# EXPLORING THE GENETIC DIVERSITY OF MARINE ORGANISMS BASED ON THE ANALYSIS OF CHROMOSOME AND GENOMIC DNA MARKERS

EDITED BY: Dongdong Xu, Zhiqiang Han, Tifeng Shan and Cassia Fernanda Yano

PUBLISHED IN: Frontiers in Marine Science

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-693-7
DOI 10.3389/978-2-88971-693-7

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

# EXPLORING THE GENETIC DIVERSITY OF MARINE ORGANISMS BASED ON THE ANALYSIS OF CHROMOSOME AND GENOMIC DNA MARKERS

Topic Editors:
**Dongdong Xu,** Zhejiang Marine Fisheries Research Institute, China
**Zhiqiang Han,** Zhejiang Ocean University, China
**Tifeng Shan,** Chinese Academy of Sciences (CAS), China
**Cassia Fernanda Yano,** Federal University of Mato Grosso, Brazil

**Citation:** Xu, D., Han, Z., Shan, T., Yano, C. F., eds. (2021). Exploring the Genetic Diversity of Marine Organisms Based on the Analysis of Chromosome and Genomic DNA Markers. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-693-7

# Table of Contents


04 ***Analysis of Genetic Structure of Wild and Cultured Giant Freshwater Prawn (*Macrobrachium rosenbergii*) Using Newly Developed Microsatellite***
Lingyun Yu, Xinping Zhu, Jianhui Liang, Jiajia Fan and Chen Chen

13 ***Assessment of the Genetic Connectivity Between Farmed Populations on a Typical Kelp Farm and Adjacent Spontaneous Populations of* Saccharina japonica *(Phaeophyceae, Laminariales) in China***
Tifeng Shan, Qianxi Li, Xuemei Wang, Li Su and Shaojun Pang

20 ***Seasonal Fish Assemblage Structure Using Environmental DNA in the Yangtze Estuary and Its Adjacent Waters***
Hui Zhang, Susumu Yoshizawa, Wataru Iwasaki and Weiwei Xian

30 ***Phylogeography and Historical Demography of Two Sympatric Atlantic Snappers:* Lutjanus analis *and* L. jocu**
Allyson Santos de Souza, Eurico Azevedo Dias Júnior, Manolo Fernandez Perez, Marcelo de Bello Cioffi, Luiz Antonio Carlos Bertollo, Erik Garcia-Machado, Marcelo Nazareno Souza Vallinoto, Pedro Manoel Galetti Jr. and Wagner Franco Molina

41 ***Cryptic Diversity and Database Errors Challenge Non-indigenous Species Surveys: An Illustration With* Botrylloides *spp. in the English Channel and Mediterranean Sea***
Frédérique Viard, Charlotte Roby, Xavier Turon, Sarah Bouchemousse and John Bishop

54 ***Overview on Karyotype Stasis in Atlantic Grunts (Eupercaria, Haemulidae) and the Evolutionary Extensions for Other Marine Fish Groups***
Clóvis Coutinho da Motta-Neto, Marcelo de Bello Cioffi, Gideão Wagner Werneck Félix da Costa, Karlla Danielle Jorge Amorim, Luiz Antonio Carlos Bertollo, Roberto Ferreira Artoni and Wagner Franco Molina

66 ***Artificial Induction and Genetic Structure Analysis of Tetraploid Turbot* Scophthalmus maximus**
Zhihao Wu, Lijuan Wang, Yunliang Lu, Xiangping Zhu, Xinlu Yue and Feng You

78 ***Genome-Wide Discovery of Single-Nucleotide Polymorphisms and Their Application in Population Genetic Studies in the Endangered Japanese Eel (*Anguilla japonica*)***
Bing-Jian Liu, Yu-Long Li, Bai-Dong Zhang and Jin-Xian Liu

89 ***Genetic Assessment of a Black Rockfish,* Sebastes schlegelii*, Stock Enhancement Program in Lidao Bay, China Based on Mitochondrial and Nuclear DNA Analysis***
Lijuan Wang, Zhihao Wu, Yingjun Wang, Mengxia Liu, Aihuan Song, Hongjun Liu and Feng You

100 ***Genetic Population Structure of the Hard Clam* Meretrix meretrix *Along the Chinese Coastlines Revealed by Microsatellite DNA Markers***
Yingying Ye, Chengrui Yan, Wansuk Senanan, Baoying Guo, Kaida Xu and Zhenming Lü

frontiers in Marine Science

ORIGINAL RESEARCH
published: 18 June 2019
doi: 10.3389/fmars.2019.00323

# Analysis of Genetic Structure of Wild and Cultured Giant Freshwater Prawn (*Macrobrachium rosenbergii*) Using Newly Developed Microsatellite

*Lingyun Yu[1]*, Xinping Zhu[1]*, Jianhui Liang[2], Jiajia Fan[1] and Chen Chen[1]*

*[1] Key Laboratory of Tropical and Subtropical Fishery Resources Application and Cultivation, Ministry of Agriculture and Rural Affairs, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, China, [2] Foshan Sanshui Baijin Aquatic Seedling Co., Ltd., Foshan, China*

The giant freshwater prawn (GFP) is one of the most critical crustacean species cultured in Southeast Asia. Investigation of the genetic structure of current commercial stocks allows GFP breeding programs to better manage crosses and germplasm banks as well as to promote the rational use of GFP. The objective of the study was to characterize genetic diversity in diverse prawn populations with emphasis on those cultured in China. Seventeen microsatellite loci, including 12 novel loci derived from GFP transcriptome data, were screened to assess genetic diversity in one wild (Myanmar) and six cultured populations (i.e., four Chinese (Zhejiang, Guangxi, and Guangdong A and B), one Malaysian, and one Thai population). The results showed that the number of alleles per locus ranged from 3 to 18. The mean observed heterozygosity (0.363 ± 0.048) was less than the expected heterozygosity (0.637 ± 0.048). The mean values of polymorphism information content among the seven populations were >0.5 (ranging from 0.110 to 0.915). These cultured populations exhibited reduced genetic diversity when compared with that of the wild population. Pair-wise genetic differentiation ranged from 0.006 to 0.131 within the seven populations. The dendrogram of the genetic distance shows that the six cultured populations were distributed on the same major branch, suggesting that they have are genetically close, whereas the wild population was distributed on an independent branch. The results provide a basic assessment of genetic diversity in some available stocks and lay a foundation for future research efforts toward genetic monitoring and selective breeding.

**Keywords: genetic diversity, genetic structure, giant freshwater prawn, *Macrobrachium rosenbergii*, microsatellite**

OPEN ACCESS

***Edited by:***
*Dongdong Xu,*
*Marine Fisheries Research Institute of Zhejiang, China*

***Reviewed by:***
*Supawadee Poompuang,*
*Kasetsart University, Thailand*
*Yanhe Li,*
*Huazhong Agricultural University, China*

****Correspondence:***
*Lingyun Yu*
*lysnp@163.com*
*Xinping Zhu*
*zhuxinping_1964@163.com*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *05 March 2019*
***Accepted:*** *28 May 2019*
***Published:*** *18 June 2019*

***Citation:***
*Yu L, Zhu X, Liang J, Fan J and Chen C (2019) Analysis of Genetic Structure of Wild and Cultured Giant Freshwater Prawn (Macrobrachium rosenbergii) Using Newly Developed Microsatellite. Front. Mar. Sci. 6:323. doi: 10.3389/fmars.2019.00323*

## INTRODUCTION

The decapod crustacean, giant freshwater prawn (GFP) *Macrobrachium rosenbergii*, also known as the Malaysian prawn or giant river prawn, is a commercially important species. It is also an important wild capture fishery species, particularly in Southeast Asia, with global production exceeding 216,856 tons in 2014 (Schneider et al., 2012; Food and Agriculture Organization of the United Nations [FAO], 2018). The natural distribution of GFP extends from Southern Vietnam in the east to Pakistan in the west, across Southeast Asia, Southern to Northern Australia, and New

Guinea and some Pacific and Indian Ocean Islands (Chand et al., 2005). With the development of modern culture techniques for GFP in 1959 at the Malaysian Fisheries Research Institute (Ling, 1969), the culture of GFP has rapidly expanded to areas outside the native range of this species. For instance, in 1965 and 1966, some individuals were imported to Hawaii of the United States from Penang, Malaysia (Fujimura and Okamoto, 1972). In 1976, a few GFPs were introduced into mainland China (Yang et al., 2008, 2012; Thanh et al., 2015). The introduction and development of this species form a significant part of aquaculture industry worldwide, especially in mainland China where the production of GFP is currently the highest globally. However, several critical issues, such as slow growth rate, size variation at harvest, disease, and deterioration of pond environment, affect the development of GFP farming industry (Chareontawee et al., 2007). Additionally, genetic deterioration in most hatchery stocks due to inbreeding might also affect the culture period, economic returns, and natural ecosystems (Thanh et al., 2015). These factors, particularly those involved in genetic degradation, hinder the sustainable development of this species. Thus, a breeding program and a strategy for proper conservation of this species are required.

Understanding genetic diversity in the species is critical for developing conservation strategies. Recognition of particular genetic diversity will also enable informed choice in breeding plans regarding the selection of genetically diverse brood-stock and the maintenance of genetic diversity in various stocks. Nowadays, the application of genetic markers has allowed rapid progress in the investigation of parentage assignments, genetic variability assessments, inbreeding determination, and species identification of commercially important candidate aquaculture species (Liu and Cordes, 2004). A few genetic studies have focused on GFP diversity, using allozymes (Hedgecock et al., 1979), mtDNA (De Bruyn et al., 2005), microsatellite (Charoentawee et al., 2006; Chareontawee et al., 2007; Divu et al., 2008; Schneider et al., 2012; Mohanty et al., 2013; Sun et al., 2015; Thanh et al., 2015), and single nucleotide polymorphisms (SNP) (Jung et al., 2014; Agarwal et al., 2016; Alam et al., 2017). Of these DNA markers used to examine genetic variation in populations, the microsatellite marker is the most suitable and preferred marker for genetic structure and conservation studies (Weber and May, 1989; Jarne and Lagoda, 1996; Vieira et al., 2016; Manechini et al., 2017). However, to date, the microsatellite resources of GFP remain deficient. Only approximately 54 pairs of microsatellite markers have been utilized to analyze genetic diversity of the species (Chand et al., 2005; Charoentawee et al., 2006; Divu et al., 2008; Mohanty et al., 2013; Thanh et al., 2015). This has become a major obstacle to understanding the genetic background and diversity of wild and cultured GFP populations. In the present study, we aimed to develop and characterize novel microsatellite markers in the GFP population and to investigate genetic diversity in one wild and six cultured populations for a clear understanding of genetic levels of several lines and to support future genetic management and brood-stock selection activities.

# MATERIALS AND METHODS

## Sample Collections and DNA Extraction

A total of seven populations were investigated in this study. The cultured populations were sampled from six farms located in Zhejiang, Guangxi, and Guangdong Provinces of China; Ayutthaya, Thailand; and Kuala Lumpur, Malaysia, which represent the major hatcheries from that region. The wild population was collected from the Yangon River of Myanmar (**Figure 1** and **Table 1**). Thirty adult prawns of each population were collected for genetic diversity analysis. All individuals were preserved directly in 95% ethanol and taken to the laboratory. Whole genomic DNA was then extracted using the E.Z.N.A.$^{TM}$ MicroElute Genomic DNA Kit (Omega Bio-Tek, Guangzhou, China) following the rapid phenol–chloroform protocol adapted from the standard procedure. The quality of extracted DNA was examined by running the sample on 0.8% agarose gel. The present study, in addition to its protocol, was approved by the Ethics Committee of Animal Experiments of Pearl River Fishery Research Institute before it began. All efforts were made to minimize the suffering of the GFPs.

## Microsatellite Markers and Genotyping

Twelve novel microsatellite sequences and primers of *M. rosenbergii* were obtained from GFP transcriptome data sources according to the method used for other species (Yu et al., 2015; **Supplementary S1** and **Table 2**). Five previously known microsatellite loci were also used in this study (**Table 2**; Mohanty et al., 2013; Sun et al., 2015; Thanh et al., 2015). The 5′ end of each forward primer was labeled with fluorescent dye (6-fluorescein amidite, HEX, or TAMRA). Multiplex polymerase chain reaction (PCR) was carried out using a 30-μL reaction mixture containing 1 U Tap DNA polymerase, 1 × PCR buffer, 0.3 mM dNTP, 0.3 μM each of forward and reverse primers, and 100 ng DNA template. The steps of PCR and genotype detection were the same as those described in the preceding subsection except the annealing temperature, which in multiplex 1 was 58°C, multiplex 2 was 52°C, and multiplex 3 was 60°C.

Amplified products were resolved via capillary electrophoresis using an ABI 3730xl DNA analyzer (Applied Biosystems). The size of fragments was determined using GeneMapper® software version 3.5 (Applied Biosystems) by comparison against the GenScan$^{TM}$ 500 ROX$^{TM}$ (Applied Biosystems) internal size standard.

