Ocean Sci. J. (2019) 54(2):245–256
http://dx.doi.org/10.1007/s12601-019-0013-y

Available online at http://link.springer.com

SpringerLink

Article

Ocean Science Journal

pISSN 1738-5261
eISSN 2005-7172

Population Genetic Structure of *Liza affinis* (Eastern Keelback Mullet), Reveals High Gene Flow Inferred from Microsatellite Analysis

Lu Liu[1], Xiumei Zhang[2], Dianrong Sun[3], Tianxiang Gao[2,4], and Na Song[1]*

[1]*The Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 276000, China*
[2]*National Engineering Research Center for Marine Aquaculture, Zhejiang Ocean University, Zhoushan 316021, China*
[3]*South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510300, China*
[4]*Key Laboratory for Technology Research on Sustainable Utilization of Marine Fishery Resources, Zhoushan 316021, China*

Received 10 August 2018; Revised 31 October 2018; Accepted 13 December 2018

Abstract – The population genetic studies of marine fishes usually show complex patterns of genetic differentiation which were influenced by both historical process and contemporary gene flow. Genetic structure of eight different populations for eastern keelback mullet, *Liza affinis*, collected from the coast along eastern and southern China, was examined using six microsatellite loci. We found strong genetic similarities among most of the samples except the Dongguan population and most microsatellite variation was found within populations. All loci were characterized by high genetic variability with expected heterozygosities ranging from 0.871 to 0.928. The Bayesian cluster analysis of the microsatellite data detected four genetic groups with no relation to geographic areas. The bottleneck results also showed no significant values. Based on these data we postulated that complex marine currents and larval dispersal shaped the genetic structured of studied populations. The present study illustrated the importance of understanding the biological significance of genetic differentiation when using molecular data in identifying units for management and protection.

Keywords – *Liza affinis*, microsatellite DNA, genetic differentiation, population genetic structure

1. Introduction

Mugil fish are widely distributed throughout tropical and temperate seas, and many species are commercially important fishery resources throughout the world (Nash and Shehadeh 1980). The eastern keelback mullet, *Liza affinis* (Güther, 1866), which belongs to Mugilidae, is a small species which is common in the coastal waters of Japan (except for the northern part of Hollaido), and also along the coasts of China, from the south Yellow Sea to the South China Sea (Lee et al. 1994; Liu et al. 2017). *L. affinis* is also one of the important components of commercially important species in China and Japan for fishery and mariculture (Li et al. 1993; Yoshimatsu et al. 1993). A few studies have been carried out fragmentarily on the early life history of this species, such as, the embryonic development and the early larval stages from reared specimens (Zhang et al. 1985), morphology research about the eggs and early-larvae of *L. affinis* from field-collected specimens (Joh et al. 1988; Ikeda and Mito 1988), and some researches about larvae and juveniles (Nakamura 1936; Mito 1966; Kinoshita 1988). Other research featured a series of reared specimens to describe early development, growth, and morphological changes (Yang and Qiu 1989; Yoshimatsu et al. 1993). Besides the studies mentioned above, there has been research on diet ingestion, absorption, utilization and the complete mitochondrial DNA sequence analysis of *L. affinis* (Li et al. 1993; Gong et al. 2016).

Knowledge of geographic structuring is a critical component in the design of informed management strategies (Selkoe et al. 2016). Besides pelagic eggs and larvae, large population sizes and wide distribution are considered to be among several general characteristics of typical marine fishes. Marine fishes commonly showed weak or even no population structure due to the lack of natural physical barriers in the oceans (Gyllensten 1985; Palumbi 1994; Ward et al. 1994; Waples 1998; Minegishi et al. 2008; Kim et al. 2010a; Gao et al. 2016). However, significant genetic structuring was still detected in some marine

*Corresponding author. E-mail: songna624@163.com

fishes. For example, *Mugil liza*, one species of mullet, is distributed along the Atlantic coast of South America, and some researchers discerned patterns of distribution of neutral nuclear genetic variation in 250 samples from the Brazilian states of Rio de Janeiro, Sao Paulo, Santa Catarina and Rio Grande do (Mai et al. 2014). Two different demographic clusters were detected in mullets from the southernmost portion of the Atlantic coast of Brazil and from Argentina. Mullet larvae are passively carried out by oceanic currents, which, in turn, may promote long distance dispersal and gene flow over a wide geographic scale. This factor was thought to be responsible for the significant differentiation observed in one study (Whitfield et al. 2012). Furthermore, the existence of specific oceanographic conditions regarding spawning and marine boundaries that could provide efficient barriers to the dispersal of larvae, and restrict gene flow, has been proposed to explain genetic structuring in other Mugilids (Rocha-Olivares et al. 2000, 2017; Liu et al. 2009; Kim et al. 2010b; Durand et al. 2013).

Microsatellite DNA is one of the useful molecular markers which can be used in uncovering population structure. It has been proven to be more useful than mitochondrial sequences, which were not apparent using less variable markers (Nielsen et al. 2001; Carlsson et al. 2004; Cheng et al. 2015). Microsatellites are characterized as biparental hypervariable markers, and it was considered to be more helpful than other molecular markers for investigating population genetic structure (Toews and Brelaford 2012). Additionally, microsatellite markers, which are influenced by geographical factors or environmental factors, are more sensitive in testing the existence of reductions in gene flow currently than mitochondrial sequences (Novembre et al. 2008). In recent years, many studies have shown that microsatellite markers had been widely used to infer population genetic differentiation of fishes and other organisms at different geographical scales (Zhang and Hewitt 2003). However, research on *Liza affinis* based on microsatellite markers is still deficient. The control region of mitochondrial DNA was employed to examine the phylogeographic pattern and dispersal ability of *L. affinis* by Liu et al. (2018), and the result showed that there exsited high genetic diversity and homogeneity of *L. affinis* across the East China Sea and the South China Sea. Egg and larval dispersal may be the key factor for frequent gene flow and may prevent population divergence among populations. It is also interesting to see that genetic differentiation between population DG and some populations were large, but gene flow calculation suggested that there still exist gene exchanges between them. Therefore, we were curious whether or not more sensitive markers could be employed to make more accurate inferences. Is it really the case that the larvae of *L. affinis* create more potential for long distance dispersal and increase gene flow among different geographic populations?.

The purpose of this paper is to assess genetic variation and population structure of *Liza affinis* along the Pacific coast of eastern and southern China using six microsatellite loci. Through this study, we can also reveal the pattern and degree of differentiation among populations, and detect the relative factors of biological characteristics and environmental influences involved in the formation of the contemporary population genetic structure of *Liza affinis*. As *Liza affinis* occupy a relatively low position in the food web, this research may relate to their food and feeding habits and this can provide guidance for biodiversity conservation (Liu et al. 2018).

