active movement. Some marine organisms distributed along the northwestern Pacific have significant population differentiation such as *Synechogobius ommatus* (Song et al. 2010). Weak dispersal ability limited admixture of *S. ommatus* individuals which ultimately corresponds to restricted gene flow, and genetic distance increases with geographic distance between populations (Song et al. 2010). In contrast, *E. japonicus* has no notable genetic difference at population level throughout the northwestern Pacific probably because of its high dispersal ability.

For pelagic fishes, the blocking effect of ocean barrier is weak, where ocean current contribute to dispersal and the exchange of genes among fish from different locations, homogenizing genetic constituent of the species (Ward et al. 2010; Grant and Bowen 1998). For anadromous fishes, such as *C. nasus*, the individual movement route is more complex, and the swimming range is larger, including freshwater basins (inland rivers, estuaries) (Masson et al. 2018), which may lead to limited gene flow among populations. The results of STRUCTURE analysis showed that the sampled *C. nasus* from China, Korea and Japan were combined as one group (Fig. 3). The RAxML tree also showed that most of the fish from populations in China, Korea and Japan were grouped

together (Fig. 2). The results of F_{ST} show that the level of genetic difference between the populations from the three countries was far less than that between them and DT and LH populations (Table 5). Therefore, we concluded that most populations of Korea, Japan and China (except LH and DT) belong to one species though morphological differences were reported between *C. nasus* populations in the three countries for characteristics such as the number of anal fin ray, number of longitudinal scales, number of scutes and vertebrae (Takita 1967).

Morphological data can be easily affected by different environmental conditions. Very often, this led to problems in species identification. Molecular studies had confirmed low genetic difference between Japanese and Chinese populations (Sun et al. 2019; Gao et al. 2014; Yang et al. 2016). Analyzing pairwise fixation index (F_{ST}) has revealed that there was more frequent gene exchange between the Korean population and Chinese population than between the Chinese and Japanese population (Table 5). In conclusion, we think that the marine physical barrier did not affect the gene flow of *C. nasus* among different geographic populations and that ocean currents may also aid gene exchange between different populations, so *C. nasus* of different marine geographic populations should be treated as one species.

Genetic structure of *Coilia nasus* among major drainage of China

Low level of population diversity often occurs in the anadromous fish (Ward et al. 2010; Grant and Bowen 1998; Boguski et al. 2016). The genetic divergence values among migratory *C. nasus* of different water systems in China were relatively low but most are significant ($F_{ST} = 0.0708 - 0.2649$; $P < 0.00064$). Thus, gene exchange should take place but is restricted among all water systems. The differentiation level between the Yalu River water system and the Yangtze River water system is small (average $F_{ST} = 0.1290$). Divergence of

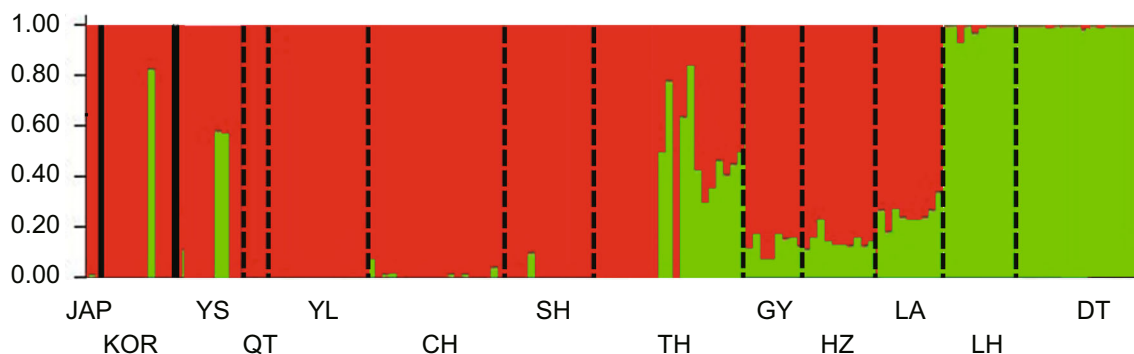


Fig. 3 Structure analyses of all populations, including Japan (JAP), Korea (KOR), the Yellow Sea (YS), the Qiantang River (QT), the Yalu River (YL), Lake Chao(CH), mouth of the Yangtze River (SH), Lake