## Statistical Analyses

Commonly derived statistics from the microsatellite genotypic data including the allele frequency, observed number of alleles (*No*), observed heterozygosity ($H_O$), expected heterozygosity ($H_E$), and polymorphic information content (PIC) were calculated using the Microsatellite Toolkit for each locus and population (Park, 2001). The Hardy–Weinberg equilibrium test was carried out using the GENEPOP computer program (Raymond and Rousset, 1995), which was also used to estimate *F* statistics ($F_{IT}$, $F_{IS}$, and $F_{ST}$) (Weir and Cockerham, 1984) for each


**FIGURE 1 |** Map of sampling sites of the seven *Macrobrachium rosenbergii* populations.

**TABLE 1 |** Location of populations and number of samples in each populations of *M. rosenbergii*.

| Number of pop | Sample name | Population | Amount of individuals | Source |
|---|---|---|---|---|
| 1 | AY | Cultured | 30 | Ayutthaya province of Thailand |
| 2 | GX | Cultured | 30 | Nanning city, Guangxi province of China |
| 3 | GDA | Cultured | 30 | Zhongshan city, Guangdong province of China |
| 4 | KL | Cultured | 30 | Kuala Lumpur city of Malaysia |
| 5 | YG | Wild | 30 | Yangon River, Yangon city of Myanmar |
| 6 | ZJ | Cultured | 30 | Huzhou city, Zhejiang province of China |
| 7 | GDB | Cultured | 30 | Foshan city, Guangdong province of China |

*Sample name refers to location in* ***Figure 1****.*

locus and the pairwise $F_{ST}$ between populations. Nei's genetic distance ($D_A$) (Saitou and Nei, 1987) between populations was measured using the Microsatellite Analyzer (Dieringer and Schlotterer, 2003). The neighbor-joining (NJ) tree was generated using MEGA 5.0 software based on $D_A$ (Nei, 1987) to reveal the genetic relationships among the populations collected in the present study.

The model-based approach for the population structure test of the seven prawn populations was carried out with the software STRUCTURE 2.3.1 (Pritchard et al., 2000), which assessed genomic clustering (K) of the sample. To obtain a representative value of K for data modeling, seven independent runs were performed for each value from one to seven. The run length was set to 100,000 burn-ins followed by 100,000 iterations. The ΔK estimated the most likely number of K that represented the population structure (Evanno et al., 2005). Principal component analysis (PCA) was performed using GENALEX v.6.501 software (Peakall and Smouse, 2006, 2012) in order to spatially plot clusters and individuals based on the distance matrix with data standardization.

# RESULTS

## Characteristics and Polymorphism of Microsatellite Loci

A polymorphism test of 12 novel loci was performed using the GFP populations. These markers comprised three hexanucleotide (Mr034642, Mr041463, and Mr036095), two

**TABLE 2 |** Primer sequence and annealing temperature for microsatellite marker amplification.

| No. | Locus name | *abbr.* | Primer sequences 5′-3′ | Lable | *Tm* (°C) | Size (bp) | Repeat motif |
|---|---|---|---|---|---|---|---|
| 1 | Unigene038721* | Mr03872 | F:TGGGCTTGGTCCATATGAGT R:TCTCCATCGAACTGACTCCC | TAMRA | 60 | 124–147 | (AAT)25 |
| 2 | Unigene082899* | Mr08289 | F:ATTTCTGCTGTGAGGGCAGT R:ACAGGTGTCGCTTTGGACTT | FAM | 58 | 210–264 | (GAA)23 |
| 3 | Unigene032300* | Mr032300 | F:TCTGAGGGCTTCTTCCATGT R:GGGGCCACTGACTCAAAAAT | HEX | 58 | 100–177 | (ATA)21 |
| 4 | Unigene001327* | Mr001327 | F:CGGGAGGGTATTGTAGCGTA R:CCTTGCAGAAAGGACCAATAG | TAMRA | 58 | 114–147 | (ATA)21 |
| 5 | Unigene041668* | Mr041668 | F:TCGTTTAAGTCCAAAGGCAG R:CTTGTCTCTTGAAGCCCCTG | FAM | 58 | 223–253 | (AATAG)5 |
| 6 | Unigene018861* | Mr018861 | F:GCAAGTGAGGAGGGTAGCAG R:AACCGTTCTCCTCGAAATCA | HEX | 59 | 204–214 | (GAAGT)5 |
| 7 | Unigene034642* | Mr034642 | F:CGAATGGGGTCACATTAGGT R:TCACGGTTTGTGCTCAGTTT | TAMRA | 59 | 218–250 | (CTGAGT)7 |
| 8 | Unigene041463* | Mr041463 | F:GTGGGCGATCTCCAAATAG R:GCTCAGGGGCACGTTTAATA | FAM | 60 | 229–247 | (GCAAAT)5 |
| 9 | Unigene036095* | Mr036095 | F:TTCCTCAGCACCGTAACCTC R:CGCTCATTCGATCACCTGTA | HEX | 58 | 126–143 | (CAACAC)4 |
| 10 | Unigene001885* | Mr001885 | F:CACATTGAGCTCGCTGAGA R:AGCCATGCCTGCTCACAT | TAMRA | 58 | 246–281 | (TAG)7 |
| 11 | Unigene015914* | Mr015914 | F:GCAGCTCCGTAGCAGCTATAA R:AGAAAAGACGAAGACGCGAG | FAM | 58 | 256–274 | (TGC)5 |
| 12 | Unigene044258* | Mr044258 | F:GTGGTTAACGTCGCTCCTTC R:CATGTTCGTGTGGGAGAGAA | HEX | 58 | 258–274 | (AG)5 |
| 13 | EMR-55 | EMR-55 | F:GAAGTCATCCGACAAACTTCAC R:AGTAATCATGTGGCCTAGCCTAG | FAM | 58 | 189–234 | (TTA)13 |
| 14 | Y6 | Y6 | F: CATCAGCATTTGGCAGT R: AGCCCTTGAACTTGTTGTAT | HEX | 52 | 141–183 | (ATT)7 |
| 15 | Y7 | Y7 | F: ATGCCTGGAAAGAATGAG R: TTGTCTGAGCCTGAAACC | TAMRA | 52 | 298–347 | (TG)11 |
| 16 | Y16 | Y16 | F: ATTCGGTATCAGCTCTGC R: AGGTCATCACCCTTTCCA | FAM | 58 | 189–216 | (CTG)7 |
| 17 | W24 | W24 | F: AGGATTTCTGCGAGGTCTTG R: CGTGTTGTTCTTCATAGGCTTC | HEX | 60 | 171–213 | (TCT)2(ATA)2 |

**Primer sequences showed in **Supplementary S1**. Five previously known microsatellite loci (EMR-55, Y6, Y7, Y16, and W24) first reported in Mohanty et al., 2013; Sun et al., 2015 and Thanh et al., 2015.*

tetranucleotide (Mr041668 and Mr018861), and one dinucleotide units (Mr044258), whereas the other contained trinucleotide units. Among five previously known makers, there were one complex repeat (W24), one dinucleotide repeat (Y7), and two trinucleotide units (EMR-55, Y6, and Y16). Seventeen loci were selected to investigate the genetic structure of cultured and wild GFP (**Table 2**).

Polymorphism was apparent at all the microsatellite loci in the seven GFP populations. The allele frequency at each locus in all the populations is shown in **Figure 2**. The genetic characteristics of the 17 microsatellite loci are listed in **Table 3**. The average *Na* was 8.94. The actual number of alleles ranged from 3 (Mr018861) to 18 (Y6). The average *Ne* ranged from 1.128 (Mr04146) to 12.570 (Y6), with an average across loci of 3.959. The PIC value ranged from 0.110 (Mr041463) to 0.9149 (Y6), with an overall average of 0.591. All the selected microsatellite loci were sufficiently polymorphic, indicating that these loci were suitable for the genetic analysis of GFP populations. The *He* among the seven microsatellite loci ranged from 0.077 (Mr041463) to 0.886 (Y6). The *Ho* among all the microsatellite loci ranged from 0.114 (Mr041463) to 0.923 (Y6), with an average of 0.637 (**Table 3**).

## Genetic Variability Among Intra- and Inter-Populations

The genetic polymorphism, including the *Ho*, *He*, PIC, and mean *Na* and *Ne*, were counted to assay the allelic diversity at each locus and each population. The results of 17 loci for the seven GFP populations are listed in **Table 4** and **Supplementary Table S1**. The *Ho* varied from 0.313 in GX to 0.377 in AY and GDB, whereas the *He* varied from 0.569 in GX to 0.630 in YG. The PIC ranged from 0.503 in GX to 0.581 in YG. The mean *Ho* of all the populations at the 17 loci was lower than the *He*, and the PIC values were all $>0.5$. The YG population displayed the highest values of *Na*, *Ne*, *I*, *He*, and PIC. The GX population had the smallest value among all the genetic parameters.

The evaluation of genetic differentiation among the seven populations, using the genetic differentiation ($F_{ST}$) and $D_A$ between populations for comparison, are provided in **Table 5**. The $F_{ST}$ of each population pair was highly statistically significant

**TABLE 3 |** Genetic variability at eighteen microsatellite loci in seven populations of *M. rosenbergii*.

| No. | Locus name | *A* | *Ae* | *I* | *Ho* | *He* | HW | PIC |
|---|---|---|---|---|---|---|---|---|
| 1 | Mr03872 | 7 | 2.206 | 0.933 | 0.452 | 0.548 | 0.005 | 0.447 |
| 2 | Mr08289 | 16 | 6.286 | 2.152 | 0.157 | 0.843 | 0.001* | 0.823 |
| 3 | Mr032300 | 8 | 2.488 | 1.090 | 0.401 | 0.600 | 0.000* | 0.516 |
| 4 | Mr001327 | 12 | 2.925 | 1.430 | 0.340 | 0.660 | 0.000* | 0.600 |
| 5 | Mr041668 | 6 | 2.620 | 1.100 | 0.380 | 0.620 | 0.994 | 0.548 |
| 6 | Mr018861 | 3 | 2.702 | 1.038 | 0.368 | 0.631 | 0.000* | 0.552 |
| 7 | Mr034642 | 6 | 1.862 | 0.9130 | 0.536 | 0.464 | 0.999 | 0.429 |
| 8 | Mr041463 | 4 | 1.128 | 0.268 | 0.886 | 0.114 | 0.992 | 0.110 |
| 9 | Mr036095 | 4 | 1.987 | 0.797 | 0.502 | 0.498 | 0.002* | 0.404 |
| 10 | Mr001885 | 6 | 2.079 | 1.040 | 0.480 | 0.520 | 0.095 | 0.483 |
| 11 | Mr015914 | 10 | 3.276 | 1.590 | 0.267 | 0.733 | 0.000* | 0.699 |
| 12 | Mr044258 | 7 | 3.865 | 1.500 | 0.257 | 0.743 | 0.514 | 0.698 |
| 13 | EMR-55 | 14 | 9.518 | 2.395 | 0.103 | 0.897 | 0.858 | 0.886 |
| 14 | Y6 | 18 | 12.570 | 2.676 | 0.077 | 0.923 | 0.756 | 0.915 |
| 15 | Y7 | 15 | 6.586 | 2.135 | 0.150 | 0.850 | 0.486 | 0.831 |
| 16 | Y16 | 8 | 1.987 | 1.017 | 0.502 | 0.497 | 0.863 | 0.466 |
| 17 | W24 | 8 | 3.214 | 1.418 | 0.310 | 0.691 | 0.000* | 0.643 |
| Mean | | 8.94 | 3.959 ± 0.747 | 1.382 ± 0.154 | 0.363 ± 0.048 | 0.637 ± 0.048 | 0.656 ± 0.117 | 0.591 ± 0.050 |

*A, number of observed alleles; Ae, effective alleles; I, shannon index; Ho, observed heterozygosity; He, expected heterozygosity; HW, Hardy–Weinberg equilibrium; PIC, polymorphism information content. *Significant ($P < 0.05$) departure from the Hardy-Weinberg equilibrium.*

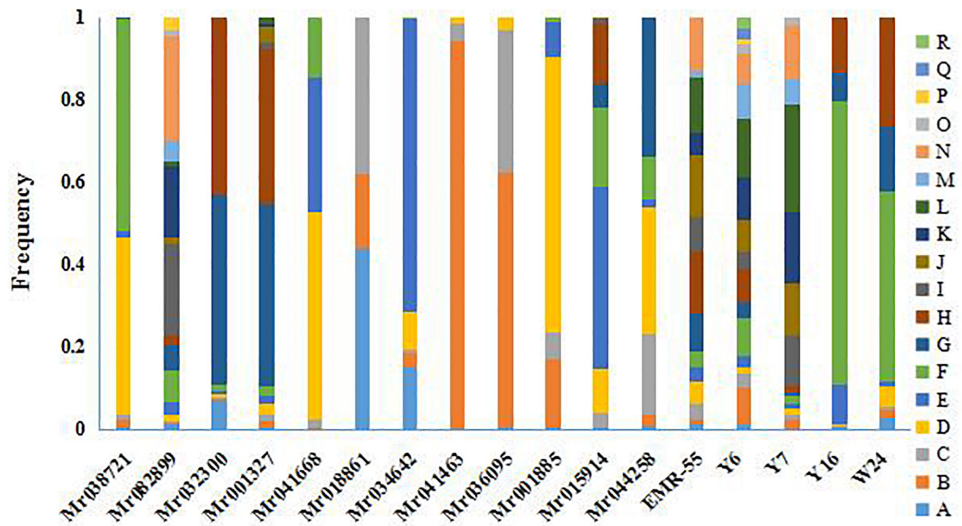


**FIGURE 2 |** Histograms presented frequency of alleles across loci in the seven populations of *M. rosenbergii*.

**TABLE 4 |** Genetic variability at seventeen microsatellite loci in seven populations of *M. rosenbergii*.

| Population | *A* | *Ae* | *I* | *Ho* | *He* | *Fis* | PIC |
|---|---|---|---|---|---|---|---|
| AY | 5.88 ± 0.94 | 3.361 ± 0.635 | 1.172 ± 0.157 | 0.377 ± 0.070 | 0.579 ± 0.056 | −0.137 ± 0.108 | 0.518 ± 0.055 |
| GX | 4.94 ± 0.725 | 2.801 ± 0.359 | 1.071 ± 0.122 | 0.313 ± 0.070 | 0.569 ± 0.49 | −0.230 ± 0.104 | 0.503 ± 0.048 |
| GDA | 5.18 ± 0.801 | 3.104 ± 0.514 | 1.135 ± 0.137 | 0.339 ± 0.068 | 0.589 ± 0.049 | −0.162 ± 0.100 | 0.525 ± 0.040 |
| KL | 6.53 ± 0.814 | 3.280 ± 0.448 | 1.284 ± 0.128 | 0.341 ± 0.055 | 0.627 ± 0.043 | −0.090 ± 0.076 | 0.569 ± 0.045 |
| YG | 6.88 ± 0.970 | 3.816 ± 0.708 | 1.308 ± 0.165 | 0.327 ± 0.070 | 0.630 ± 0.052 | −0.140 ± 0.066 | 0.581 ± 0.052 |
| ZJ | 5.35 ± 0.712 | 3.239 ± 0.422 | 1.140 ± 0.141 | 0.316 ± 0.061 | 0.621 ± 0.048 | −0.171 ± 0.099 | 0.555 ± 0.049 |
| GDB | 6.29 ± 0.873 | 3.310 ± 0.491 | 1.240 ± 0.139 | 0.377 ± 0.058 | 0.611 ± 0.048 | −0.027 ± 0.083 | 0.554 ± 0.050 |
| Total | 5.87 ± 0.278 | 3.273 ± 0.115 | 1.193 ± 0.033 | 0.342 ± 0.010 | 0.604 ± 0.010 | −0.137 ± 0.024 | 0.544 ± 0.011 |