2. Materials and Methods

Sampling

Liza affinis was sampled from eight localities in the Pacific coast of eastern and southern China between November 2013 and November 2015 (Table 1; Fig. 1). Taxonomic status of the fishes was identified morphologically (Liu et al. 2016a). The muscle tissue was excised from each individual and preserved in 95% ethanol for DNA extraction.

Microsatellite genotyping

Genomic DNA was isolated from the muscle tissue by proteinase K digestion followed by the standard phenol/chloroform method (Sambrook and Russell 2001). Six microsatellite loci were amplified using primers La83, La139, La140, La195, La200, La217 developed from *Liza affinis* (Liu et al. 2016b) and 5'of the primers were fluorescelin-labeled. All the primers were synthesized by Invitrogen Corp. The information for the six microsatellite loci and primers is shown in Table 2. The PCR reactions were carried out in 25 μL reaction mixture, containing 17.25 μL ultrapure water, 2.5 μL 10×PCR buffer, 2 μL dNTPs, 1 μL of each primer (5 μM), 0.25 μL Taq polymerase, and 1 μL of DNA template. The thermal regime consisted of an initial step of 2 min at 95°C followed by 40 cycles of 30 s at 94°C, 45 s at annealing temperature (*Ta*) (Table 2), and 1 min at 72°C; with a final step of 10 min at 72°C, after which the reaction was held at 7°C. Negative controls were included in all PCR reactions to

Table 1. Sampling sites, date of collection and sample size of each population of *Liza affinis*

ID	Sampling sites	Sample size	Date of collection
ND	Ningde, Fujian Province	24	2014.10
JJ	Jinjiang, Fujian Province	24	2013.11
JM	Jinmen, Fujian Province	24	2013.10
XM	Xiamen, Fujian Province	24	2014.02
DG	Dongguan, Guangdong Province	24	2014.11
ZJ	Zhanjiang, Guangdong Province	24	2014.12
BH	Beihai, Guangxi Province	24	2014.05/2014.12
FCG	Fangchenggang, Guangxi Province	24	2015.11
	Total	192	-

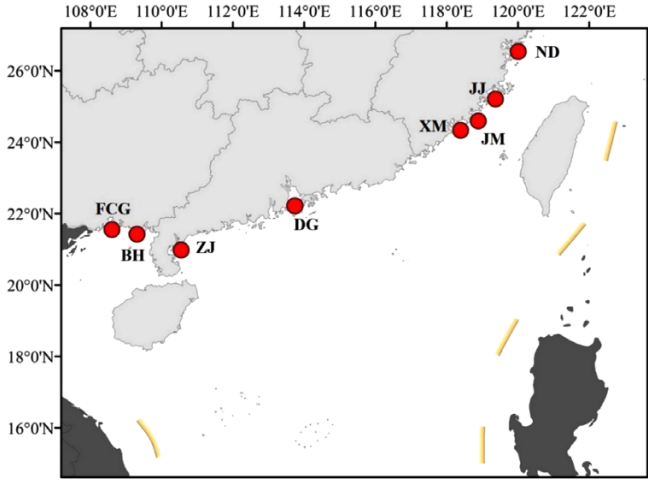

Fig. 1. Sampling sites of *Liza affinis* in the present study.

confirm the absence of contaminants. All PCR were performed in an Eppendorf thermal cycler following optimized reaction conditions. Electrophoresis and Genotyping were carried out by ABI3730xl and GENEMARKER software version 2.2.0 (SoftGenetics, State College, PA, USA) using com-LIZ ™500 as size standard in the Thermo Fisher Scientific Company.

Statistical analysis

Microsatellite genetic diversity was qualified as the number of unique alleles (A), observed heterozygosity (H_O) and expected heterozygosity (H_E) for each locus and sample site using Excel Microsatellite Toolkit (Park 2001). Polymorphism information content (PIC) was calculated using allele frequencies according to the formula given by Botstein et al (Botstein et al. 1980). The allelic richness (R_S), which is a standardized index of the mean number of alleles per locus irrespective of sample size, was calculated by Fstat 2.9.3 (Goudet et al. 2001). Deviations from the Hardy–Weinberg equilibrium (HWE) expectation and linkage disequilibrium were performed for each site at each locus using GENEPOP 3.4 (Raymond and Rousset 1995) with 10,000 burn-in steps, and 500 batches of 5,000 Monte Carlo Markov Chain (MCMC) steps per batch. The software MICRO-CHECKER 2.2.0 (Van et al. 2004) was used to test for technical artefacts such

Table 2. Information for six microsatellite loci and primers in microsatellite analysis of *Liza affinis* populations

Locus	Primer sequence (5'-3')	Fluorescent dye	Repeat motif	T_a (°C)	Allele size range (bp)	N_A	H_O	H_E	GenBank accession
La83	F: CTGTCCCCTGCTGACCTAA R: GTTACCAAACATTAAAAGCCAT	FAM	(TG)15	48	254–380	6	0.999	0.803	KT355832
La139-1	F: ATTGTTCATACCGTAGTGC R: GATTAAGCCAAGAAGAGG	FAM	(CTCCT)5	51	201–242	7	0.500	0.793	KT355836
La139-3	F: CTTTATGTCTCCTCCAGCGTT R: TGTGCCTTTTCTAATGTCAGC	TAM	(CTCCT)5···(GT)5	60	210–240	5	0.417	0.794	KT355837
La195-1	F: ACTATTCCCGAGGAGTCT R: TTCCAGAATTTTGTCTTAAC	TAM	(CA)5···(CA)17···(CA)5	48	190–250	7	0.875	0.824	KT355845
La200-3	F: TCCCTGACAGAATACCAAC R: GGTGTAGTAGAGCGAGGC	HEX	(GT)11	50	240–270	7	0.958	0.785	KT355846
La217-3	F: ATCAGTCATTATTGCCTTTT R: ACACGCATGAGTAAACTACAT	HEX	(AC)6···(TG)8	48	250–300	8	0.870	0.854	KT355849

as null alleles, stuttering and large allele dropout. Critical significance levels for multiple simultaneous tests were adjusted using the sequential Bonferroni correction (Rice 1989). Pairwise genetic divergence values between populations were estimated using F_{ST} values for microsatellite data and with Arlequin 3.5 (Excoffier and Lischer 2010). The output of a list of genetic distances on 8 populations was calculated by the software package POPULATION1.2 (Oliver 2002) based on UPGMA algorithm method. In addition, we conducted a hierarchical analysis of molecular variance (AMOVA) in ARLEQUIN 3.5 (Excoffier and Lischer 2010) to investigate genetic differentiation across the total sample range and among putative regional grouping of samples.