Tai(TH), Lake Gaoyou (GY), Lake Hongze (HZ), Luan(LA), the Liaohe River (LH), and Lake Dongting (DT). Countries are separated by solid lines, and populations in China are separated by dashed lines

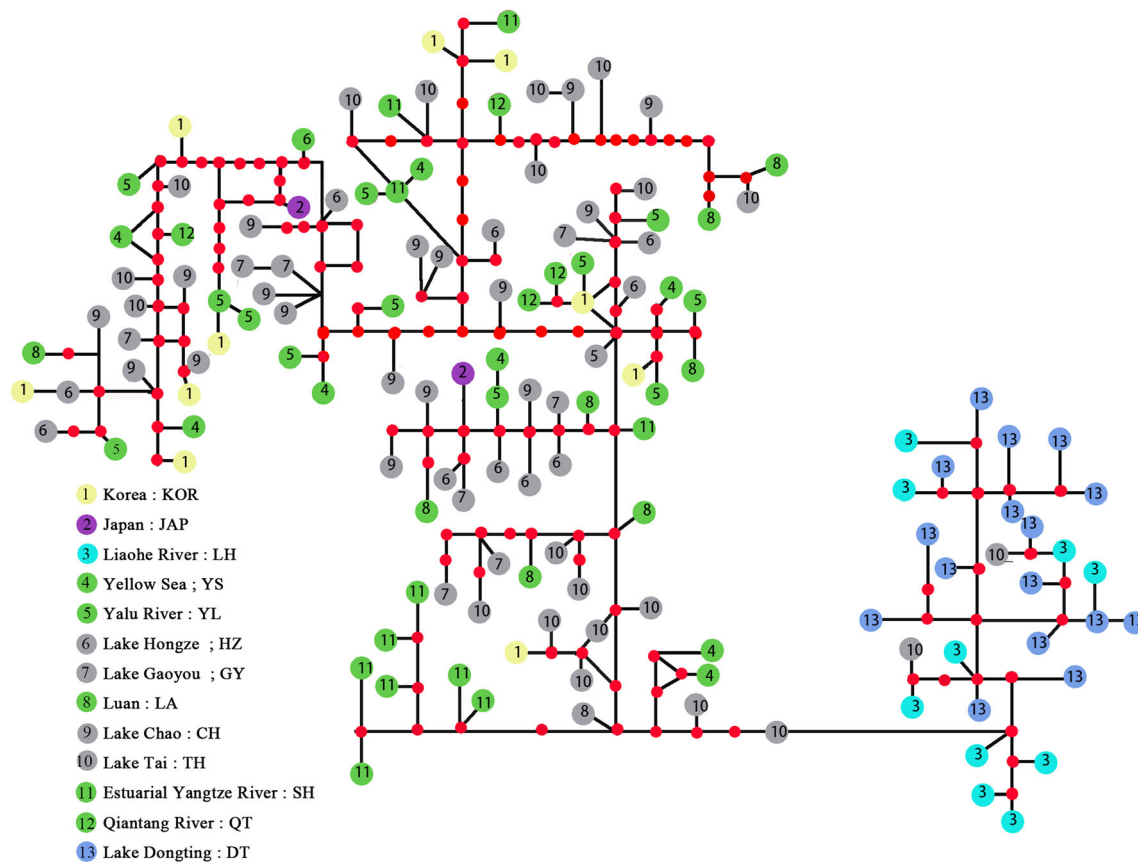


Fig. 4 A network of 145 individuals based on 934 SNP loci. Gray circle (6, 7, 9, 10), freshwater *Coilia nasus*; green circle (4, 5, 8, 11, 12), migratory *C. nasus*; yellow circle (1), *C. nasus* from Korea; purple circle (2), Japanese

samples; blue circle (13), *C. brachygnathus* from Lake Dongting and light blue (3), *C. brachygnathus* from the Liaohe River