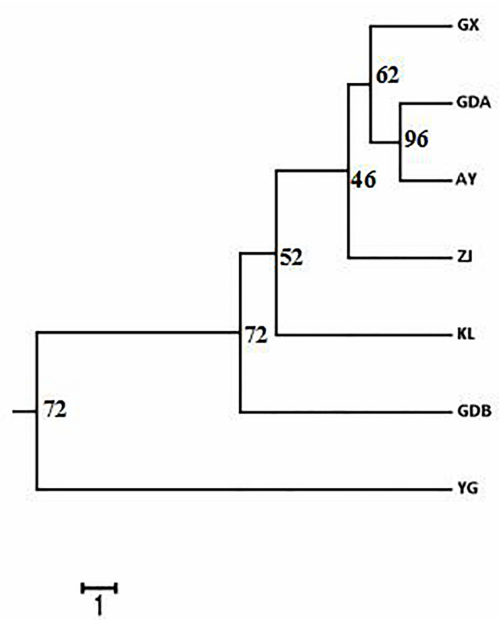


**FIGURE 3 |** Representation of neighbor-joining Nei's (1972) standard genetic distance among the seven populations of *M. rosenbergii* based on 1,000 replicates (numbers in nodes are percentage bootstrap values). See **Table 1** for population abbreviations.

**TABLE 5 |** Matrix of pair-wise $F_{ST}$ values (below diagonal) and genetic distances ($D_A$: above diagonal) between seven populations of *M. rosenbergii*.

| | AY | GX | GDA | KL | YG | ZJ | GDB |
|---|---|---|---|---|---|---|---|
| AY | | 0.049 | 0.049 | 0.074 | 0.226 | 0.074 | 0.107 |
| GX | 0.032* | | 0.031 | 0.133 | 0.259 | 0.052 | 0.142 |
| GDA | 0.018* | 0.068* | | 0.109 | 0.274 | 0.061 | 0.146 |
| KL | 0.022* | 0.052* | 0.006* | | 0.237 | 0.107 | 0.121 |
| YG | 0.044* | 0.040* | 0.070* | 0.071* | | 0.239 | 0.267 |
| ZJ | 0.032* | 0.049* | 0.021* | 0.027* | 0.050* | | 0.122 |
| GDB | 0.105* | 0.105* | 0.123* | 0.131* | 0.113* | 0.108* | |

**Pair-wise $F_{ST}$ was significant at $p < 0.05$.*

($P < 0.05$). The GDB population presented the highest $F_{ST}$ as to the other population, whereas the $F_{ST}$ value was the smallest for pair-wise comparisons between the KL and GDA populations. The $F_{ST}$ values of the remaining population pair ranged from 0.018 (for the GDA and AY population pair) to 0.071 (for the YG and KL population pair). The genetic distance between GDA and YG was the highest (0.274), and the minimum distance was observed between YG and GDA (0.031). The Analysis of Molecular Variance (AMOVA) indicated that 7.772% of the total variance was between the groups (i.e., wild and cultured populations), whereas 3.869% of the total variation and 8.218% of the variation belonged to differences among populations and within populations, respectively (**Table 6**).

## Population Differentiation Analysis

The $D_A$ of these seven GFP populations was calculated, and the $D_A$ distance matrix was utilized to construct the NJ tree. Chinese cultured populations (GX, GDA, ZJ, and GDB) and the other cultured populations (AY and KL) were gathered together in a single major category. Among them, the GDA and AY populations were close to each other. YG of the wild population was clearly depicted as an independent taxon from the cultured populations (**Figure 3**). The PCA of pair-wise genetic distances in the seven populations was employed to represent the relative position of the individuals and populations. The first PC1, second PC2, and third PC3 accounted, respectively for 4.35, 3.47, and 2.9% of the total variation (**Figure 4**). The PCA result was similar to those of the phylogenetic tree drawn using the NJ method. The cultured populations were grouped in a cluster, whereas the wild population formed a distinct group.

The STRUCTURE software program using Bayesian model-based clustering algorithms of multi-locus genotypes was employed to assign individuals into populations by estimated individual admixture proportions and to infer the number of populations (K) for a given sample. When all the seven populations were separated into two parts ($K = 2$), one included four populations (i.e., AY, GX, GDA, and ZJ) and the other included three populations (KL, YG, and GDB). When $K = 3$, all the populations were divided into three parts, the first one included three populations (AY, KL, and GDB), another also included three populations (GX, GDA, and ZJ), whereas the wild population (YG) was in an independent group (**Figure 5**). At $K = 4$, the GDB population was separated into a new branch, whereas the AY population was disordered with no clear structure. However, no new bunch appeared, and the figures were not much different for values of $K > 5$ (**Supplementary Figure S1**). Based on the above results of STRUCTURE and the phylogenetic tree, the best *K*-value was 3.

# DISCUSSION

There were 12 novel SSR loci that were obtained in the study. Among these markers, several hypervariable loci in the GFP population, including Mr08289 and Mr001327, were found (**Table 2**), which had been previously obtained in this species (Chand et al., 2005; Charoentawee et al., 2006; Divu et al., 2008; Mohanty et al., 2013; Thanh et al., 2015). These complex loci might be beneficial for the analysis of populations due to their excellent diversity. Besides these novel loci, five known markers with hypervariable loci were also utilized in this study to examine the alleles of the GFP populations to better estimate genetic diversity. The mean values of *Na*, *Ho*, *He*, and PIC across all the markers were 8.94, 0.363, 0.637, and 0.591, respectively. Previous studies have suggested that microsatellite markers used in studies of genetic variation and distance should have *Ho* values between 0.3 and 0.8 in the population, and the *Na* value should be at least five alleles in order to reduce the standard error of the distance estimates (Excoffier and Lischer, 2010; Lai et al., 2018). Our results suggested that the complexity of all the SSR markers utilized in the current study made them suitable for assessing the genetic structure of the GFP.

**TABLE 6 |** Analysis of molecular variance (AMOVA) results for seven populations of *M. rosenbergii*.

| Source of variation | Sum of squares | Variance components | Percentage variation | Fixation index Average *F*-statistics over all loci |
|---|---|---|---|---|
| Among groups | 64.624 | 0.4529 | 7.772 | $F_{SC}$: 0.042 |
| Among populations within groups | 90.764 | 0.225 | 3.869 | $F_{CT}$: 0.078 |
| Total | 155.388 | 0.6779 | | |

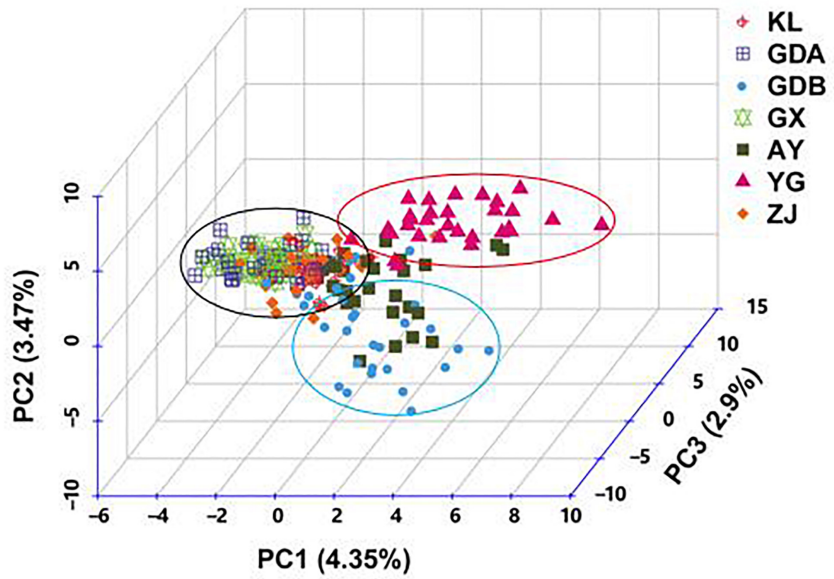


**FIGURE 4 |** Principal coordinate analysis of all the individuals of *M. rosenbergii* based on dissimilarity matrix (1-Jaccard) between pairs of genotypes. Capital letter indicates the name of sample. PC, Principal coordinate.

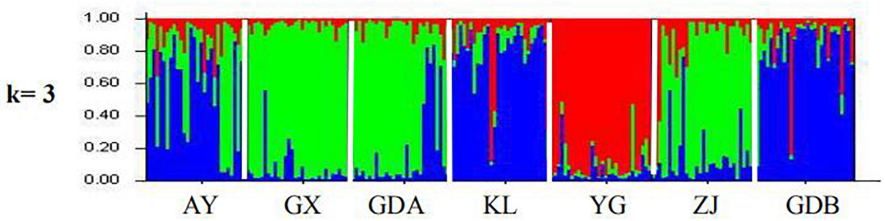


**FIGURE 5 |** Clustering analysis of cultured and wild populations of *M. rosenbergii*.

Information on genetic structure and connectivity of populations is critical for sustainable harvest of populations and the administration of diversity. A major finding of the study was the recognition that diversity in the cultured populations was low when compared with that in the wild population. The results showed that all the cultured populations had relatively low levels of genetic diversity, which was less than that reported previously (Chareontawee et al., 2007; Schneider et al., 2012), especially in genetic parameter comparison (**Supplementary Table S1**) using these published microsatellite loci (i.e., EMR-55, Y7, Y16, and W24) except Y6 (Sun et al., 2015; Thanh et al., 2015). The situation regarding the genetic background of GFP is a concern for the future. This might have several potential reasons, such as null alleles (Van Oosterhout et al., 2004), random genetic drift, founder effects, and artificial and natural selection in the cultured environments (Lacy, 1987; Thanh et al., 2015). Furthermore, it was probably caused by a rapid decline in wild stocks due to over harvesting, habitat loss, and increased population (Ng, 1997; New, 2000). In the present study, the wild population (YG) was collected in the Yangon River (Myanmar), which is a rich source of genetic diversity based on genetic allelic richness and *He* (**Table 4**). The phenomenon suggested that the wild population may be utilized to improve genetic diversity for their breeding program. In contrast, the GX and AY populations of cultured populations from China and Thailand possessed relatively fewer genetic resources. In the GFP aquaculture history of China, approximately 90% of the GFP seeds used in production are the offspring of the stock introduced about 25 years ago (Miao and Ge, 2002), and numerous private hatcheries have been utilizing broodstock identified as local strains to produce larvae in

Thailand (Chareontawee et al., 2007). These cultured populations are probably a result of being both closed and small for a long period. Additional investigation on the genetic resources of these cultured populations should therefore be performed to verify the specific causes.

Another major outcome of the study was gaining a better understanding of the level and pattern of genetic diversity present in some available GFP strains in China and some populations in Southeast Asia, which would be suitable for a stock improvement program or intraspecific hybridization. Based on the AMOVA results, the proportion of genetic variation attributed to differences among groups in this study was approximately 7.772%, whereas the proportion attributed to 'among populations within groups' was 3.869% in the total genetic variation (**Table 6**), indicating that wild population and cultured populations should not be derived from the same source. This phenomenon was also confirmed with the NJ tree, which showed two clusters of wild populations and all the cultured populations (**Figure 3**). However, the wild population (YG) was an independent group, and all the cultured populations were divided into two parts (part 1 including the AY, KL, and GDB populations, and part 2 including the GX, GDA, and ZJ populations) when $K = 3$ based on the STRUCTURE analysis (**Figure 5**). Furthermore, similar results were obtained using the PCA (**Figure 4**). It is thus speculated that all the cultured populations were derived from other resources rather than the resources of Myanmar.

In the cultured populations, the analysis of GFP genetic identity confirmed that the cultured populations in China have relatively high similarity with those in Thailand and Malaysia (**Table 5**), indicating these commercial hatchery stocks in China are still derived mainly from the Southeast Asian resource. On the contrary, the AY, KL, and GDB populations were of the same origin, whereas the GX, GDA, and ZJ populations were from another similar resource based on the STRUCTURE analysis (**Figure 5**). The results further suggested that these cultured populations of Southeast Asia were not derived from a single resource. Although we could not confirm the origin of cultured GFP populations based on the current research results, Thanh et al. (2015) thought that some cultured populations in China might have originated from a wild source of the Mekong or the Dong Nai River in Vietnam. Thus, further studies on the origin of GFP cultured resources are needed using representative wild populations coming from different regions of Southeast Asia.

## CONCLUSION

In conclusion, this study documented that the 17 SSR loci, including the 12 novel SSR loci obtained in this study, are useful in analyzing the relationships and genetic diversity in the cultured and wild GFP populations in Southeast Asia. Furthermore, the study provides a baseline for genetic investigation of GFP resources and practical guidelines for the hatchery industry. A proper management program of local hatchery stocks is required to preserve the genetic diversity for continued and sustainable development of GFP aquaculture.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

LY and XZ designed the research. JL supplied the samples. JF and CC performed the analysis. LY wrote the manuscript.