Isolation-by-distance (IBD) analyses were conducted for microsatellite data. Shoreline distances between sampled populations were estimated in km using Google Earth version 4.4 and plotted against genetic distance, pairwise $F_{ST}/(1-F_{ST})$. IBD regression analysis was performed online using the IBD web service (http://ibdws.sdsu.edu/*ibdws/) (Bohonak 2002; Jensen et al. 2005) with 10,000 randomizations of the data.

To investigate recent genetic bottlenecks, the Wilcoxon sign-rank test for hetretrozygosity excess was applied under three different models, namely, infinite alleles model (IAM), two-phase model (TPM) and stepwise mutation model (SMM), using the program Bottleneck 1.2.02 (Piry et al. 1999). Furthermore, the qualitative test of model shift was performed to calculate the allele frequency distribution using Bottleneck 1.2.02 (Piry et al. 1999).

Three dimensional factorial correspondence analysis (3D-FCA) was performed in GENETIX 4.05 (Belkhir et al. 2002) to explore population divisions and relationships of *Liza affinis*, independent from prior knowledge of their relationships. To identify clusters of genetically similar populations, a model-based Bayesian clustering algorithm was applied by using STRUCTURE 2.1 (Pritchard et al. 2000), five replicates were run for all possible values of the maximum number of clusters (K) up to K = 8 (total sites). For each run, 1,000,000 iterations were carried out after a burn-in period of 100,000 iterations. We used the Structure Harvester website (Earl 2012) to determine the number of genetic discrete populations (K) that best fit the data, which implemented the Evanno method (Evanno et al. 2005). The simulated K values ranged from 2 to 8. Five independent runs were carried out for each specific K-value in order to verify the consistency of the results. The simulations were conducted assuming an admixture model with correlated allele frequencies, which was considered as the superior model for detecting structure among closely related populations (Falush et al. 2003). MCMC consisted of 100,000 burn-in iterations followed by 1,000,000 iterations. For all calculations, significance was assessed by 1,000 random permutations and P values from multiple comparisons were Bonferroni adjusted (Rice 1989).

3. Results

Summary statistics for six microsatellite loci in the eight populations from the eastern and southern coast of China are shown in Table 3. A total of 152 alleles were detected across all six loci, with 17 alleles (La200-3) to 34 alleles (La195-1). The total allelic richness (*R*s) per locus ranged from 11.7 (La200-3) to 21.1 (La195-1). Average *R*s value was highest in BH and ND (18.7) and lowest in JJ (15.2). The observed and excepted heterozygosities per population varied from 0.5903 (DG) to 0.8056 (BH), and from 0.8466 (JJ) to 0.9279 (ND), respectively. No linkage disequilibrium was detected for six loci in the eight populations. There were 16 of 48 locus-by-population tests that exhibited significant departure from HWE, but none of them remained significant after correction for multiple tests. Analysis using MICRO-CHECKER suggested the presence of null alleles in FCG population at locus La83 and La217-3.

The values of pairwise F_{ST} were calculated in the program *F*stat and the results indicated weak gene differentiation among *L. affinis* populations (F_{ST} = -0.0153~0.0180) (Table 4, below diagonal). All the numerical values were generally low, even negative in some populations. Most of them were not significant after Bonferroni correction except that the DG population was significantly differentiated from the JM, JJ, XM, and ND populations. According to the Mantel test, no significant correlation ($R^2 = 0.209$, $P > 0.05$) was observed between genetic distance determined as $F_{ST}/(1-F_{ST})$ and geographical distance based on all loci (Fig. 2). The genetic divergence between *L. affinis* based on Nei's genetic distance is shown in Table 4 above diagonal divide. The results showed that there were no obvious cluster pattern of cultivars.

The hypothetical grouping of populations was further explored by an analysis of molecular variance (AMOVA) tests (Table 5). Based on the result of population genetic distances, we used four hypotheses for grouping the 8 populations. DG was singled out alone because of it was significantly differentiated from other populations based on the values of pairwise F_{ST}.

Table 3. Summary statistics for the variability of 6 polymorphic microsatellite loci in 8 *Liza affinis* populations

Location ID	Parameters	Locus						Average
		La139-1	La195-1	La217-3	La83	La139-3	La200-3	
ZJ	A	12	19	20	17	20	9	16.2
	*R*s	11.7	18.9	19.5	16.8	19.3	8.7	15.8
	H_O	0.5000	0.7500	0.7917	0.7083	0.6667	0.3333	0.6250
	H_E	0.8768	0.8520	0.9441	0.9415	0.9291	0.6809	0.8707
	PIC	0.8449	0.8265	0.9196	0.9168	0.9033	0.6422	0.8422
JJ	A	10	22	17	19	10	13	15.2
	*R*s	9.9	21.4	16.6	19.0	9.8	12.3	14.8
	H_O	0.6667	0.8333	0.6667	0.7083	0.5417	0.6250	0.6736
	H_E	0.8511	0.8635	0.8688	0.9193	0.7863	0.7908	0.8466
	PIC	0.8162	0.8375	0.8410	0.8942	0.7480	0.7548	0.8153
JM	A	16	21	17	16	18	14	17.0
	*R*s	15.6	20.2	16.2	15.6	17.5	13.6	16.45
	H_O	0.5833	0.8750	0.7917	0.5833	0.7083	0.6250	0.6944
	H_E	0.9025	0.9122	0.9167	0.8901	0.9184	0.8528	0.8988
	PIC	0.8737	0.8873	0.8894	0.8509	0.8909	0.8210	0.8704
XM	A	16	21	21	19	16	14	17.8
	*R*s	15.3	19.9	20.1	18.6	15.6	13.1	17.1
	H_O	0.7083	0.8333	0.9583	0.5417	0.7500	0.6667	0.7431
	H_E	0.8918	0.8635	0.9504	0.8989	0.9069	0.8440	0.8926
	PIC	0.8634	0.8375	0.9263	0.8732	0.8789	0.8100	0.8649
DG	A	18	18	16	17	17	9	15.8
	*R*s	17.7	17.2	15.4	16.1	16.7	8.2	15.2
	H_O	0.5000	0.9583	0.5417	0.5417	0.5833	0.4167	0.5903
	H_E	0.9344	0.9220	0.9051	0.8954	0.9043	0.7030	0.8774
	PIC	0.9087	0.8963	0.8770	0.8665	0.8755	0.6653	0.8482
BH	A	17	24	19	20	22	10	18.7
	*R*s	16.3	23.1	19.0	19.4	21.6	10.0	18.2
	H_O	0.7500	0.9999	0.9167	0.8333	0.7500	0.5833	0.8056
	H_E	0.9406	0.9548	0.9397	0.9309	0.9371	0.8342	0.9229
	PIC	0.9156	0.9314	0.9146	0.9052	0.9121	0.7958	0.8958
ND	A	16	22	21	22	18	13	18.7
	*R*s	15.5	21.7	20.4	21.6	17.3	12.4	18.2
	H_O	0.7917	0.9583	0.9167	0.7083	0.7917	0.6250	0.7986
	H_E	0.9317	0.9335	0.9548	0.9441	0.9291	0.8741	0.9279
	PIC	0.9058	0.9095	0.9311	0.9202	0.9032	0.8409	0.9018
FCG	A	19	22	18	14	19	12	17.3
	*R*s	18.5	21.6	17.9	13.6	18.6	11.9	17.0
	H_O	0.7083	0.9583	0.9167	0.7917	0.7083	0.5417	0.7708
	H_E	0.9441	0.9512	0.9397	0.8998	0.9255	0.8592	0.9198
	PIC	0.9196	0.9274	0.9147	0.8700	0.8997	0.8226	0.8923