the QT population from other water systems was slightly higher than that among other water systems (average $F_{ST} = 0.2270$), but far less than the difference between species. It may be due to the different environment of the Qiantang River and its locations. Although the Qiantang River estuary and the Yangtze River estuary are close to each other, the life history pattern of *C. nasus* in the Qiantang River is different from that in the Yangtze River. In terms of difference between populations, HZ and GY had the lowest F_{ST} value (0.0101). These two lakes are close to each other and both of them have freshwater resident *C. nasus*, so the barrier to gene exchange would be small. On the contrary, there is a large distance between the QT and HZ locations, and their populations belong to different ecotypes, so their F_{ST} value was the largest (0.2649). The result of Mantel test suggested moderate correlation between geographic distance and genetic distance of *C. nasus* distributed in different drainages, so isolation by distance would be one but not the only cause for the genetic patterns among populations.

Some anadromous fishes have a unique habit of returning to their birthplace to lay eggs, such as *Oncorhynchus keta* (Tallman and Healey 1994) and *Salmo gairdneri* (Lindsey et al. 1959). These fish migrate upstream to the spawning site

during the breeding season and rarely deviate from the route. Even if they enter the spawning site of established populations of conspecifics by mistake, they may not be able to reproduce (Tallman and Healey 1994). In our study, the level of differentiation between *C. nasus* populations of different drainages was low but some were significant ($P < 0.00064$), indicating that there was differentiation but also restricted gene exchange between different populations. There was no monophyletic clade corresponding to river drainages in the maximum likelihood tree probably due to their short divergence time or strayed individuals with imperfect natal homing (Fig. 2). Jiang et al. (2016) used laser ablation inductively coupled plasma mass spectrometry (LA-ICPMS) to analyze the element composition in the otolith nucleus region of *C. nasus* in the Yangtze River and showed that *C. nasus* use Lake Poyang, the Yangtze River and the Yellow Sea as their natal lake, migration pathway, and feeding ground, respectively. They also revealed different life history patterns among populations of *C. nasus* from four river drainages in China, and suggested that the distribution of Sr concentrations and Sr:Ca ratios should be a better tool to estimate natal homing and populations structure than

genetic data (Jiang et al. 2014a). In conclusion, we consider *C. nasus* as a migratory species with imperfect natal homing behavior.

Fisheries management

Coilia nasus is one of the most economically important fishes in China and East Asia. In recent years, the natural populations of *C. nasus* are declining rapidly (Ma et al. 2010; Xue et al. 2019), in most cases due to anthropogenic activities (Xue et al. 2019). For instance, a number of water conservancy projects, such as hydroelectric dams have been constructed, which could be barriers on the migratory route of *C. nasus* (Mao et al. 2015). The construction of the Isahaya bay reclamation dike has led to disappearance of some spawning grounds and emergence of freshwater isolated populations (Simanjuntak et al. 2015). Restricted gene flow among populations may contribute to the reduction of genetic diversity. Higher genetic diversity promotes population persistence and resilience, as well as their ability to adapt a changing environment (Presti and Wasko 2014; Li et al. 2017). Overfishing is also becoming a concern for the wild stock of *C. nasus* (Mao et al. 2015). It has been reported that the total catch and average body size of *C. nasus* has been decreasing (Mao et al. 2015; Dai et al. 2020). Tremendous efforts have been made on research and conservation of the resources of *C. nasus*. For example, *C. nasus* has been listed in the "critical species" list of rare wild aquatic organisms in Japan (Liu et al. 2018). In order to protect the fish resources of the Yangtze River, China has implemented a fishing ban on the Yangtze River since 2003. In 2005, all fishing vessels are required to hold a fishing license. In 2019, the issuing of *C. nasus* fishing licenses was stopped, and the fishing of *C. nasus* was banned totally (He 2019).