## FUNDING

This study was funded by the Central Public Interest Scientific Institution Basal Research Fund, CAFS (Nos. 2018SJ-YZ01 and 2019ZX0304), the China-ASEAN Maritime Cooperation Fund (CAMC-2018F), and the Scientific Innovation Fund, PRFRI (No. CZ-2019-04).

## ACKNOWLEDGMENTS

We would like to thank Mr. Yongqing Ye from the Guangdong Daynew Aquatic Sci-Tech Co., Ltd., for his help and Editage (www.editage.cn) for English language editing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00323/full#supplementary-material

**FIGURE S1 |** Clustering analysis of cultured and wild populations of *M. rosenbergii*.

**TABLE S1 |** Genetic variability at seventeen microsatellite loci in seven populations of *M. rosenbergii*.

**SUPPLEMENTARY S1 |** Sequences information of microsatellite markers in GFP transcriptome data.

## REFERENCES

Agarwal, D., Aich, N., Pavan-Kumar, A., Kumar, S., Sabnis, S., Joshi, C. G., et al. (2016). SNP mining in transcripts and concomitant estimation of genetic variation in *Macrobrachium rosenbergii* stocks. *Conserv. Genet. Resour.* 8, 1–10.

Alam, M. M. M., Westfall, K. M., and Palsson, S. (2017). Historical demography and genetic differentiation of the giant freshwater prawn *Macrobrachium rosenbergii* in Bangladesh based on mitochondrial and ddRAD sequence variation. *Ecol. Evol.* 7, 4326–4335. doi: 10.1002/ece3.3023

Chand, V., de Bruyn, M., and Mather, P. B. (2005). Microsatellite loci in the eastern form of the giant freshwater prawn (*Macrobrachium rosenbergii*). *Mol. Ecol. Notes* 5, 308–310. doi: 10.1111/j.1471-8286.2005.00910.x

Chareontawee, K., Poompuang, S., Na-Nakorn, U., and Kamonrat, W. (2007). Genetic diversity of hatchery stocks of giant freshwater prawn (*Macrobrachium rosenbergii*) in Thailand. *Aquaculture* 271, 121–129. doi: 10.1016/j.aquaculture.2007.07.001

Charoentawee, K., Poompuang, S., and Nakorn, U. (2006). Isolation and characterization of microsatellite in giant freshwater prawn *Macrobrachium rosenbergii*. *Mol. Ecol. Notes* 6, 823–825. doi: 10.1111/j.1755-0998.2008.02127.x

De Bruyn, M., Nugroho, E., Hossain, M. M., Wilson, J. C., and Mather, P. B. (2005). Phylogeographic evidence for the existence of an ancient biogeographic barrier: the Isthmus of Kra Seaway. *Heredity* 94, 370–378. doi: 10.1038/sj.hdy.6800613

Dieringer, D., and Schlotterer, C. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol.* 3, 167–169. doi: 10.1046/j.1471-8286.2003.00351.x

Divu, D., Karunasagar, I., and Karunasagar, I. (2008). Microsatellite DNA markers in the giant freshwater prawn, *Macrobrachium rosenbergii*: a tool for genetic analysis. *Mol. Ecol. Resour.* 8, 1040–1042. doi: 10.1111/j.1755-0998.2008.02148.x

Evanno, G. G., Regnaut, S. S., and Gouder, J. J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (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. doi: 10.1111/j.1755-0998.2010.02847.x

Food and Agriculture Organization of the United Nations [FAO] (2018). *Cultured Aquatic Species Fact Sheets: Giant Freshwater Prawn (Macrobrachium rosenbergii)*. Rome: Food and Agriculture Organization of the United Nations.

Fujimura, T., and Okamoto, H. (1972). "Notes on progress made in developing a mass culturing technique for *Macrobrachium rosenbergii* in Hawaii," in *Coastal Aquaculture in the Indo-Pacific Region*, ed. T. V. R. Pilay (West Byfleet: Fishing News Books), 313–327.

Hedgecock, D., Stelmach, D. J., Nelson, K., Lindenfelser, M. E., and Malecha, S. R. (1979). Genetic divergence and biogeography of natural populations of *Macrobrachium rosenbergii*. *Proc. World Maricult. Soc.* 10, 873–879. doi: 10.1111/j.1749-7345.1979.tb00084.x

Jarne, P., and Lagoda, P. J. (1996). Microsatellites, from molecules to population and back. *Trends Ecol. Evol.* 11, 424–429. doi: 10.1016/0169-5347(96)10049-5

Jung, H., Lyon, R. E., Li, Y., Thanh, N. M., Dinh, H., Hurwood, D. A., et al. (2014). A candidate gene association study for growth performance in an improved giant freshwater prawn (*Macrobrachium rosenbergii*) culture line. *Mar. Biotechnol.* 16, 161–180. doi: 10.1007/s10126-013-9555-7

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, selection, and population subdivision. *Conserv. Biol.* 1, 143–158. doi: 10.1111/j.1523-1739.1987.tb00023.x

Lai, F. Y., Tu, P. A., Ding, S. T., Lin, M. J., Chang, S. C., Lin, E. C., et al. (2018). Survey of genetic structure of geese using novel microsatellite markers. *Asian-Australas J. Anim. Sci.* 31, 167–179. doi: 10.5713/ajas.17.0224

Ling, S. W. (1969). *The General Biology and Development of Macrobrachium resonbergii*. Rome: FAO. 589–606.

Liu, Z. J., and Cordes, J. F. (2004). DNA marker technologies and their application in aquaculture genetics. *Aquaculture* 238, 1–37. doi: 10.1016/j.aquaculture.2004.05.027

Manechini, J. R. V., da Costa, J. B., Pereira, B. T., Carlini-Garcia, L. A., Xavier, M. A., Landell, M. G. A., et al. (2017). Unraveling the genetic structure of Brazilian commercial sugarcane cultivars through microsatellite markers. *PLoS One* 13:e0195623. doi: 10.1371/journal.pone.0195623

Miao, W. M., and Ge, X. P. (2002). Freshwater prawn culture in China: an overview. *Aquacult. Asia* 7, 7–12.

Mohanty, P., Sahoo, L., Parida, K., and Das, P. (2013). Development of polymorphic EST-SSR markers in *Macrobrachium rosenbergii* by data mining. *Conserv. Genet. Resour.* 5, 133–136. doi: 10.1007/s12686-012-9751-1

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

New, M. B. (2000). "History and global status of freshwater prawn farming," in *Freshwater Prawn Culture; the Farming of Macrobrachium rosenbergii*, eds M. B. New and W. C. Valenti (Oxford: Blackwell Science), 1–11. doi: 10.1002/9780470999554.ch1

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292. doi: 10.1086/282771

Ng, P. K. L. (1997). The conservation status of freshwater prawn and crabs in Singapore with emphasis on the nature reserves. *Proc. Natl. Res. Surv. Sem. Gard. Bull. Singap.* 49, 267–272.

Park, S. D. E. (2001). *Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection*. [Ph.D.Thesis]. Dublin University: Dublin

Peakall, R. R., and Smouse, P. E. P. (2006). GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6, 288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Peakall, R. R., and Smouse, P. E. P. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Mol. Ecol. Notes* 28, 2537–2549.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–949.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249. doi: 10.1093/oxfordjournals.jhered.a111573

Saitou, N., and Nei, M. (1987). The neighboring-joining method: a new method for reconstructing phylogenetic tress. *Mol. Biol. Evol.* 4, 406–425.

Schneider, K., Tidwell, J., Gomelsky, B., Pomper, K., Waldbieser, G., Saillant, E., et al. (2012). Genetic diversity of cultured and wild populations of the giant freshwater prawn *Macrobrachium rosenbergii* (de Man, 1879) based on microsatellite analysis. *Aquacult. Res.* 44, 1425–1437. doi: 10.1111/j.1365-2109.2012.03147.x

Sun, C., Ye, X., Dong, J., Tian, Y., and Liang, J. (2015). Genetic diversity analysis of six cultured populations of *Macrobrachium rosenbergii* using microsatellite markers. *South China Fish. Sci.* 11, 20–26.

Thanh, H., Liu, Q., Zhao, L., Zhang, H., and Liu, J. (2015). Genetic diversity of the cultured giant freshwater prawn (*Macrobrachium rosenbergii*) in China based on microsatellite markers. *Biochem. Syst. Ecol.* 59, 144–154. doi: 10.1016/j.bse.2014.12.023

Van Oosterhout, C., William, F. H., Derek, P. M. W., and Peter, S. (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Vieira, M. L. C., Santini, L., Diniz, A. L., Munhoz, C., and de, F. (2016). Microsatellite markers: what they mean and why they are so useful. *Genet. Mol. Biol.* 39, 312–328. doi: 10.1590/1678-4685-gmb-2016-0027

Weber, J. L., and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain-reaction. *Am. J. Hum. Genet.* 33, 388–396.

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Yang, G., Frinsko, M., Chen, X., Wang, J., Hu, G., and Gao, Q. (2008). New technique for induced breeding and culture of giant freshwater prawn (part one). *Sci. Fish Farm.* 10, 12–13.

Yang, G., Frinsko, M., Chen, X., Wang, J., Hu, G., and Gao, Q. (2012). Current status of the giant freshwater prawn (*Macrobrachium rosenbergii*) industry in China, with special reference to live transportation. *Aquacult. Res.* 43, 1049–1055. doi: 10.1111/j.1365-2109.2011.03009.x

Yu, L. Y., Bai, J. J., Fan, J. J., Ma, D. M., Quan, Y. C., and Jiang, P. (2015). Transcriptome analysis of the grass carp (*Ctenopharyngodon idella*) using 454 pyrosequencing methodology for gene and marker discovery. *Genet. Mol. Res.* 14, 19249–19263. doi: 10.4238/2015.December.29.35

**Conflict of Interest Statement:** JL author was employed by Foshan Sanshui Baijin Aquatic Seedling Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 02 August 2019
doi: 10.3389/fmars.2019.00494

# Assessment of the Genetic Connectivity Between Farmed Populations on a Typical Kelp Farm and Adjacent Spontaneous Populations of *Saccharina japonica* (Phaeophyceae, Laminariales) in China

*Tifeng Shan[1,2,3]*, Qianxi Li[1,2,3,4], Xuemei Wang[1,2,3,4], Li Su[1,2,3] and Shaojun Pang[1,2,3]**

[1] Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2] Center for Ocean Mega-Science, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [3] Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology, Qingdao, China, [4] University of Chinese Academy of Sciences, Beijing, China

The commercially important brown alga *Saccharina japonica* in China has been believed to be accidentally introduced from Japan in 1920s. Meanwhile, spontaneous populations in the wild are assumed to be derived from the locally farmed populations. Spontaneous populations are often observed in the subtidal zones and on the cultivation infrastructure near farmed populations in the north of China. However, the genetic connectivity between these sympatric spontaneous and farmed populations remains unclear. Here, three commonly farmed cultivars (farmed populations) and three spontaneous populations (two from subtidal zones and one from cultivation rafts) were sampled from a typical kelp farm in Dalian, China, and analyzed with ten polymorphic microsatellite markers. Genetic diversity of farmed populations was found to be higher than that of the subtidal spontaneous populations. Neighbor joining cluster analysis based on genetic distance, Bayesian model-based structure analysis, and discriminant analysis of principal components revealed significant genetic divergence between the farmed populations and the subtidal spontaneous ones. Gene flow out of farmed populations to the subtidal spontaneous populations was revealed to be very limited, but gene flow in the contrary direction was more prominent. The spontaneous sporophytes on the structural rafts contained pedigree from both farmed and subtidal spontaneous populations. Results of this study may help us to understand reciprocal impacts between sympatric spontaneous and farmed populations of *S. japonica.*

**Keywords: kelp, seaweed farming, simple sequence repeat, genetic structure, gene flow**

OPEN ACCESS

***Edited by:***
*Zhijun Dong,*
*Yantai Institute of Coastal Zone Research (CAS), China*

***Reviewed by:***
*Naihao Ye,*
*Yellow Sea Fisheries Research Institute (CAFS), China*
*Andrew Anthony David,*
*Clarkson University, United States*

****Correspondence:***
*Tifeng Shan*
*shantifeng@qdio.ac.cn*
*Shaojun Pang*
*sjpang@qdio.ac.cn*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *10 June 2019*
***Accepted:*** *22 July 2019*
***Published:*** *02 August 2019*

***Citation:***
*Shan T, Li Q, Wang X, Su L and Pang S (2019) Assessment of the Genetic Connectivity Between Farmed Populations on a Typical Kelp Farm and Adjacent Spontaneous Populations of Saccharina japonica (Phaeophyceae, Laminariales) in China. Front. Mar. Sci. 6:494. doi: 10.3389/fmars.2019.00494*

## INTRODUCTION

The important economic kelp species *Saccharina japonica* (J. E. Areschoug) C. E. Lane, C. Mayes, Druehl, and G.W. Saunders has been assumed to be non-native to China, but unintentionally introduced from the north of Japan in 1920s (Tseng and Zhang, 1952; Tseng, 2001). This assumption is supported by the recent genetic diversity analysis of *S. japonica* populations from Japan and China as founder effect is revealed to be evident in Chinese populations (Shan et al., 2017). Originally, small scale cultivation experiments of this kelp was conducted using traditional Japanese method by throwing stones with attached spores to the sea (Hasegawa, 1976; Tseng, 2001). The experiments were first performed in Dalian, and later expanded to Yantai, Weihai and Qingdao in Shandong Province. Original spontaneous (wild) populations of *S. japonica* in rocky shores of these regions were thought to be derived from the zoospores escaping from cultivated individuals (Tseng and Zhang, 1952). Now the southernmost distribution of wild *S. japonica* populations in China is around 36°N latitude (Pang et al., 2007).