A: allelic number, R_S: allelic richness, H_O: observed heterozygosity, H_E: expected heterozygosity, PIC: polymorphism information conten

The other three classification methods were selected based on the different distribution of sea areas. It revealed that the majority of variance was attributable to differences within sampled populations. In addition, only a small amount of difference was explained by the differences among populations. The hierarchical AMOVA, a measure of population variation based on F_{ST} analysis, showed a genetic variance of 0.75% (P = 0.249) among the groups, while the intraspecies

Table 4. Pair-wise F_{ST} (below diagonal) and the shared allele distance based on UPGMA algorithm method (above diagonal divide) among populations of *L. affinis*

	ZJ	JJ	JM	XM	DG	BH	ND	FCG
ZJ		0.1715	0.1978	0.1785	0.1894	0.1144	0.1404	0.1873
JJ	0.0043		0.2440	0.1773	0.2683	0.1920	0.1584	0.2354
JM	-0.0029	0.0108		0.1990	0.2062	0.2022	0.1761	0.2220
XM	-0.0035	0.0033	-0.0021		0.2439	0.1625	0.1170	0.1819
DG	-0.0030	0.0165*	0.0031 *	0.0112*		0.2222	0.1977	0.2365
BH	-0.0018	0.0101	0.0014	-0.0019	0.0082		0.1399	0.0909
ND	-0.0009	0.0053	-0.0044	-0.0086	0.0028*	-0.0096		0.1534
FCG	0.0112	0.0180	0.0085	0.0070	0.0128	-0.0153	-0.0060	

*, significant at $P < 0.0018$ by the permutation test

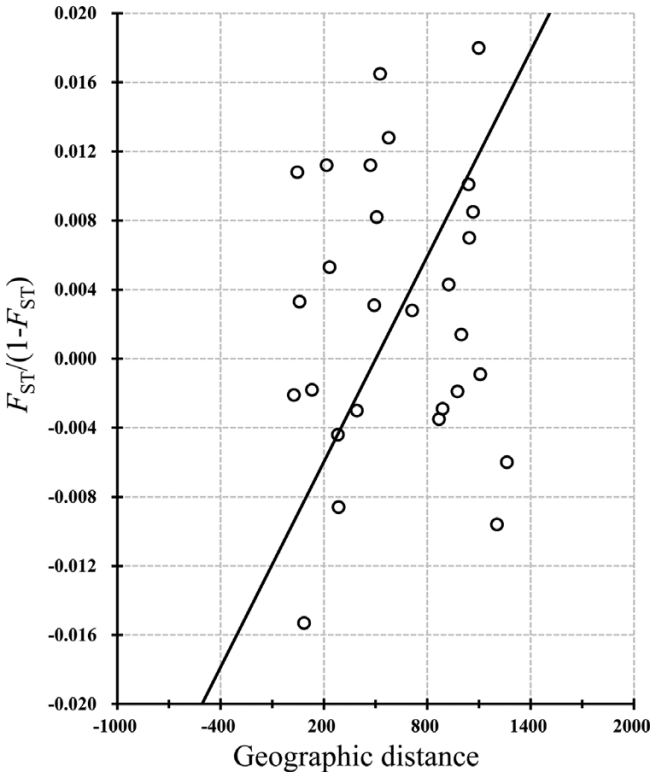

Fig. 2. Plot of pairwise estimates of $F_{ST}/(1-F_{ST})$ and geographic distance between 8 *L. affinis* populations

genetic variation was 98.74% with a difference of 0.51% among the populations. When the AMOVA was performed without considering population disjunction, 0.70% of the genetic variation was assigned in the population and 99.3% (P = 0.097) was detected among individuals in the population.

The result of three-dimensional factorial correspondence analysis (3D-FCA) explained 47.08% of the overall variation, and the eight *L. affinis* populations in the present study showed that there was no obvious clustering (Fig. 3), and that the spatial distribution was more dispersed. Each population was quite widely separated from each other. The Bayesian cluster analysis performed in the program STRUCTURE indicated that the model with K = 4 resulted in the highest ΔK value and can explain the data in a satisfactory manner (Figs. 4, 5). As shown in Table 5, nearly each individual *L. affinis* was assigned to each of the four main genetic clusters (where K was set to 4, respectively) in roughly equal proportions (i.e., apparently at random). Individuals from all sample areas were therefore pooled for follow-up analysis (Table 6).

The Wilcoxon test also detected no excess of heterozygotes assuming the same mutation model. Changing the mutation model assumed in the test had no effect on the results ($P > 0.05$), which indicated *L. affinis* should be in mutation-drift equilibrium with no signature of recent population bottleneck (Table 7). The microsatellite allele frequency distribution was in a normal L-shape ('mode-shift' indicator), suggesting that *L. affinis* resembled a stable population.

4. Discussion

Genetic data serve as essential elements for understanding the shaping progresses of genetic structure as well as assisting in devising conservation and management strategies regarding marine species (Diáz-Ferguson et al. 2010; Avise 2000). The purpose of our research was to expand our understanding of the genetic population structure regarding *Liza affinis*, an important fish resource along the Pacific coast of eastern and southern China.