Most conservation efforts so far have been focused on a few populations or a particular river drainage, but *C. nasus* of other regions have not been considered. The environment of the Qiantang River is very similar to that of the Ariake Sea in Japan, with a large confluence area of seawater and fresh water (Liu et al. 2020). *Coilia nasus* from the Qiantang River and Japan can be divided into three types as short-term, long-term and non-freshwater dependent according to the degree of freshwater dependence (Khumbanyiwa et al. 2018; Liu et al. 2020; Liu et al. 2018). The life history of *C. nasus* in the Qiantang River is different from that of the Yangtze River and the Yellow Sea (Jiang et al. 2014a; Khumbanyiwa et al. 2018; Jiang et al. 2014b). The nucleotide diversity ($\pi=0.00415$) of the *C. nasus* population of the Qiantang River was higher than that of other migratory populations in China (Table 2). We also found that nucleotide diversity was the highest in Korean samples ($\pi=0.0063$) (Table 2). Based on these result, we recommend establishing a "fishery sanctuary"

or "genetic resource protection area" for *C. nasus* near the sampling locations of Korea (Hongdo) and the Qiantang River as well. Integrated management strategy should be taken by the countries of this region to protect the valuable fisheries resource of *C. nasus*.

Acknowledgements We thank Ms Shuli Song for helping with sample collection. We thank the anonymous reviewers for the constructive comments and suggestions.

Funding This work was supported by "Science and Technology Commission of Shanghai Municipality (19050501900)", "Science and Technology Commission of Shanghai Municipality (19410740500)."

Declarations

Conflict of interest The authors declare no competing interests.

Ethics approval All experimental procedures involving fish were approved by the Animal Ethics Committee of Shanghai Ocean University, China.

Sampling and field studies All necessary permits for sampling have been obtained by the authors.

Data availability The raw reads and sequencing data have been uploaded in the SRA under Bioproject number PRJNA751281.

Author contribution The raw reads and sequencing data have been uploaded in the SRA under Bioproject number PRJNA751281. FC, CL and HH conceived the research plan. HH, FC and JJ did experiments and collected data. HH, JJ and KS analyzed the data. CL, KS, JK and HH wrote the draft. All authors edited and approved the final version of the manuscript.

References

- Bandelt HJ, Forster P, Rohlf A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Boguski DA, Gallagher CP, Howland KL, Harris LN (2016) Genetic stock identification and mixed-stock fishery analysis of Arctic Char (*Salvelinus alpinus*) in Darnley Bay, Northwest Territories. *DFO Can Sci Advis Sec Res Doc* 2015/023.v + 18 p
- Cheng FY, Wang Q, Delser PM, Li CH (2019) Multiple freshwater invasions of the tapertail anchovy (Clupeiformes: Engraulidae) of the Yangtze River. *Ecol Evol* 9:12202–12215. <https://doi.org/10.1002/ece3.5708>
- Dai P, Yan Y, Zhu XY, Tian JL, Ma FJ, Liu K (2020) Status of *Coilia nasus* resources in the National Aquatic Germplasm Resources Conservation Area in the Anqing Section of the Yangtze River. *J Fish Sci China* 27:1267–1276. <https://doi.org/10.3724/SP.J.1118.2020.20130>
- Depristo MA, Banks E, Poplin R, Garimella KV, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498. <https://doi.org/10.1038/ng.806>
- Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and

- implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- ESRI (2011) ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA
- Excoffier L, Lischer HEL (2010) 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>
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genet* 131: 479–491. <https://doi.org/10.1093/genetics/131.2.479>
- Gao TX, Wan ZZ, Song N, Zhang XM, Han ZQ (2014) Evolutionary mechanisms shaping the genetic population structure of coastal fish: insight from populations of *Coilia nasus* in Northwestern Pacific. *Mitochondr DNA* 25:464–472. <https://doi.org/10.3109/19401736.2013.814109>
- Grant WS, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *J Hered* 89:415–426. <https://doi.org/10.1093/jhered/89.5.415>
- Gwak WS, Lee YD, Nakayama K (2015) Population structure and sequence divergence in the mitochondrial DNA control region of gizzard shad *Konosirus punctatus* in Korea and Japan. *Ichthyol Res* 62: 379–385. <https://doi.org/10.1007/s10228-014-0450-7>
- He Z (2019) Save bigmouth grenadier anchovy. *World Environ*:61–63
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913. <https://doi.org/10.1038/35016000>
- Jiang T, Liu HB, Shen XQ, Shimasaki Y, Oshima Y, Yang J (2014a) Life history variations among different coast inferred from Otolith Microchemistry. *J Fac Agr Kyushu U* 59:383–389. <https://doi.org/10.5109/1467650>
- Jiang T, Yang J, Liu H, Shen XQ (2014b) Life history of *Coilia nasus* from the Yellow Sea inferred from otolith Sr:Ca ratios. *Environ Biol Fishes* 95:503–508. <https://doi.org/10.1007/s10641-012-0066-6>
- Jiang T, Liu HB, Lu MJ, Chen TT, Yang J (2016) A possible connectivity among Estuarine Tapertail Anchovy (*Coilia nasus*) populations in the Yangtze River, Yellow Sea, and Poyang Lake. *Estuar Coast* 39: 1762–1768. <https://doi.org/10.1007/s12237-016-0107-z>
- Jiang T, Liu HB, Li MM, Yang J (2018) Investigation on shrimp feeding of *Coilia nasus* during its anadromous migration along the Yangtze River. *J Lake Sci* 30:458–463. <https://doi.org/10.18307/2018.0217>. (in Chinese with English abstract)
- Jiang JM, Yuan H, Zheng X, Wang Q, Kuang T, Li JY, Liu JN, Song SL, Wang WC, Cheng FY, Li HJ, Huang JM, Li CH (2019) Gene markers for exon capture and phylogenomics in ray-finned fishes. *Ecol Evol* 9:3973–3983. <https://doi.org/10.1002/ece3.5026>
- Khumbanyiya DD, Li M, Tao J, Liu H, Jian Y (2018) Unraveling habitat use of *Coilia nasus* from Qiantang River of China by otolith microchemistry. *Reg Stud Mar Sci* 18:122–128. <https://doi.org/10.1016/j.rsma.2018.02.001>
- Kristjansson BK, Skulason S, Noakes DLG (2002) Rapid divergence in a recently isolated population of threespine stickleback (*Gasterosteus aculeatus* L.). *Evol Ecol Res* 4:659–672. <https://doi.org/10.1046/j.1525-142X.2002.02015.x>
- Lavin PA, Mcphail JD (1993) Parapatric lake and stream sticklebacks on Northern Vancouver Island—disjunct distribution or parallel evolution? *Can J Zool* 71:11–17. <https://doi.org/10.1139/z93-003>
- Lee JY, Edwards SV (2008) Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution* 62:3117–3134. <https://doi.org/10.1111/j.1558-5646.2008.00543.x>
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li C, Michael H, Nicolas S, Shannon C, Naylor G (2013) Capturing protein-coding genes across highly divergent species. *Biotechniques* 54:321–326. <https://doi.org/10.2144/000114039>