Conspicuous spontaneous populations of *S. japonica* are often observed in rocky subtidal zones and on the cultivation infrastructures adjacent to farmed populations in northern China (**Figure 1**). The sympatric distribution of farmed and spontaneous populations provides possibility of gene flow between them. On one hand, farmers and breeders are most concerned with the potential for contamination of wild pedigree into the farmed populations, which may alter the agronomical traits of the latter and thus degrade their quality or/and yield. On the other hand, spontaneous populations play important ecological roles in near shore rocky ecosystem, and may represent specific gene repertoire which can be very valuable materials in breeding programs (Fei, 2004; Zhang et al., 2017). From these perspectives, measures should be taken to prevent gene flow between them. It is therefore very important to evaluate the level of genetic connectivity between these sympatric farmed and spontaneous populations. However, no such studies have been conducted for *S. japonica* in China.

Supposedly at least two alternative prerequisites should be satisfied for the possibility of reciprocal gene flow between farmed and subtidal spontaneous populations. One is overlap of the reproduction season which makes their crossing possible. Otherwise, the alternative one is that zoospores released from the spontaneous or farmed populations are able to colonize the cultivation longlines or subtidal zones, respectively, and grow into mature sporophytes. Subtidal spontaneous sporophytes of *S. japonica* are biennial. They usually reproduce twice in their life history, in autumns when they are 1 year and 2 years old, respectively, and then the sporophytes die away (Tseng and Wu, 1962). In earlier times the "autumn sporeling" method was used, in which seedlings are produced by employing mature subtidal spontaneous individuals as parents (Shan et al., 2011). After it was devised in the 1950s, the "summer sporeling" method was gradually accepted by the cultivation industry and has become the dominant method for large-scale production of seedlings in China (Tseng et al., 1955; Tseng, 2001). The farmed populations derived from this method produce sori in summer (mainly from July to August), different from the reproductive time of subtidal spontaneous populations (Shan et al., 2011). Most of the farmed populations are harvested before July and only hundreds of selected parental sporophytes are hung in deep waters until August for the purpose of seedling production. Even though the zoospores released from these parental sporophytes are able to colonize in the subtidal regions, their derived gametophytes and sporophytes are unlikely to survive the high temperature (up to 27–28°C) in summer (Pang et al., 2007). Strings, on which grow young seedlings about 3–4 cm long and the remnant gametophytes, are transported from hatcheries to the open sea for transitional cultivation when seawater temperature drops below 20°C in late October. The transitional cultivation lasts for 1 month until the young sporophytes reach an average length of 20 cm. The subtidal spontaneous populations reproduce at this period (**Figure 1**). Zoospores released from them are likely to attach to the seedling strings and their derived gametophytes might self-cross or cross with the remnant gametophytes of the farmed populations. However, selection of larger sporophytes in the insertion process for final cultivation is expected to minimize the pedigree contamination from spontaneous populations. According to the above analysis, the genetic connectivity between subtidal spontaneous and farmed populations is expected to be very limited. Spontaneous sporophytes are also often observed growing on the structural rafts of longline cultivation system (**Figure 1**). They usually become conspicuous from January and February. They are hypothesized to be derived from the zoospores released by the subtidal spontaneous populations in autumn.

Microsatellites have been the markers of choice in genetic structure analysis due to its codominant nature and high polymorphism information content per locus owing to its multiallelic nature (Liu and Cordes, 2004). They were successfully used to assess the genetic connectivity between farmed and the adjacent spontaneous populations of *Undaria pinnatifida* on the cultivation infrastructure in our previous study (Shan et al., 2018). By employing ten highly polymorphic microsatellites, our objective in the present study is to assess the genetic connectivity between farmed populations on a typical kelp farm and the adjacent spontaneous populations of *S. japonica* in China.

## MATERIALS AND METHODS

### Sample Collection

Sampling of *S. japonica* were conducted from a typical kelp farm and the adjacent subtidal regions on May 10, 2018 in a bay of Lüshun district, Dalian city, China (**Figure 2**). The sampling range was from 38°47′N, 121°15′E to 38°47′N, 121°16′E. Three commonly farmed cultivars, which are designated as FB, FJ, and XS, were sampled from longlines. All of them were produced through "summer sporeling" method and grown on 8-m long cultivation ropes, with both ends tied to the structural rafts (Figure 1 in Shan et al., 2018). Subtidal spontaneous populations were collected from two rocky locations adjacent to the farmed populations, with the individuals sampled from the west and east regarded as two separate populations, designated as SW

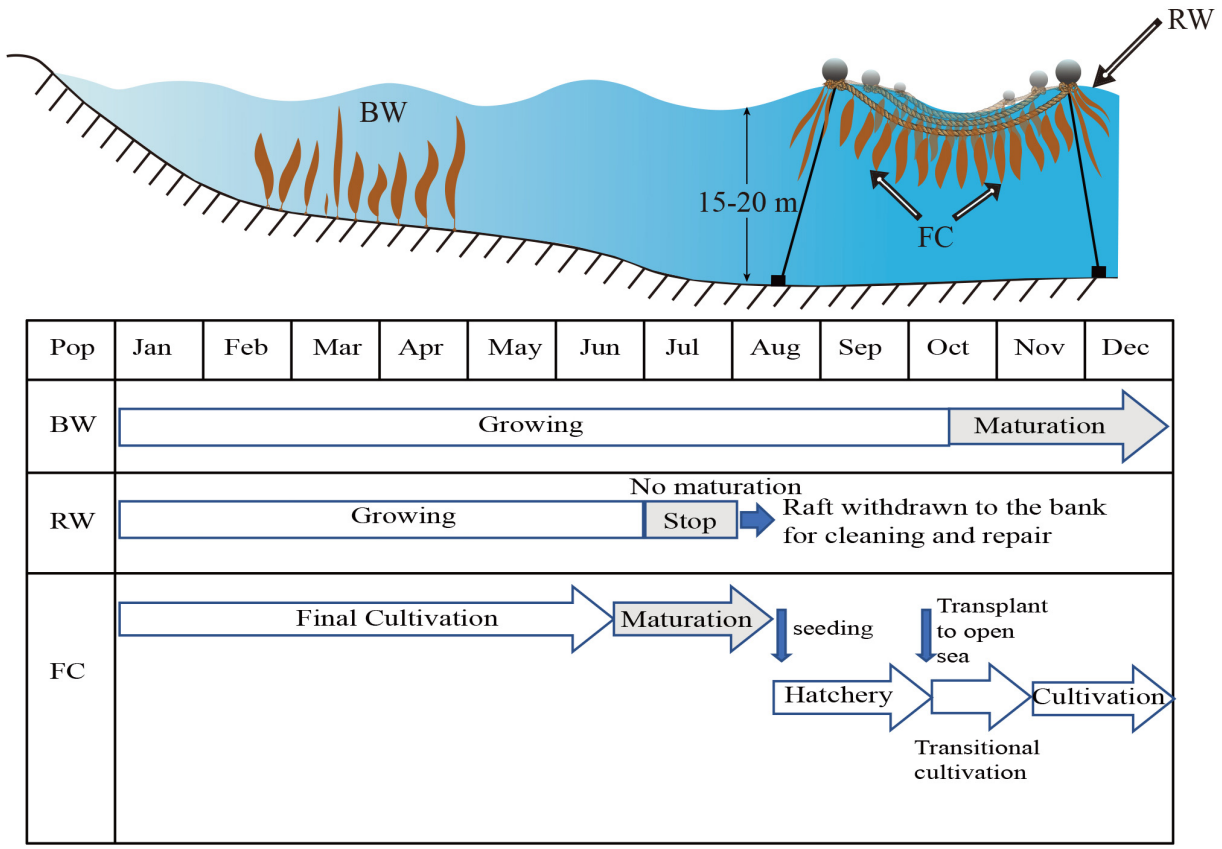


**FIGURE 1 |** Schematic drawing of the sympatric spontaneous populations in subtidal benthic zone (BW) and on the cultivation rafts (RW), and farmed cultivars (FC) of *Saccharina japonica* on a typical kelp farm in Dalian China. Their growing and maturation periods are also shown.

and SE, respectively (**Figure 2**). The distance between them was ca. 400 m. Spontaneous sporophytes were also collected from structural rafts adjacent to the farmed populations and designated as population RW. The distance between any two sampled individuals was at least 3 m to reduce the chance of collecting siblings or close relatives. Thirty individuals were sampled from each population except SE population, which included 24 individuals. A small part of the blade was cut from each individual, cleaned by sterilized seawater, blotted with paper towels, and then dried in silica gels in plastic bags.

## Microsatellite Genotyping

Genomic DNA was isolated using Plant Genomic DNA kit Ca.#DP305 (TIANGEN, Beijing, China).Ten polymorphic microsatellite markers H123 (Shi et al., 2007), Zspj6, Zspj9, Zspj14, Zspj17, Zspj20, Zspj26, Zspj28, Zspj39, and Zspj40 (Zhang et al., 2014) were employed for genetic analysis. The polymerase chain reaction (PCR) and microsatellite genotyping were conducted as described in Shan et al. (2019).

## Data Analysis

The number of alleles ($N_a$), observed and expected heterozygosities ($H_o$ and $H_e$), inbreeding coefficient $F_{is}$ and Nei's standard genetic distance (Nei, 1972) were computed with GenAlEx 6.5 (Peakall and Smouse, 2006, 2012). The deviation from Hardy–Weinberg equilibrium (HWE) was evaluated by GENEPOP version 4.7.0 (Rousset, 2008). The adjusted $P$-value $< 0.05$ after Bonferroni correction (Rice, 1989) was regarded to be significant. The population selfing rate was estimated by the index $g_2$ using the *inbreedR* package (Stoffel et al., 2016). A neighbor-joining (NJ) unrooted phylogenetic tree was constructed with calculated genetic distance between populations by the POPTREE software (Takezaki et al., 2010). During bootstrapping, 1000 permutations were performed to assess the robustness of the clusters. The $F_{st}$ values were computed by ARLEQUIN version 3.11 with 1000 permutations to assess pairwise population genetic differentiation (Laurent et al., 2005). To adjust for multiple comparisons, the false discovery rate (FDR) was controlled by using "BH" method (Benjamini and Hochberg, 1995) in the *p. adjust* function in R. The adjusted $P$-value $< 0.05$ was considered to be significant.

A Bayesian model-based clustering analysis was conducted with STRUCTURE 2.3.4 to evaluate the most possible number of genetic clusters (Pritchard et al., 2000). This clustering method was applied in identification of genetically distinct subpopulations based on the allele frequencies.

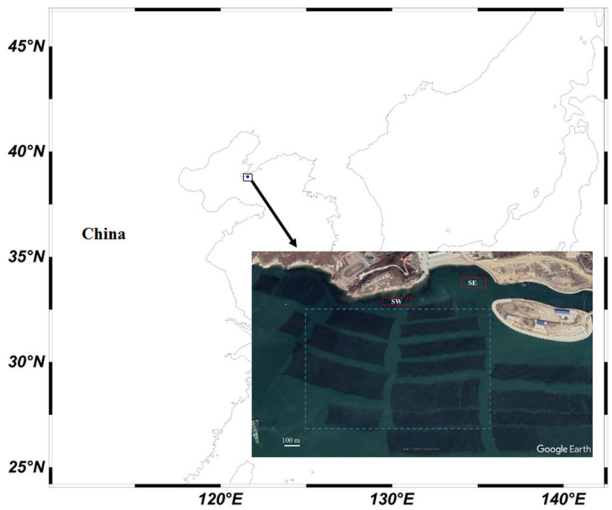


**FIGURE 2 |** Sampling site of the farmed populations and the adjacent spontaneous populations of *Saccharina japonica.* Red rectangles indicate the sampling range of the subtidal spontaneous populations SW and SE. Blue dashed rectangle indicates the sampling range of farmed populations and the spontaneous population on the cultivation structural rafts. The dark rectangular areas in the sea are the cultivation infrastructure.

The number of clusters ($K$ value) was set from 1 to 6, and 20 independent runs were carried out for each fixed $K$ value using the admixture model and allele frequencies correlated model. Each run included a burn-in length of 100,000 followed by 1,000,000 Mote Carlo Markov Chain (MCMC) repetitions. The optimal $K$ value was determined by submitting all results files of $K$ = 1–6 to STRUCTURE HARVESTER (Earl and Vonholdt, 2012) according to the method of Evanno et al. (2005). The run with the highest Ln Pr (X| $K$) value (log probability) among the 20 independent runs was chosen and the graphical result was displayed with DISTRUCT 1.1 (Rosenberg, 2004). A multivariate discriminant analysis of principal components (DAPC) was also performed using the *adegenet* package in R in order to infer genetic relationships among populations (Jombart, 2008; Jombart et al., 2010).

**TABLE 1 |** Genetic diversity and selfing rates of the populations of *Saccharina japonica* from a typical farm and the adjacent subtidal zones in Dalian, China.

| Population | $N_a$ | $H_e$ | $F_{is}$ | Selfing rate ($g_2$) | $P$ ($g_2$ = 0) |
|---|---|---|---|---|---|
| FB | 2.9 | 0.400 | 0.063 | 0.013 | 0.354 |
| FJ | 3.8 | 0.496 | 0.039 | −0.007 | 0.617 |
| XS | 3.5 | 0.477 | −0.023 | −0.041 | 0.964 |
| RW | 3.3 | 0.429 | −0.001 | 0.071 | 0.036 |
| SE | 2.8 | 0.389 | −0.022 | 0.042 | 0.211 |
| SW | 2.6 | 0.328 | 0.038 | 0.062 | 0.149 |

*$N_a$ number of alleles, $H_e$ expected heterozygosity, $F_{is}$ inbreeding coefficient, Selfing rate ($g_2$) indicates selfing rate estimated from $g_2$ in inbreedR package, $P$ ($g_2$ = 0) the probability value for $g_2$ = 0.*

## RESULTS

The data of microsatellite genotyping and allele frequencies was provided in **Supplementary Tables S1**, **S2**, respectively. Both the average $N_a$ and $H_e$ across the 10 microsatellite loci were found to be highest in FJ ($N_a$ = 3.8, $H_e$ = 0.496) and lowest in SW ($N_a$ = 2.6, $H_e$ = 0.328) (**Table 1**). The genetic diversity of farmed populations was higher than that of the subtidal spontaneous populations in terms of average $N_a$ and $H_e$. No significant deviation from HWE was detected for all populations either at each locus or across all the 10 loci (**Table 1** and **Supplementary Table S3**). The average $F_{is}$ values across loci were close to zero for all populations (**Table 1**). The selfing rate $g_2$ was only detected to be significant in RW ($P$ = 0.036).