Genetic diversity

The six microsatellite loci examined in this study showed high genetic diversity. The ranges of the observed and excepted heterozygosities in the overall samples at each locus ranged from 0.5903 to 0.8056 and from 0.6809 to 0.8741. The PIC

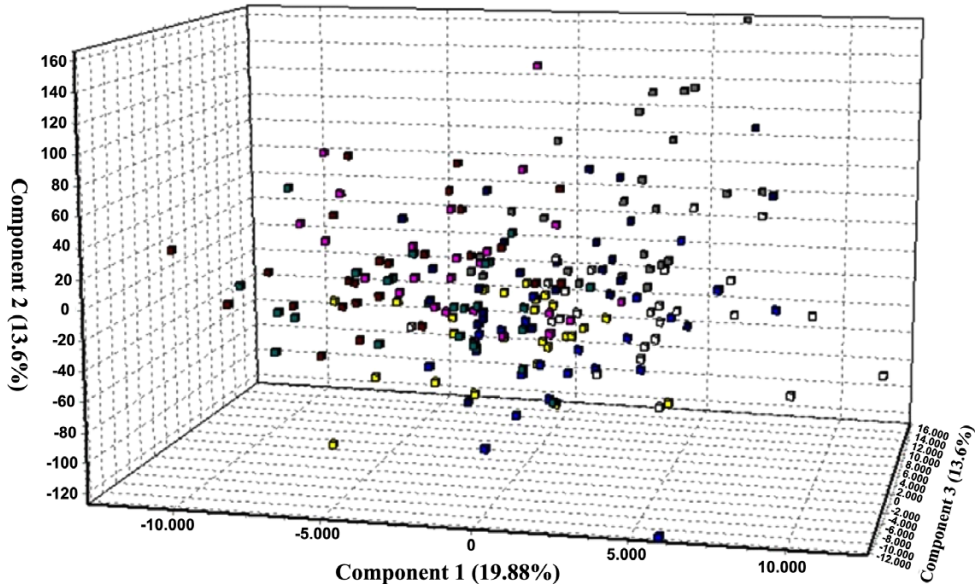

Fig. 3. Three dimensional factorial correspondence analysis (3D-FCA) showing relationships among *L. affinis* popualtions and individuals based on six microsatellite loci

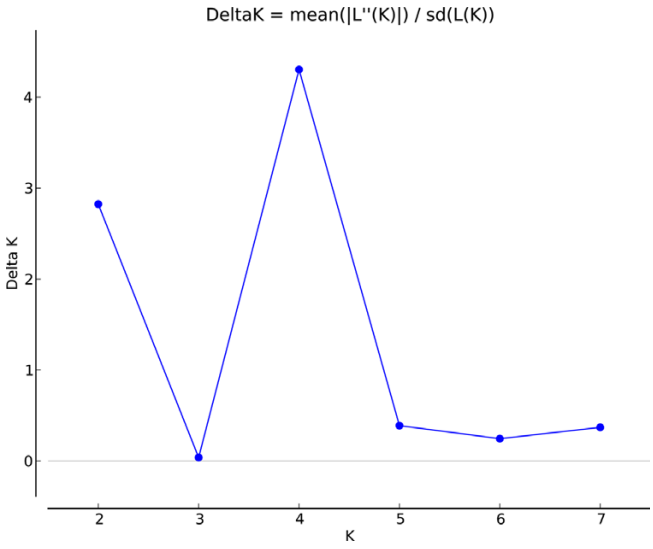

Fig. 4. Scatter plot of possible number of clusters (K; horizontal axis) against ad hoc statistic ΔK (vertical axis) based on rate of change in logarithm probability of date between successive K values

index was considered to be helpful in measuring the usefulness of molecular markers (Botstein et al. 1980). Individual uniformity of genetic variation for a population can also be reflected by the PIC index (Liao et al. 2006). Botstein et al. (1980) thought that PIC > 0.5 meant a high polymorphism level, 0.25 > PIC > 0 meant a moderate polymorphism level and PIC < 0.25 meant a low polymorphism level. Our results showed that the PIC indexes of all the loci we detected were more than 0.5, which means a high polymorphism level. The microsatellite markers have codominant expression and multiallelism features, so exhibition of high PIC was endemic. Meanwhile, a quite rich variation was shown by all the loci, which was suitable for use in analyzing the genetic diversity of *L. affinis* as microsatellite DNA markers. This variation also indicated that due to differences in origin, ecotype and species formation, the genotypes used in this study were more diverse (Ferreira and Grattapaglia 1998). There were no geographical trends detected based on the levels of microsatellite diversity.

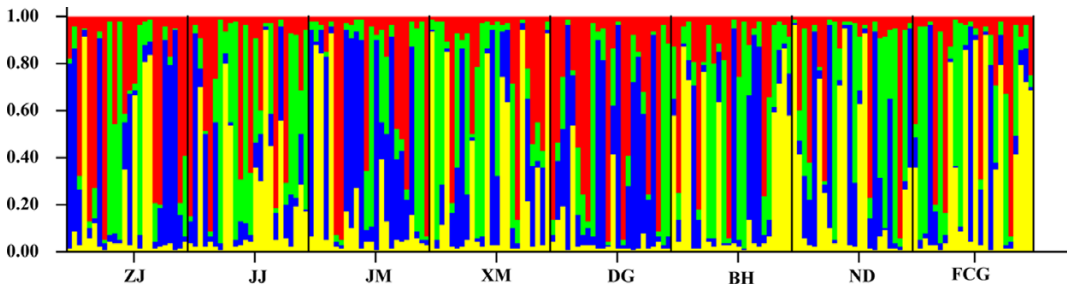

Fig. 5. Graphical results of the STRUCTURE analysis of six microsatellite loci in *L.affinis* populations. Vertical lines are proportional to the probability of individual membership in simulated cluster