- Li XQ, Liu F, Leng CM, Dong GC, Wang YN, Shi JH (2015) Genetic structure and diffusion of population of *Coilia ectenes taihuensis* in Lake Nansi inferred from the mitochondrial control region. *J Lake Sci* 27: 686–692. <https://doi.org/10.18307/2015.0417>. (in Chinese with English abstract)
- Li L, Lin H, Tang W, Dong L, Bao B, Yang J (2017) Population genetic structure in wild and aquaculture populations of *Hemibarbus maculatus* inferred from microsatellites markers. *Aquac Fish* 002: 78–83. <https://doi.org/10.1016/j.aaf.2017.03.004>
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Lindsey CC, Northcote TG, Hartman GF (1959) Homing of rainbow trout to inlet and outlet Spawning streams at Loon Lake, British Columbia. *J Fish Res Board Canada* 16:695–719. <https://doi.org/10.1139/f59-051>
- Liu JX, Gao TX, Zhuang ZM, Jin XS, Yokogawa K, Zhang YP (2006) Late Pleistocene divergence and subsequent population expansion of two closely related fish species, Japanese anchovy (*Engraulis japonicus*) and Australian anchovy (*Engraulis australis*). *Mol Phylogenet Evol* 40:712–723. <https://doi.org/10.1016/j.ympev.2006.04.019>
- Liu HB, Jiang T, Yang J (2018) Unravelling habitat use of *Coilia nasus* from the Rokkaku River and Chikugo River estuaries of Japan by otolith strontium and calcium. *Acta Oceanol Sin* 37:52–60. <https://doi.org/10.1007/s13131-018-1190-8>
- Liu HB, Jiang T, Xuan ZY, Qiu C, Yang J (2020) Otolith microchemical analysis of Tapertail Anchovy *Coilia nasus* from Ariake Sea and its adjacent tributaries in Japan. *Fish Sci* 39(04):500–508. <https://doi.org/10.16378/j.cnki.1003-1111.2020.04.005>. (in Chinese with English abstract)
- Ma CY, Cheng QQ, Zhang QY, Zhuang P, Zhao YL (2010) Genetic variation of *Coilia ectenes* (Clupeiformes: Engraulidae) revealed by the complete cytochrome b sequences of mitochondrial DNA. *J Exp Mar Biol Ecol* 385:14–19. <https://doi.org/10.1016/j.jembe.2010.01.015>
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220. <https://doi.org/10.1007/s00253-002-1013-9>
- Mao CZ, Jiao XM, Zhong JS, Hua WH, Zhang XY, Wu JX (2015) Research Progress on Resource Status and Protection of *Coilia nasus* in Yangtze River Estuary. *J Jiangsu Ocean Univ(Nat Sci)* 78-83. <https://doi.org/10.3969/j.issn.1672-6685.2015.03.018>. (in Chinese with English abstract)
- Masson S, Lepais O, Manicki A, Prevost E, Chat J (2018) Disentangling individual movement between populations from effective dispersal in the facultative anadromous *Salmo trutta* L. *Ecol Freshw Fish* 27: 323–338. <https://doi.org/10.1111/eff.12349>
- Ni Y, Wu HL (2006) Fishes of Jiangsu Province. China Agriculture Press Co., Ltd., Beijing
- Presti FT, Wasko AP (2014) A review of microsatellite markers and their application on genetic diversity studies in parrots. *Open J Genet* 4: 69–77. <https://doi.org/10.4236/ojgen.2014.42010>
- Pritchard JK, Stephens MJ, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. *Genet* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Simanjuntak C, Kinoshita I, Fujita S, Takeuchi K (2015) Reproduction of the endemic engraulid, *Coilia nasus*, in freshwaters inside a reclamation dike of Ariake Bay, western Japan. *Ichthyol Res* 62:374–378. <https://doi.org/10.1007/s10228-014-0448-1>
- Smouse P, Long J, Sokal R (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* 35:627–632. <https://doi.org/10.2307/2413122>