$F_{st}$ values were detected to be significant among all populations except that between FJ and XS (**Table 2**). Values of both $F_{st}$ and Nei's genetic distance showed that the genetic differentiation between the farmed and the subtidal spontaneous populations was much higher than those within the farmed or subtidal spontaneous populations. The NJ phylogenetic tree, which was constructed based on pair-wise genetic distance, exhibited that SW and SE were grouped into a cluster, and FJ, XS, and FB were grouped into another cluster (**Figure 3**). These two clusters were clearly separated with large distance. RW was shown to be intermediate between them.

The most likely number of $K$ was identified to be 2 by the use of STRUCTURE HARVESTER according to the $\Delta K$ values (**Figure 4** and **Supplementary Figure S1**). Subtidal spontaneous populations were assigned to the first cluster

**TABLE 2 |** Pairwise genetic distance (below diagonal) and $F_{st}$ values (above diagonal) in six populations of *Saccharina japonica* from a kelp farm and the adjacent subtidal zones in Dalian, China.

| Population | FB | FJ | SE | SW | RW | XS |
|---|---|---|---|---|---|---|
| FB | | 0.042** | 0.127** | 0.177** | 0.041** | 0.057** |
| FJ | 0.047 | | 0.106** | 0.128** | 0.032** | 0.009 |
| SE | 0.116 | 0.115 | | 0.023* | 0.026** | 0.146** |
| SW | 0.143 | 0.114 | 0.023 | | 0.048** | 0.155** |
| RW | 0.043 | 0.042 | 0.032 | 0.039 | | 0.050** |
| XS | 0.061 | 0.025 | 0.159 | 0.140 | 0.059 | |

**and ** indicate significance at $P$ < 0.05 and $P$ < 0.01 level for $F_{st}$ values, respectively.*

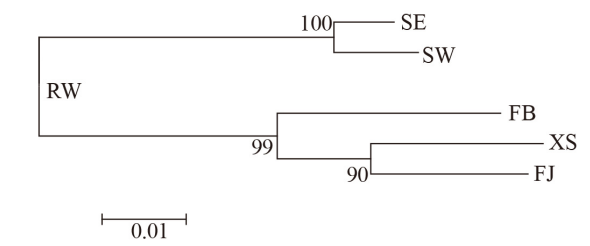


**FIGURE 3 |** Neighbor-joining phylogenetic tree based on genetic distance among the farmed and spontaneous populations of *Saccharina japonica* from a kelp farm in China. The bar indicates the genetic distance.

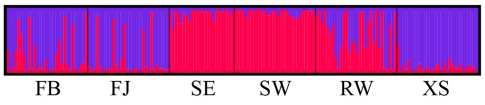


**FIGURE 4 |** Genetic structure revealed by the Bayesian model-based analysis using STRUCTURE 2.3.4 for farmed and spontaneous populations of *Saccharina japonica* from a kelp farm in China. Each individual is indicated by a vertical colored bar, and the proportion of the color in each bar represents the probability of membership in the corresponding cluster.

with high proportion of membership (>0.90), with only three individuals containing a proportion of membership >20% from farmed populations. Most individuals of the farmed populations were assigned to the second cluster, but some individuals of FB and FJ possessed high proportion of membership originated from the subtidal spontaneous populations. The total proportion of membership originated from subtidal spontaneous populations was 0.28 and 0.19 in FB and FJ, respectively. By comparison, XS contained less membership derived from subtidal spontaneous populations (<0.10). Admixture of membership was most marked in RW, with the membership proportion from the first and second clusters being 0.62 and 0.39, respectively. The genetic relationships among populations revealed by the DAPC analysis were very consistent with those revealed by STURCTURE and the NJ phylogeny (**Figure 5**).

## DISCUSSION

As expected, farmed populations were found to have exerted very limited impact on the genetic composition of the sympatric subtidal spontaneous populations. This was consistent with the results found by Zhang et al. (2017), although the spontaneous population appeared to be far from the farmed populations investigated in that study. The subtidal spontaneous populations are therefore speculated to evolve almost without the interference of the farmed population under the current farming system. They are potentially valuable breeding materials for they can integrate such characteristics as tolerance to high temperature that has evolved to adapt to the local environment (Pang et al., 2007). However, the gene flow from subtidal spontaneous populations to the farmed ones was revealed to be larger than we had expected. As is expected in the section "Introduction," the zoospores discharged by the subtidal spontaneous populations may attach to the seedling ropes during the 1-month transitional cultivation of the young sporelings in the open sea (**Figure 1**). The derived gametophytes can produce sporophytes through selfing or crossing with the remnant gametophytes of the farmed cultivars on the ropes. On one hand, growth of some of these sporophytes are likely to catch up with that of the earlier born larger sporelings during the 1-month period, and thus are exploited in the sporeling-insertion process for final farming. On the other hand, because some sporelings usually detach from the cultivation rope after insertion due to the loose attachment of the holdfast, farmers need to supplement them by inserting new sporelings from the back-up, and thus the smaller sporophytes will obtain the opportunity of being used. These scenarios may account for the relatively high level of gene flow from subtidal spontaneous populations to the farmed ones. The existence of more gene flow from subtidal spontaneous populations to the farmed ones is likely to be one reason that results in higher genetic diversity in farmed populations.

What is also unexpected is that the spontaneous population on the structural rafts (RW) was revealed to contain a large proportion of pedigree from farmed populations (39%), in addition to that from subtidal spontaneous populations. From June to July, zoospores released from the mature farmed sporophytes may attach to the cultivation infrastructure. After the harvest of *S. japonica* is finished, however, the cultivation infrastructure including the cultivation ropes and structural rafts are withdrawn to the bank for cleaning and repair. The attached spores and gametophytes are impossible to survive this long period of drought. Furthermore, the sporophytes of RW were observed to be conspicuous from January and February. So they were supposedly established in last November and December. But during this period there were no mature farmed sporophytes on longlines. One potential explanation is that they were derived from the farmed sporophytes that detached from the cultivation rope and sank in the sea bottom. From the very beginning of transitional cultivation of young sporelings to the time when they grow into adults, many of

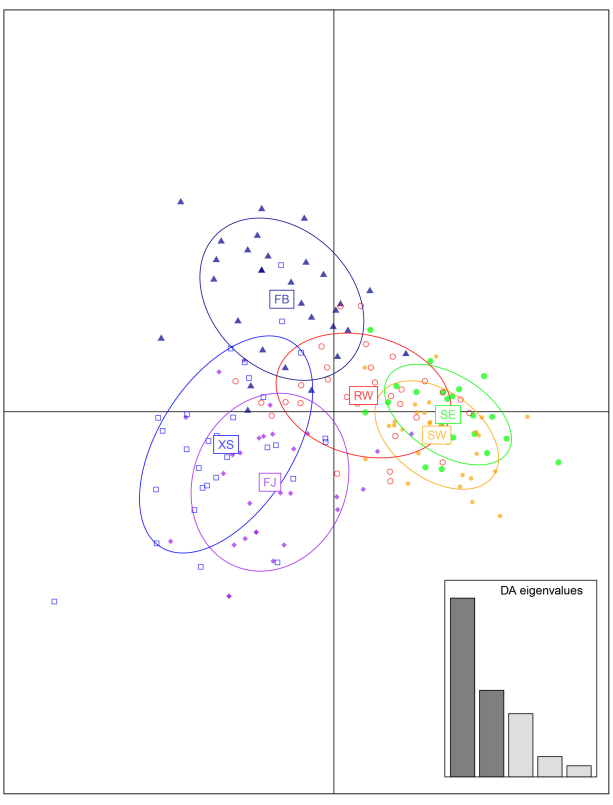


**FIGURE 5 |** Discriminant analysis of principal components for farmed and spontaneous populations of *Saccharina japonica* from a kelp farm in China.

sporophytes detach from the ropes due to currents, storm and unsteady fastening of their hapteras. The detachment rate of the adult sporophytes was especially higher, up to 50% at the most severe situation in recent years (M. F. Zhang, 2015–2019). Some of them might be tangled with the substratum or the flora at the bottom of the sea, and became sessile. Some might go back and forth near the farm under driving force of the tides and currents. It has been revealed that low temperature and low light intensity can delay the mature time of sporophytes (Tseng and Wu, 1962; Mizuta et al., 1999). This principle has been exploited by farmers to adjust the reproductive time in farming and seedling production practice. In the north of China, farmers usually select desirable sporophytes as parental individuals during the harvest season and hang them deeply in remote areas where the currents are fast, and both the temperature and light intensity are lower. In the south of China (mainly in Fujian) where the high summer temperature makes it unfeasible to preserve parental sporophytes in the sea, the selected parental sporophytes are transported to hatchery rooms, and low temperature and light are given (Pang et al., 2007). Likewise, the sporophytes that have sunken in the sea bottom experience poor light and low temperature, and thus their mature time is likely to be delayed to autumn, similar to that of the subtidal spontaneous populations. These detached sporophytes could also be potential media for the gene flow from the farmed populations to the subtidal spontaneous ones, which may account for the small amount of pedigree of the former detected in the latter in this study.

The genetic diversity of the farmed and spontaneous populations was consistent with those detected in the previous studies (Li et al., 2017; Shan et al., 2017; Zhang et al., 2017). Interestingly, the subtidal spontaneous populations exhibited lower genetic diversity than the farmed populations. As suggested by Zhang et al. (2017), introduction of alien germplasm, intra- and inter-specific crossing in the breeding process might increase the genetic diversity of farmed populations. By comparison, the subtidal spontaneous populations have supposedly been isolated from immigrants since their original foundation. Reduction in genetic diversity is potentially caused by founder effect and the subsequent genetic drift. Meanwhile, the three farmed cultivars were demonstrated to be genetically close to each other. As was suggested and revealed in our previous studies, mixing among different cultivars were most likely caused by sharing of refrigerated seawaters during seedling production in the hatchery (Shan et al., 2011; Li et al., 2017). The result was the homogenization of genetic composition and agronomical features of the cultivars.

In conclusion, very significant genetic divergence was demonstrated to exist between the farmed populations on a typical kelp farm and the adjacent subtidal spontaneous populations of *S. japonica* in China. Gene flow out of farmed populations to the subtidal spontaneous populations was revealed to be very limited, but gene flow in the contrary direction was more prominent. The pedigree of spontaneous sporophytes occurring on the structural rafts was exhibited to be admixture of both farmed and subtidal spontaneous populations. This study provides useful information for management of breeding programs and conservation of stock resources. In the long run, genetic diversity of farmed and spontaneous populations, and the genetic connectivity between them ought to be monitored continuously for the benefit of sustainable development of cultivation industry of *S. japonica*.

## DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

## AUTHOR CONTRIBUTIONS

SP and TS conceived the study. TS, QL, XW, and LS did the sampling work in the field and the analysis in the lab. All authors contributed to the writing of the manuscript.

## FUNDING

This research was supported by grants from China Agriculture Research System (CARS-50), the State's Key Project of Research and Development Plan (2016YFC1402507), the National Key Technology Support Program (2015BAD13B05), a special supporting program for maintaining seaweed stock culture from the Chinese Academy of Sciences (CZBZX-1), the National Infrastructure of Fishery Germplasm Resource (2019DKA30470), the Taishan Scholar Program of Shandong Province, and Huiquan Scholar Program of Institute of Oceanology, Chinese Academy of Sciences.

## ACKNOWLEDGMENTS

We are grateful to Mingfu Zhang for his kind help in samples collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00494/full#supplementary-material

**FIGURE S1 |** Delta-K graph obtained by submitting all results files of $K$ = 1–6 to STRUCTURE HARVESTER.

**TABLE S1 |** Microsatellite genotyping data at 10 loci in six populations of *Saccharina japonica*.

**TABLE S2 |** Allele frequencies at 10 microsatellite loci in six populations of *Saccharina japonica*.

**TABLE S3 |** Genetic diversity of the farmed and spontaneous populations of *Saccharina japonica* from a kelp farm in China at each microsatellite locus.

## REFERENCES


Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Earl, D. A., and Vonholdt, B. M. (2012). Structure harvester: a website and program for visualizing structure output and implementing the evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Fei, X. G. (2004). Solving the coastal eutrophication problem by large scale seaweed cultivation. *Hydrobiologia* 512, 145–151. doi: 10.1007/978-94-007-0944-7_19

Hasegawa, Y. (1976). Progress of *Laminaria* cultivation in Japan. *J. Fish. Board Can.* 33, 1002–1006. doi: 10.1139/f76-127

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94

Laurent, E., Guillaume, L., and Stefan, S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50.

Li, X., Pang, S. J., and Shan, T. F. (2017). Genetic diversity and population structure among cultivars of *Saccharina japonica* currently farmed in northern China. *Phycol. Res.* 65, 111–117. doi: 10.1111/pre.12167

Liu, Z. J., and Cordes, J. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238, 1–37. doi: 10.1016/j.aquaculture.2004.05.027

Mizuta, H., Nimura, K., and Yamamoto, H. (1999). Inducible conditions for sorus formation of the sporophyte discs of *Laminaria japonica* areschoug (Phaeophyceae). *Fish. Sci.* 65, 104–108. doi: 10.2331/fishsci.65.104

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292.