Table 5. AMOVA of *L. affinis* populations

Structure tested	Observed partition			
	Variance	% total	Φ Statistics	P
1. *One gene pool* (ND, JJ, JM, XM, DG, ZJ, BH, FCG)				
Among populations	0.0190	0.70	Φ_{ST} = 0.0070	0.0972 ± 0.0032
Within populations	2.6837	99.30		
2. *Two gene pool* (ND, JJ, JM, XM, DG, ZJ) (BH, FCG)				
Among groups	0.0216	0.80	Φ_{CT} = 0.0080	0.0394 ± 0.0020
Among populations within groups	0.0100	0.36	Φ_{SC} = 0.0036	0.6458 ± 0.0049
Within populations	2.6837	98.84	Φ_{ST} = 0.0116	0.0873 ± 0.0026
3. *Two gene pool* (ND, JJ, JM, XM, ZJ, BH, FCG) (DG)				
Among groups	0.0203	0.75	Φ_{CT} = -0.0075	0.2492 ± 0.0057
Among populations within groups	0.0140	0.51	Φ_{SC} = 0.0052	0.2953 ± 0.0048
Within populations	2.6837	98.74	Φ_{ST} = 0.0126	0.0923 ± 0.0028
4. *Three gene pool* (ND, JJ, JM, XM) (DG, ZJ) (BH, FCG)				
Among groups	0.0228	0.84	Φ_{CT} = -0.0084	0.1050 ± 0.0027
Among populations within groups	0.0028	0.10	Φ_{SC} = 0.0010	0.0000 ± 0.0000
Within populations	2.6837	99.06	Φ_{ST} = 0.0094	0.0000 ± 0.0000

Table 6. Proportion of eight populations in each of the four inferred clusters

Population	Inferred cluster				Number of individuals
	1	2	3	4	
ZJ	0.181	0.188	0.352	0.279	24
JJ	0.314	0.250	0.272	0.163	24
JM	0.185	0.191	0.239	0.384	24
XM	0.300	0.312	0.194	0.194	24
DG	0.175	0.073	0.408	0.344	24
BH	0.284	0.297	0.212	0.207	24
ND	0.253	0.288	0.199	0.259	24
FCG	0.261	0.360	0.210	0.169	24

Table 7. Results of Wilcoxon's heterozygosity excess test, mode shift indicator for a genetic bottleneck of *L. affinis*

Population	Wilcoxon sign-rank test			Mode shift[b]
	IAM	TPM	SMM	
ZJ	0.578	0.977	0.984	L
JJ	0.977	1.000	1.000	L
JM	0.656	1.000	1.000	L
XM	0.922	0.992	0.992	L
DG	0.422	1.000	1.000	L
BH	0.023[a]	0.977	0.984	L
ND	0.023[a]	0.977	0.977	L
FCG	0.008[a]	0.945	0.977	L

[a]Significant evidence of a population decline from bottleneck, [b]Normal L-shaped allele frequency distribution

In the study of Falcone and Mackayt (1996) it was postulated that if one population was considered to be in accordance with Hardly-Weinberg equilibrium, this population would have stable gene frequency in the next generation under optimal population conditions. In our results, only 16 of totally 48 locus by populations tests deviated from the HWE condition ($d > 0$), but none of them remained significant after correction for multiple tests. Null alleles might probably be caused by priming site variation (Jones et al. 1998; Banks et al. 1999). This indicates that the population is exhibiting a heterozygote

excess phenomenon. This phenomenon is thought to be caused by non-random mating, overfishing, small population size or natural selection (Ward 2006).

Population structure

L. affinis populations along coastal waters of eastern and southern China appear to be panmictic with no genetic structure. In addition, F_{ST} detected weak but significant genetic differentiation in *L. affinis* between DG and other populations, which may be attributed to its unique geographical location. DG is located in the Pearl River Estuary (PRE). The PRE, with an area of over 2000 km^2, is located along a micro-tidal coast with an average tidal range of about 1 m. The combined effect of a small tidal range and large bay area has resulted in interesting circulation and plume behaviour in the PRE (Su 2004). In addition, the waters of the sea have alternated between freshwater and salty waters all year round (Yuan et al. 2017), and the depth of the water is deeper, which may induce the possibility of frequent gene exchange, then hinder the interaction of genes in the region with the outside world. On that basis, the *L. affinis* inhabiting this region evolved in relative independent geographical isolation and occupied separate spawning grounds which led to a reduction of gene exchange with other populations.

History and current gene flow

Despite high polymorphism, an adequate number of loci and a large sample size, no genetic bottlenecks were dectected in all of the tests conducted with regard to *L. affinis*. The genetic patterns revealed in *L. affinis* can be explained by marine environment, the pertinent features of species biology in addition to their interactions.

Our results suggest that the strong connectivity and weak level of genetic structure of *L. affinis* along the Pacific coast of eastern and southern China has probably come about as a result of ongoing gene flow and historical contacts between present-day populations. Furthermore, the lack of isolation-by-distance, 3D-FCA and Bayesian clustering analysis also support the view that the lack of significant genetic population structure was caused by high contemporary gene flow.

The control region of the mitochondrial DNA possesses a higher mutation rate than other gene fragments among the mtDNA and is suitable for examing intraspecific population relationships (Moritz et al. 1987; Stoneking et al. 1991). Genetic divergence among populations may thus be caused by differences that are neutral within the species, representing a historical mark of divergent lineages that have remerged (Kvie et al. 2013; Webb et al. 2011). Although we can learn information about the history of populations based on mitochondrial DNA, it is still difficult to distinguish the current from past gene flow.

The microsatellite loci are more sensitive to detecting weak genetic differentiations, and what's more, they can provide more information about gene flow regarding current-day or recent generations. This present study utilized microsatellite loci and appears to be the first to report a genetically isolated research and to assess the population structure of *L. affinis*. Microsatellite DNA data confirmed our previous results about genetic differentiation based on sequences of mitochondrial DNA (mtDNA) gene fragment (Liu et al. 2018). The combination of the assignment data based on microsatellite genotypes with mtDNA sequence information provided more powerful evidence regarding the high genetic diversity and homogeneity of *L. affinis*.

Influence factors and deficiencies

The egg and larval dispersal of *L. affinis* may be the key factor for gene flow across large geographic distances (Liu et al. 2018). Strong dispersal ability of larvae (more than three months) was predicted for *L. affinis* (Yoshimatsu et al. 1993; Yang and Qui 1989), and it might greatly strengthen the dispersal potential and increase gene flow. It's still very hard to establish genetic estimates about dispersal distances, connectivity of marine populations and origins of larvae and adults (Bohonak 1999; Largier 2003). Many reseaches have indicated that the complex marine currents of the studied area are likely tohave an important effect on planktonic larvae (Hu et al. 2000; Hwang and Wong 2005). The China Coastal Current flows southward during winter monsoon winds and the South China Sea Warm Current can carry the planktonic larvae northward in summer (Hu et al. 2000; Guan and Fang 2006; Tsang et al. 2008). At the mercy of these marine currents, the larvae of *L. affinis* can spread long distances despite their limited dispersal capabilities.