- Song N, Zhang XM, Sun XF, Yanagimoto T, Gao TX (2010) Population genetic structure and larval dispersal potential of spottedtail goby *Synechogobius ommaturus* in the north-west Pacific. *J Fish Biol* 77: 388–402. <https://doi.org/10.1111/j.1095-8649.2010.02694.x>
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Sun C, Xuan ZY, Liu HB, Jiang T, Yang J (2019) Cyt b gene and D-loop sequence analyses of *Coilia nasus* from the Rokkaku River of Japan. *Reg Stud Mar Sci* 32. ARTN 100840
- Takita T (1967) The spawning and the early life history of the Engraulid fish *Coilia* sp. distributed in Ariake Sound. *Bull Fac Fish Nagasaki Univ* 23:107–122
- Tallman RF, Healey MC (1994) Homing, Straying, and Gene Flow among Seasonally Separated Populations of Chum Salmon (*Oncorhynchus keta*). *J Can Des Sci Hal Et Aqa* 51:577–588. <https://doi.org/10.1139/f94-060>
- Tang WQ, Xue LH, Yang JQ (2007) Species validities of *Coilia brachygnathus* and *C. nasus taihuensis* based on sequence variations of complete mtDNA control region. *Biodiver Sci* 224–231. <https://doi.org/10.1360/biodiv.060263>
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Pro Bioinform* 43(1110):11.10.1–11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>
- Wang Q, Cheng FY, Li CH (2017) Developing SNP markers for *Coilia* and its application in species delimitation. *J Shanghai Ocean Univ* 26(01):8–16. <https://doi.org/10.12024/jsou.20160401724> (in Chinese with English abstract)
- Ward RD, Woodwark M, Skibinski D (2010) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J Fish Biol* 44:213–232. <https://doi.org/10.1111/j.1095-8649.1994.tb01200.x>
- Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population structure. *Evolution* 38:1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
- Whitehead PJP, Nelson G, Wongratana, T (1988) FAO species catalogue. Vol. 7. Clupeoid fishes of the world (Suborder Clupeoidei). An annotated and illustrated catalogue of the herrings, pilchards, sprats, shads, anchovies and wolf-herrings. Part 2 - Engraulidae. FAO Fisheries Synopsis, No. 125 (7 Part 2)
- Xuan ZY, Jiang T, Liu HB, Chen XB, Yang J (2021) Mitochondrial DNA and microsatellite analyses reveal strong genetic differentiation between two types of estuarine tapertail anchovies (*Coilia*) in Yangtze River Basin, China. *Hydrobiologia* 848:1409–1431. <https://doi.org/10.1007/s10750-021-04541-w>
- Xue DX, Yang QL, Li YL, Zong SB, Gao TX, Liu JX (2019) Comprehensive assessment of population genetic structure of the overexploited Japanese grenadier anchovy (*Coilia nasus*): Implications for fisheries management and conservation. *Fish Res* 213:113–120. <https://doi.org/10.1016/j.fishres.2019.01.012>
- Xue DX, Yang QL, Zong SB, Gao TX, Liu JX (2020) Genetic variation within and among range-wide populations of three ecotypes of the Japanese grenadier anchovy *Coilia nasus* with implications to its conservation and management. *J Oceanol Limnol* 38:851–861. <https://doi.org/10.1007/s00343-019-9091-z>
- Yang QL, Han ZQ, Sun DR, Xie SG, Lin LS, Gao TX (2010) Genetics and phylogeny of genus *Coilia* in China based on AFLP markers. *Chin J Oceanol Limn* 28:795–801. <https://doi.org/10.1007/s00343-010-9093-3>
- Yang QL, Gao TX, Miao ZQ (2011) Differentiation between populations of Japanese grenadier anchovy (*Coilia nasus*) in Northwestern Pacific based on ISSR markers: Implications for biogeography. *Biochem Syst Ecol* 39:286–296. <https://doi.org/10.1016/j.bse.2011.07.019>
- Yang JQ, Hsu KC, Zhou XD, Kuo PH, Lin HD, Liu D, Bao BL, Tang WQ (2016) New insights on geographical/ecological populations within *Coilia nasus* (Clupeiformes: Engraulidae) based on mitochondrial DNA and microsatellites. *Mitochondrial Dna A* 1–7. <https://doi.org/10.1080/24701394.2016.1261850>
- Yang JQ, Hsu KC, Zhou XD, Kuo PH, Lin HD, Liu D, Bao BL, Tang WQ (2018) New insights on geographical/ecological populations within *Coilia nasus* (Clupeiformes: Engraulidae) based on mitochondrial DNA and microsatellites. *Mitochondrial DNA A* 29: 158–164. <https://doi.org/10.1080/24701394.2016.1261850>
- Yuan CM, Lin JB, Qin AL, Liu RH (1976) The history and current situation of Chinese classification of *Coilia* sp. *J Nanjing Univ* (in Chinese with English abstract)
- Yuan CM, Qin AL, Liu RH, Lin JB (1980) On the classification of the anchovies, *Coilia*, from the lower Yangtze River and the southeast coast of China. *J Nanjing Univ(Nat. Sci)*. (in Chinese with English abstract)
- Yuan H, Atta C, Tornabene L, Li CH (2019) Assexon: assembling exon using gene capture data. *Evol Bioinform* 15. <https://doi.org/10.1177/1176934319874792>
- Zhang SS, Luo J, Chen BJ, Cheng ZL, Sun PF, Jiang XT, Dong AR, Jiang ZQ, Wang W (2020) A study on morphological differences of *Coilia nasus* from different populations. *Freshw Fish* 50:26–33. <https://doi.org/10.13721/j.cnki.dsyy.2020.03.004> (in Chinese with English abstract)

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