Pang, S. J., Jin, Z. H., Sun, J. Z., and Gao, S. Q. (2007). Temperature tolerance of young sporophytes from two populations of *Laminaria japonica* revealed by chlorophyll fluorescence measurements and short-term growth and survival performances in tank culture. *Aquaculture* 262, 493–503. doi: 10.1016/j.aquaculture.2006.11.018

Peakall, R., and Smouse, P. E. (2006). genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol. Ecol. Resour.* 6, 288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225. doi: 10.1111/j.1558-5646.1989.tb04220.x

Rosenberg, N. A. (2004). Distruct: a program for the graphical display of population structure. *Mol. Ecol. Resour.* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Rousset, F. (2008). Genepop'007: a complete re-implementation of the genepop software for windows and linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Shan, T., Liu, F., Liu, Q., and Pang, S. (2011). Review and prospect of "summer sporeling" technique of *Saccharina japonica* in China. *J. Agric. Sci. Technol.* 13, 129–134. (in Chinese with English abstract)

Shan, T., Pang, S., Wang, X., Li, J., and Su, L. (2018). Assessment of the genetic connectivity between farmed and wild populations of *Undaria pinnatifida* (Phaeophyceae) in a representative traditional farming region of China by using newly developed microsatellite markers. *J. Appl. Phycol.* 30, 2707–2714. doi: 10.1007/s10811-018-1449-1447

Shan, T., Yotsukura, N., and Pang, S. (2017). Novel implications on the genetic structure of representative populations of *Saccharina japonica* (Phaeophyceae) in the northwest pacific as revealed by highly polymorphic microsatellite markers. *J. Appl. Phycol.* 29, 631–638. doi: 10.1007/s10811-016-0888-882

Shan, T. F., Pang, S. J., Wang, X. M., Li, J., Su, L., Schiller, J., et al. (2019). Genetic analysis of a recently established *Undaria pinnatifida* (Laminariales: Alariaceae) population in the northern wadden sea reveals close proximity between drifting thalli and the attached population. *Eur. J. Phycol.* 54, 154–161. doi: 10.1080/09670262.2018.1532116

Shi, Y., Yang, G., Liu, Y., Liao, M., Li, X., and Cong, Y. (2007). Development of 18 polymorphic microsatellite DNA markers of *Laminaria japonica* (Phaeophyceae). *Mol. Ecol. Notes* 7, 620–622. doi: 10.1111/j.1471-8286.2006.01652.x

Stoffel, M. A., Esser, M., Kardos, M., Humble, E., Nichols, H., David, P., et al. (2016). inbreedR: an R package for the analysis of inbreeding based on genetic markers. *Methods Ecol. Evol.* 7, 1331–1339. doi: 10.1111/2041-210x.12588

Takezaki, N., Nei, M., and Tamura, K. (2010). Poptree2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Mol. Biol. Evol.* 27, 747–752. doi: 10.1093/molbev/msp312

Tseng, C., Sun, K., and Wu, C. (1955). On the cultivation of Haitai (*Laminaria japonica* Aresch.) by summering young sporophytes at low temperature. *Acta Bot. Sin.* 4, 255–264. (in Chinese with English abstract)

Tseng, C. K. (2001). Algal biotechnology industries and research activities in China. *J. Appl. Phycol.* 13, 375–380.

Tseng, C. K., and Wu, C. Y. (1962). *Mariculture of Saccharina Japonica*. Beijing: Beijing Science Press, (in Chinese).

Tseng, C. K., and Zhang, J. F. (1952). Economic seaweeds in the north of China. *J. Shandong Univ.* 2, 57–82. doi: 10.1111/j.1529-8817.2012.01229.x (in Chinese)

Zhang, J., Wang, X., Yao, J., Li, Q., Liu, F., Yotsukura, N., et al. (2017). Effect of domestication on the genetic diversity and structure of *Saccharina japonica* populations in China. *Sci. Rep.* 7:42158. doi: 10.1038/srep42158

Zhang, L., Peng, J., Li, X., Liu, Y., Cui, C., Wu, H., et al. (2014). Development of 27 trinucleotide microsatellite markers for *Saccharina japonica* using next generation sequencing technology. *Conserv. Genet. Resour.* 6, 341–344. doi: 10.1007/s12686-013-0089-80


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 03 September 2019
doi: 10.3389/fmars.2019.00515

# Seasonal Fish Assemblage Structure Using Environmental DNA in the Yangtze Estuary and Its Adjacent Waters

*Hui Zhang*[1,2,3,4]*, *Susumu Yoshizawa*[4], *Wataru Iwasaki*[4,5] *and Weiwei Xian*[1,2,3]*

[1] CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2] Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, [3] Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China, [4] Atmosphere and Ocean Research Institute, The University of Tokyo, Chiba, Japan, [5] Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan

The fisheries resources of the Yangtze Estuary and its adjacent waters have undergone dramatic declines as a consequence of environmental changes and human activities, with traditional ecological investigations demonstrating progressive decreases in species diversity and annual numbers in the fisheries resource. Environmental DNA (eDNA) technology has been demonstrated as an effective tool by many studies for detecting fish species, monitoring fish biodiversity, and indicating the abundance of fish. In the present study, we chose the Yangtze Estuary and its adjacent waters as a primary research area to investigate fish assemblage structure using eDNA technology. A total of 50 eDNA samples were collected in the estuary in 2018. The results showed that 41 operational taxonomic units were identified from three seasons, with 18, 12, and 33 fish species associated specifically with spring, summer and autumn, respectively. The fish assemblage differed significantly among seasons. *Canonical correspondence analysis* showed that water temperature, salinity, and dissolved oxygen were the main environmental factors affecting structure of the seasonal assemblages. Results of the present study indicate that eDNA technology can be an effective tool not only for fisheries monitoring, but might also importantly assist marine resources conservation, sustainable exploitation of fisheries, the aquatic products processing industry, eco-friendly development, and socioeconomic stability.

**Keywords: Yangtze Estuary, marine fish, assemblage structure, environmental DNA, MiSeq, seasonal assemblages**

OPEN ACCESS

***Edited by:***
*Zhiqiang Han,*
*Zhejiang Ocean University, China*

***Reviewed by:***
*Xiaoshou Liu,*
*Ocean University of China, China*
*Xiubao Li,*
*Hainan University, China*
*Sher Khan Panhwar,*
*University of Karachi, Pakistan*

***Correspondence:**
*Hui Zhang*
*zhanghui@qdio.ac.cn*
*Weiwei Xian*
*wwxian@qdio.ac.cn*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *28 May 2019*
***Accepted:*** *05 August 2019*
***Published:*** *03 September 2019*

***Citation:***
*Zhang H, Yoshizawa S, Iwasaki W and Xian W (2019) Seasonal Fish Assemblage Structure Using Environmental DNA in the Yangtze Estuary and Its Adjacent Waters. Front. Mar. Sci. 6:515. doi: 10.3389/fmars.2019.00515*

## INTRODUCTION

The Yangtze Estuary and its adjacent waters is an important spawning and nursery ground for many commercial fish species, and is also a large and important bait field for many species in summer and autumn (Luo and Shen, 1994). Additionally, the Yangtze Estuary is the only migration channel of the endangered species such as Chinese sturgeon, roughskin sculpin, and reeves shad. Thus, the Yangtze Estuary occupies an important ecological and economical position (Luo and Shen, 1994). In recent years, with the environmental changes and human activities, fisheries resources of the

Yangtze Estuary seriously decline, and the species and numbers of fisheries resources based on the traditional ecological investigation is decreasing year by year (Zhang et al., 2015, 2016, 2019).

Environmental DNA (eDNA) refers to DNA that can be extracted from environmental samples without first isolating any target organisms (Taberlet et al., 2012). Ficetola et al. (2008) were the first to report on eDNA. By collecting eDNA samples, based on genetic analysis methods (e.g., next-generation sequencing technologies), it is possible to detect species that appear in the survey area. Recently, research conducted in various aquatic environments and concerning different species have verified the effectiveness and sensitivity of eDNA in species detection. For example, a large study conducted in the United States by the Water Conservation Center developed eDNA technology for the detection of silver carp *Hypophthalmichthys molitrix* and bighead carp *H. nobilis* in waterways connecting to the Great Lakes (Jerde et al., 2011). Goldberg et al. (2011) demonstrated the use of eDNA techniques involving DNA extraction and PCR methods for the detection of Idaho giant salamander *Dicamptodon aterrimus* and Rocky Mountain tailed frog *Ascaphus montanus* in source-water streams. Thomsen et al. (2012) showed that eDNA techniques can be effective for monitoring crustaceans, aquatic insects, aquatic mammals, birds, and terrestrial mammals in freshwater environments. Since 2012, there has been a significant increase in articles on the use of eDNA for the detection of vertebrates, including several review papers, notably those of Bohmann et al. (2014), Barnes and Turner (2016), and Rees et al. (2014), plus at least 12 papers on biosafety as mentioned by Goldberg et al. (2015). Together, these studies provide a necessary foundation for further using eDNA technology for detecting fish species and monitoring fish biodiversity in marine environments.

Marine fish stock surveys are usually carried out using sampling with nets, which is most reliable when monitoring highly abundant species but has less probability of catching valued though low-abundance fish or endangered species. Thus, as many wild fish resources continue to decrease, traditional resource survey methods may lack reliability (Magnuson et al., 1994). Numerous studies have now shown that the eDNA method can be more sensitive and efficient than traditional survey methods. For instance, Dejean et al. (2012) investigated two traditional bullfrog species in a study of alien species invasion, and compared a traditional survey method with eDNA technology; their results showed that eDNA was the more convenient and effective choice. Takahara et al. (2012) developed a method using eDNA technology for estimating the biomass of carp in experimental environments and in waters in the wild, and concluded that eDNA technology might be aptly applied to estimations of fish biomass in the wild. Davy et al. (2015) used eDNA to monitor eight species of freshwater turtles, thereby demonstrating the method's applicability to potential investigations and monitoring of endangered species. Sigsgaard et al. (2015) monitored a near-extinct species of loach in Danish waters using both traditional survey methods and eDNA techniques; the results showed that the eDNA method was reliable and cost less. Smart et al. (2016) likewise concluded that the eDNA method would be more cost-efficient than traditional methods for the purposes of aquatic taxonomy, and they explored the costs of optimizing eDNA technology. Evans et al. (2017) compared the advantages and disadvantages of the electroshock method with eDNA metabarcoding for quantifying brook trout *Salvelinus fontinalis*, showing that the eDNA method was more time- and labor-efficient (saving about 67% of the costs); thereby suggesting that eDNA can valuably complement other methods of investigation.

Spatial and temporal variation in fish assemblages of the Yangtze Estuary has been widely studied (Yang et al., 1990; Zhu et al., 2002; Zhong et al., 2007; Zhang et al., 2015, 2016, 2019). The changing environmental status of the Yangtze Estuary has compelled numerous studies of the seasonal variation in species composition and biodiversity, including the characteristics of the region's ichthyoplankton assemblage structure and its relationship with environmental factors, such as water temperature, depth, dissolved oxygen, and salinity.

This study utilized water samples and data on environmental conditions collected during three cruises on the Yangtze Estuary and its adjacent waters, in 2018, to characterize the estuary's fish assemblage. Especially, we applied eDNA technology to determine species composition and biodiversity, and thereby further reveal relationships between the spatial-temporal distribution patterns in the fish assemblage and prevailing environmental factors. To our knowledge, this investigation is the first attempt to use eDNA in the Yangtze Estuary and its adjacent waters for the benefit of fishery science in the region. The results should provide an improved scientific basis for management and sustainable utilization of the estuary's fisheries resources.

# MATERIALS AND METHODS

## Sample Collection

A total of 50 eDNA samples were collected from the Yangtze Estuary and its adjacent waters (30°45′-32°00′ N, 121°00′-123°20′ E), during spring, summer and autumn cruises, in 2018. We collected 15 eDNA samples in spring and summer, and 20 in autumn. Briefly, 2-L water samples were collected at the surface using a bucket, and from bottom waters using a van Dorn sampler (**Table 1** and **Figure 1**). Water samples were immediately filtered on the research vessel using a Sterivex$^{TM}$-GP filter unit without a filling bell (pore size 0.22 μm; EMD Millipore Corp.). To minimize cross-contamination, the filter funnels and measuring cups were bleached after every filtration, and 50 artificial seawaters were filtered. We then added DNA preservation buffer (Tiandz Inc.) into the filter unit. Total eDNA was extracted from each filter using a DNeasy Blood and Tissue Kit (Qiagen). To check for cross-contamination during eDNA extraction, eDNA was simultaneously extracted from deionized water. These eDNA samples and negative-control samples were obtained specifically for this study.

Real-time data on the environmental parameters of the water column were measured, including water temperature (T), salinity (S), total nitrogen (TN), total phosphorus (TP), pH, suspended matter (SPM), depth (D), dissolved oxygen (DO), and chemical oxygen demand (COD). Collection of all

data followed the guidelines in "Specification of Oceanographic Investigation" (GB12763-2007).

## Paired-End Library Preparation on the MiSeq Platform

Amplicon libraries of partial 12S rRNA genes were obtained by PCR amplification using the universal primer pairs for fish eDNA, MiFish-U/E (Miya et al., 2015). The first PCR was performed using the two universal primer pairs. The total reaction volume was 12 μl and comprised 6.0 μl of 2× KAPA HiFi HotStart ReadyMix, 3.6 pmol of each MiFish primer, 1 μl of template, and water. The thermal-cycle profile was: 95°C for 3 min; 35 cycles of 98°C for 20 s, 65°C for 15 s, and 72°C for 15 s; and 72°C for 5 min. The first PCR products were diluted 10 times using Milli-Q water, and used as a template for the following PCR. The second PCR was performed to add MiSeq adaptor sequences and 8 bp index sequences to both amplicon ends. The total reaction volume of the second PCR was also 12 μl, comprising 6.0 μl of 2× KAPA HiFi HotStart ReadyMix, 3.6 pmol each of the forward and reverse primers, 1 μl of template, and water. The thermal-cycle profile for the second PCR was: 95°C for 3 min; 12 cycles of 98°C for 20 s, and 72°C for 30 s; and 72°C for 5 min. PCR amplifications were performed in triplicate for each eDNA sample. As a result, three replications of a single eDNA sample had different index sequences, allowing us to assess whether PCR replication increased the number denoting the species detected. All the indexed PCR products were pooled in an equal volume and the pooled libraries were purified by agarose gel electrophoresis. Finally, the libraries were sequenced using an Illumina MiSeq v2 Reagent Kit for 2 × 150 bp paired-end reads (Illumina, San Diego, CA, USA). We acknowledge that all samples analyzed in the present study were sequenced on a single MiSeq run, but that samples analyzed for other research projects were simultaneously sequenced on this run. The total number of reads obtained from the run was 2,308,620 for spring, 1,475,304 for summer, and 4,320,556 for autumn.