Although sample sizes for microsatellites are widely distributed along the eastern and southern coast of China, these types of sample sites were still few in the Yellow Sea coast of China and in the waters of Japan (and non existent in the northern part of Hollaido), from which more samples should be collected. Based on our results, we suggest that measures restricting the total volume of *L. affinis* caputured or protecting the species in its reproductive period should be implemented and the population DG be treated as a special management unit. In addition, we further call for a better

understanding of such relationships, given that the health, economic and ecological costs associated with infectious diseases are substantial and proper methods to estimate the effective size of stocks and monitor genetic diversity should be employed. Our findings can effectively contribute to the proper and appropriate design of management plans and aid in the conservation of this important and overexploited fish resource in the Northwest Pacific.

5. Conclusion

The results of our research revealed that there is a high level of genetic diversity within populations of *L. affnis* along the coasts of eastern and southern China. There is no significant genetic differentiations among populations except that there are small but significant genetic differentiations between DG and most other populations, which might have arisen as a result of complex marine currents and differences in larval dispersal. Therefore, molecular data is particular useful in understanding the biological significance of genetic differentiation and identifying units for management and protection. In addition, our study has positive reference value with regard to calculating population structure and dispersal in the same marine region for species with similar life-history traits.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (41506158, 41776171) and the Special Fund for Agro-scientific Research in the Public Interest (201303048). We thank Dr. Yuan Li, Dr. Liangmin Huang, Mr. Binbin Shan, Mr. Long Yan for collecting the samples. We also thank Dr. Jiaguang Xiao and Mr. Wei Zhou for their help in data processing.

References

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, 447 p

Banks MA, Blouin MS, Baldwin BA, Rashbrook VK, Fitzgerald HA, Blankenship SM, Hedgecock D (1999) Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tschawytscha*). J Hered **90**(2):281–288

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2002) GENETIX 4.05, logiciel sous Windows TM pour la Génétique des Populations. Université de Montpellier II, Montpellier (in French)

Bohonak AJ (1999) Dispersal, gene flow, and population structure. Q Rev Biol **74**(1):21–45

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. J Hered **93**:153–154

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet **32**:314–331

Carlsson J, McDowell JR, Diaz–Jaimes P, Carlsson JE, Bole SB, Gold JR, Graves JE (2004) Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. Mol Ecol **13**:3345–3356

Cheng J, Yanagimoto T, Song N, Gao TX (2015) Population genetic structure of chub mackerel *Scomber japonicus* in the Northwestern Pacific inferred from microsatellite analysis. Mol Biol Rep **42**:373–382

Diáz–Ferguson E, Haney RA, Wares JP, Silliman BR (2010) Population genetics of a trochid gastropod broadens picture of Caribbean Sea connectivity. PLoS One **5**:e12675. doi:10.1371/journal.pone.0012675

Durand JD, Blel H, Shen KN, Koutrakis ET, Guinand B (2013) Population genetic structure of *Mugil cephalus* in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers. Mar Ecol-Prog Ser **474**:243–261

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour **4**:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol **14**:2611–2620

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 Res **10**:564–567

Falconer DS, Mackayt FC (1996) Introduction to quantitative genetics. The Ronald Press Company, New York, 480 p

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics **164**(4):1567–1587

Ferreira ME, Grattapaglia D (1998) Introdução ao uso de marcadores moleculares em análise genética. Embrapa, Brasília, 220 p

Gao TX, Han ZQ, Zhang XM, Luo J, Yanagimoto T, Zhang H (2016) Population genetic differentiation of the black rockfish *Sebastes schlegelii* revealed by microsatellites. Biochem Syst Ecol **68**:170–177

Gong XL, Zhu WJ, Bao BL (2016) Complete mitochondrial DNA sequence of the Eastern keelback mullet *Liza affinis*. Mitochondr DNA **27**:2178–2179

Goudet J (2001) FSTAT V.2.9.3, a program to estimate and test gene diversities and fixation indices (Updated from Goudet 1995). University of Lausanne. http://www.unil.ch/izea/softwares/fstat.html Accessed 9 Sep 2017

Guan BX, Fang GH (2006) Winter counter–wind currents off the southeastern China coast: a review. J Oceanogr **62**:1–24

Gyllensten U (1985) The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous and freshwater species. J Fish Biol **26**:691–699

Hu J, Kawamura H, Hong H, Qi Y (2000) A review on the currents in the South China Sea: seasonal circulation, South China Sea Warm Current and Kuroshio intrusion. J Oceanogr **56**(6): 607–624

Hwang J, Wong C (2005) The China Coastal Current as a driving force for transporting *Calanus sinicus* (Copepoda: Calanoida) from its population centers to waters off Taiwan and Hong Kong during the winter northeast monsoon period. J Plankton Res **27**(2):205–210

Lee CL, Joo DS (1994) Synopsis of family Mugilidae (Perciformes) from Korea. Bull Korean Fish Soc **27**:814–824

Ikeda T, Mito S (1988) The identification of fish eggs and emergent larvae. In: Okiyama M (ed) An atlas of the early stage fishes in Japan. Tokai University Press, Tokyo, pp 999–1083 (in Japanese)

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. BMC Genet **6**:13. doi:10.1186/1471-2156-6-13

Joh YG, Lee KH, Cho YJ (1988) Lipid composition of roe, muscle and viscus of *Liza carinata*, a species of the Mugilidae family. J Food Sci Tech Mys **20**:674–682

Jones AG, Stockwell CA, Walker D, Avise JC (1998) The molecular basis of a microsatellite null allele from the white sands pupfish. J Hered **89**:339–342

Kim MJ, An HS, Choi KH (2010a) Genetic characteristics of Pacific cod populations in Korea based on microsatellite markers. Fisheries Sci **76**:595–603

Kim WJ, Kim KK, Han HS, Nam BH, Kim YO, Kong HJ, Noh JK, Yoon, M (2010b) Population structure of the olive flounder (*Paralichthys olivaceus*) in Korea inferred from microsatellite marker analysis. J Fish Biol **76**:1958–1971

Kinoshita I (1988) *Liza carinata carinata* in Mugilidae. In: Okiyama M (ed) An atlas of the early stage fishes in Japan. Tokai University Press, Tokyo, pp 390–391 (in Japanese)

Kvie KS, Hogner S, Aarvik L, Lifjeld JT, Johnsen A (2013) Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*). Ecol Evol **3**(1):126–144

Largier JL (2003) Considerations in estimating larval dispersal distances from oceanographic data. Ecol Appl **13**(1):S71–S89

Liao XJ, Chang H, Sun W (2006) Study of genetic diversity in six sheep/goat population with microsatellite markers. Acta Ecol Ani Dom **27**(2):13–18 (in Chinese)

Li WQ, Zheng WY, Cai AG, Chen WQ (1993) The study on the diet ingestion, absorption and utilization of *Liza carinatus*. Mar Sci **2**:43–47 (in Chinese)

Liu JY, Brown CL, Yang TB (2009) Population genetic structure and historical demography of grey mullet, *Mugil cephalus*, along the coast of China, inferred by analysis of the mitochondrial control region. Biochem Syst Ecol **37**:556–566

Liu L, Gao TX, Han ZQ, Li CH, Sun DR, Song N (2016a) Latin name correction of *Liza affinis* from China. J Fish Sci China **23**(5):1108–1116 (in Chinese)

Liu L, Gao TX, Han ZQ, Li CH, Sun DR, Song N (2016b) Development and characterization of microsatellite markers of the eastern keelback mullet (*Liza affinis*). Genet Mol Res **15**(2):1–4

Liu L, Panhwar SK, Gao TX, Han ZQ, Li CH, Sun DR, Song N (2017) New genetic evidence from three keel–backed *Liza* species based on DNA barcoding confirms morphology–based Identification. Pak J Zool **49**(5):1901–1907

Liu L, Liu LQ, Gao TX, Song N (2018) Phylogeographic pattern of *Liza affinis* populations in Chinese coastal waters: estimation of larval dispersal potential. Mitochondr DNA **29**:1253–1260. doi:10.1080/24701394.2018.1444038

Mai AC, Mino CI, Marins LF, Monteiro–Neto C, Miranda L, Schwingel PR, Lemos VM, Gonzalez–Castro M, Castello JP, Vieira JP (2014) Microsatellite variation and genetic structuring in *Mugil liza* (Teleostei: Mugilidae) populations from Argentina and Brazil. Estuar Coast Shelf S **149**:80–86

Minegishi Y, Aoyama J, Tsukamoto K (2008) Multiple population structure of the giant mottled eel, *Anguilla marmorata*. Mol Ecol **17**:3109–3122

Mito S (1966) Fish eggs and larvae. In: Okiyama M (ed) An atlas of the early stage fishes in Japan. Tokai University Press, Tokyo, pp 44–45 (in Japanese)

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu Rev Ecol Syst **18**:269–292

Nakamura S (1936) Larvae and young of fishes found in the vicinity of Kominato –XV. J Imp Fish Inst 7:135–144 (in Japanese)

Nash CE, Shehadeh ZH (1980) Review of breeding and propagation techniques for grey mullet, *Mugil cephalus* L. ICLARM Stu Rev **3**:11–77

Nielsen EE, Hansen MM, Schmidt C, Meldrup D, Grønkjaer P (2001) Population of origin of Atlantic cod. Nature **413**:272

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD (2008) Genes mirror geography within Europe. Nature **456**:98–101

Oliver L (2002) Populations 1.2.28. http://www.cnrs–gif.fr/pge/bioinfo/populations Accessed 9 Sep 2017

Olivier L, Wolf YI, Koonin EV, Aravind L (2012) The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Res **12**:1048–1059

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Annul Rev Ecol Syst **25**:547–572

Park SDE (2001) Trypanotolerance in West African cattle and the population genetic effects of selection. Ph.D. Thesis, University of Dublin

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J Hered **90**(4):502–503

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genet **155**(2):945–959

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. J Hered **86**:248–249

Rocha–Olivares A, Garber NM, Stuck KC (2000) High genetic diversity, large inter-oceanic divergence and historical demography of the striped mullet. J Fish Biol **57**:1134–1149

Rocha–Olivares A, Garber NM, Garber AF, Stuck KC (2017) Structure of the mitochondrial control region and flanking tRNA genes of *Mugil cephalus*. Hidrobiologica **15**:139–149

Rice WR (1989) Analysing tables of statistical tests. Evolution **43**:223–225

Sambrook J, Russell DW (2001) Gel electrophoresis of DNA and pulsed–field agarose gel electrophoresis. In: Green MR, Sambrook J (eds) Molecular cloning: a laboratory manual. Cold Spring Harbor laboratory press, New York, pp 5.4–5.86

Selkoe KA, Aloia CCD, Crandall ED, Lacchei M, Liggins L, Puritz JB, Heyden SVD, Toonen RJ (2016) A decade of seascape genetics: contributions to basic and applied marine connectivity. Mar Ecol-Prog Ser **554**:1–19

Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, Erlich HA (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence–specific oligonucleotide probes. Am J Hum Genet **48**(2):370–382

Su JL (2004) Overview of the South China Sea circulation and its influence on the coastal physical oceanography outside the Pearl River Estuary. Cont Shelf Res **24**:1745–1760

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol **21**:3907–3930

Tsang LM, Chan BKK, Ma KY, Chu KH (2008) Genetic differentiation, hybridization and adaptive divergence in two subspecies of the acorn barnacle *Tetraclita japonica* in the northwestern Pacific. Mol Ecol **17**:4151–4163

Van OC, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes **4**:535–538

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered **89**:439–450

Ward RD (2006) The importance of identifying spatial population structure in restocking and stock enhancement programmes. Fish Res **80**:9–18

Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. J Fish Biol **44**:213–232

Webb WC, Marzluff JM, Omland KE (2011) Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. Mol Ecol **20**:2390–2402

Whitfield AK, Panfili J, Durand JD (2012) A global review of the cosmopolitan flathead mullet *Mugil cephalus* Linnaeus 1758 (Teleostei: Mugilidae), with emphasis on the biology, genetics, ecology and fisheries aspects of this apparent species complex. Rev Fish Biol Fisher **22**:641–681

Yang SY, Qiu SY (1989) A preliminary study on growth and feeding habits of Mullet (*Liza carinata*) underyearlings in Luoyuan Bay. J Xiamen Univ **28**:110–119 (in Chinese)

Yoshimatsu T, Matsui S, Kitajima C (1993) Early development of laboratory–reared keelback mullet. Nippon Suisan Gakk **59**(5): 765–776

Yuan M, Tang Y, Xu SN, Chen ZZ, Yang YT, Jiang YE (2017) Community structure of fishery resources from the Nansha Waters of Pearl River Estuary in autumn. South China Fisher Sci **13**:18–25 (in Chinese)

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol **12**:563–584

Zhang R, Lu S, Zhao CY, Chen L, Zang Z, Zhang X (1985) Fish eggs and larvae in the offshore waters of China. Shanghai Sci Tech Press, Shanghai, 206 p