**TABLE 1 |** Sampling location.

| Station | Longitude | Latitude | Spring | Summer | Autumn |
|---|---|---|---|---|---|
| 01 | 122.00 | 30.75 | + | + | + |
| 02 | 122.17 | 30.75 | + | + | + |
| 03 | 122.33 | 30.75 | + | + | + |
| 04 | 122.00 | 31.00 | + | + | + |
| 05 | 122.17 | 31.00 | + | + | + |
| 06 | 122.33 | 31.00 | + | + | + |
| 07 | 122.50 | 31.00 | + | + | + |
| 08 | 122.67 | 31.00 | + | + | + |
| 09 | 122.00 | 31.13 | + | + | + |
| 10 | 122.17 | 31.25 | + | + | + |
| 11 | 122.33 | 31.25 | + | + | + |
| 12 | 122.50 | 31.25 | + | + | + |
| 13 | 122.67 | 31.25 | + | + | + |
| 14 | 122.50 | 31.50 | + | + | + |
| 15 | 122.67 | 31.50 | + | + | + |
| 16 | 122.67 | 31.75 | | | + |
| 17 | 123.00 | 31.75 | | | + |
| 18 | 123.33 | 31.75 | | | + |
| 19 | 123.00 | 31.50 | | | + |
| 20 | 123.00 | 31.25 | | | + |

## Data Quality Control and Reads Assembly

Using the program FastQC (Andrews, 2010), the tails of each MiSeq read were trimmed until the Phred quality score (related to the base-calling accuracy) of the last base was ≥20. The paired-end reads (R1 and R2 in the MiSeq platform) were then assembled using the program Flash (Magoč and Salzberg, 2011) when read pairs overlapped by more than 9 bp; reads that could not be assembled were discarded. Next, we discarded reads containing ambiguous bases (Ns). After that, because the expected amplicon length (target region + 127 bp of the first PCR primer sequences) was 297 ± 25 bp, according to comparisons of fish 12S rRNA gene sequences, reads with sequence lengths outside the range 272–322 bp were similarly discarded. In addition, chimeric reads were searched and removed using UCHIME (Edgar et al., 2011). Finally, primer sequences were removed from each read using TagCleaner. In this process, we allowed for mismatches in <4 bases in the search for primer

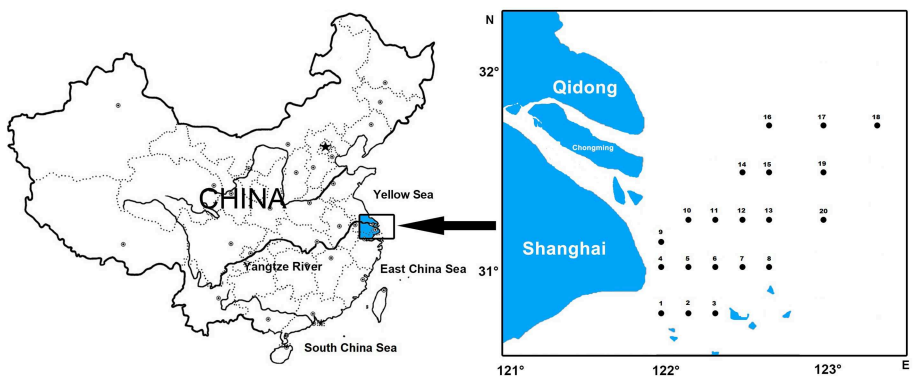


**FIGURE 1 |** Sampling location.

sequences, for two reasons: the sequence of MiFish-U/E primer sets show a two-base difference in forward primers and a one-base difference in reverse primers; and, PCR can amplify fish 12S rRNA sequences even if the sequences have a few mismatches in the primer binding sites. When primer sequences were not found, the read was discarded. This data processing was implemented by the MitoFish database (Iwasaki et al., 2013) and MiFish pipeline (Sato et al., 2018; available at http://mitofish.aori.u-tokyo.ac.jp/).

## Fish Species Identification

We used the same pipeline program mentioned above for taxonomic assignment of the obtained sequences. Before taxonomic assignment, MiSeq reads with an identical sequence (97% sequence similarity, E-value = $10^{-5}$) were assembled using UCLAST (Edgar et al., 2011), and assembled sequences with ≥2 MiSeq reads were subjected to a BLAST search (Camacho et al., 2009). If the sequence similarity between queries and the top BLAST hit was ≥99% and the E-value was ≤$10^{-5}$, the assembled sequence was assigned to the top-hit species. Conversely, if the top-hit sequence was <99%, the unique sequence was not subjected to the following analyses. Note that ≥99% similarity indicates a less than two-base difference between the query and reference sequences because the maximum sequence length subjected to taxonomic assignment is 195 bp. This procedure also works as a filter for erroneous reads because erroneous reads are expected to never match the reference species DNA at ≥97% similarity by chance. After BLAST searches, assembled sequences assigned to the same species were clustered, and we considered the clustered sequences as an operational taxonomic unit (OTU). The reliability of each assignment was evaluated and classified as high, moderate, or low (Miya et al., 2015). Of those three classes, low-confidence assignments suggest that the taxon assigned to an OTU cannot be distinguished from a second candidate taxon. In the present study, MitoFish, NCBI Organelle Genome Resources were used for the BLAST search. After blast, we confirm the species name according to supplementary of Miya et al. (2015).

## Fish Assemblage Structure

Fish abundance was analyzed according to the total number of reads of each species that was detected in the Yangtze River in the present study.

The Simpson index ($D$) and Shannon–Wiener index ($H'$, $\log_e$) were calculated for each season. The related equations were as follows (Ludwig and Reynolds, 1988):

$$D = (S - 1)/\ln N \quad (1)$$

$$H' = -\sum_{i=1}^{S} P_i \cdot \ln P_i \quad (2)$$

Where $S$ is the number of species, $N$ is the total individuals, and $Pi$ is the proportion of individuals belonging to a species to the total individuals. Calculations were carried out using the diversity function in the vegan package of R software (v3.5.3).

Nonmetric multidimensional scaling (NMDS) was used to show the seasonal difference. NMDS analysis was performed on the weighted UniFrac distance and unweighted UniFrac distance using R software (v3.5.3), and assemblage structure was described by a two-dimensional sorting graph. The thermal image was used to show the difference of assemblage composition based on the abundance distribution of the OTUs or the degree of similarity between the seasons.

Canonical correspondence analysis (CCA) was applied to analyze the correlation between environmental factors and the distribution pattern of ichthyoplankton assemblages.

**TABLE 2 |** OTUs and species information for three seasons.

| OTU ID | Species | Spring | Summer | Autumn |
|---|---|---|---|---|
| AB969962 | *Pennahia argentata* | | | † |
| AB970004 | *Harpadon nehereus* | † | † | † |
| AB972202 | *Upeneus japonicus* | | | † |
| AB972205 | *Larimichthys crocea* | | | † |
| AB972229 | *Trichiurus japonicus* | † | | † |
| AB972233 | *Nibea mitsukurii* | | | † |
| AB974486 | *Benthosema pterotum* | † | | † |
| AB974524 | *Acropoma japonicum* | | | † |
| AB974683 | *Liparis tanakae* | † | | † |
| NC002333 | *Danio rerio* | | | † |
| NC003196 | *Pagrus major* | | | † |
| NC006131 | *Acanthogobius hasta* | † | † | † |
| NC006291 | *Carassius carassius* | | | † |
| NC009579 | *Coilia nasus* | † | † | † |
| NC010194 | *Hypophthalmichthys nobilis* | | | † |
| NC011707 | *Pampus* sp. | † | † | † |
| NC011710 | *Larimichthys polyactis* | † | † | † |
| NC014263 | *Collichthys niveatus* | | † | † |
| NC014350 | *Collichthys lucidus* | | | † |
| NC014351 | *Miichthys miiuy* | † | | † |
| NC015205 | *Nibea albiflora* | † | | |
| NC016693 | *Trypauchen vagina* | | | † |
| NC018347 | *Bahaba taipingensis* | | | † |
| NC020466 | *Cyprinidae* sp. | | | † |
| NC020468 | *Setipinna taty* | † | † | † |
| NC021130 | *Johnius grypotus* | | | † |
| NC021460 | *Psenopsis anomala* | | | † |
| NC022464 | *Johnius belangerii* | † | | † |
| NC023538 | *Coilia grayii* | | | † |
| NC023980 | *Saurida microlepis* | † | | † |
| NC028228 | *Sillago japonica* | | | † |
| NC029228 | *Amblychaeturichthys hexanema* | | | † |
| NC029341 | *Phoxinus semotilus* | | | † |
| NC030374 | *Odontamblyopus lacepedii* | | | † |
| AB969925 | *Decapterus maruadsi* | | † | |
| AB972113 | *Scomber japonicus* | † | † | |
| KM257636 | *Coilia mystus* | † | † | |
| NC009057 | *Oreochromis* sp. | † | † | |
| NC021597 | *Glyptosternon maculatum* | | † | |
| NC024184 | *Trachinotus carolinus* | † | | |
| NC025669 | *Oreochromis niloticus* | † | | |

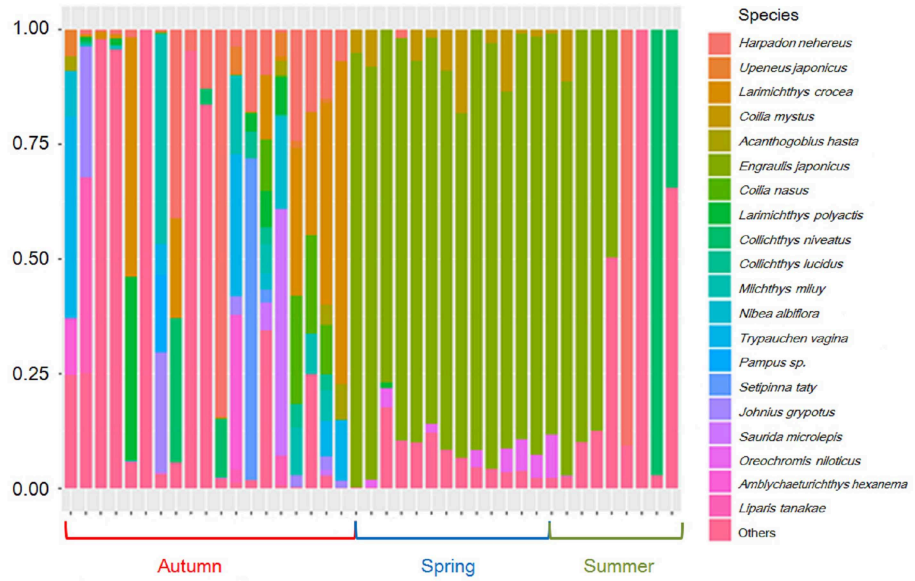


**FIGURE 2 |** Species distribution in three seasons.

To eliminate the effect of few dominant species and plenty of zeros in the species data and highly variable values in the environment data, all data matrices were transformed by log(x + 1). This analysis was performed with CANOCO 5.0.

**TABLE 3 |** Diversity index in different seasons.

| Index | Spring | Summer | Autumn |
|---|---|---|---|
| Simpson index (*D*) | $0.16 \pm 0.02^{A}$ | $0.08 \pm 0.01^{B}$ | $0.56 \pm 0.05^{C}$ |
| Shannon-Wiener index (*H'*) | $0.31 \pm 0.02^{A}$ | $0.12 \pm 0.01^{B}$ | $1.15 \pm 0.11^{C}$ |

*$P < 0.01$, Numbers with different superscript are significantly different with each other.*

# RESULTS

## Species Composition

After the quality-control process, the total number of reads obtained from the run was 2,308,620 for spring, 1,475,304 for summer, and 4,320,556 for autumn. All the reads were blasted in MitoFish, NCBI Organelle Genome Resources, and M.M.'s laboratory (the raw data is available online as **Supplementary Material**). A total of 41 OTUs were identified from the three seasons in the Yangtze Estuary and its adjacent waters in 2018; there were 18, 12, and 33 species separately belonging to spring, summer, and autumn, respectively (**Table 2** and **Figure 2**), 17 species were found in at least two seasons. Due to the strong biases of species composition in different seasons, we selected the first 20 species with read numbers higher than 4,000 for further analysis, because others were detected only in autumn.

The dominant fish species were large yellow croaker *Larimichthys crocea*, Bombay duck *Harpadon nehereus*, Chinese drum *Miichthys miiuy*, burrowing goby *Trypauchen vagina*, Japanese grenadier anchovy *Coilia nasus*, the lizardfish *Saurida microlepis*, croaker *Johnius grypotus*, scaly hairfin anchovy *Setipinna taty*, Osbeck's grenadier anchovy *Coilia mystus*, and pinkgray goby *Amblychaeturichthys hexanema*, with read numbers all higher than 10,000.

## Diversity and Seasonal Assemblage Structure

Indices *D* and *H'* were calculated in each season to show differences among the three seasons (**Table 3** and **Figure 3**). Autumn presented the highest diversity indexes, and summer presented the lowest (**Table 3**). Furthermore, according to the results of multiple comparisons, the diversity indices all differed significantly among spring, summer, and autumn (**Figure 3**).

Seasonal assemblage structure was described in a two-dimensional sorting graph based on the NMDS analysis. These results showed that the assemblage structure differed significantly among the three seasons (**Figure 4**). The thermal image was also used to show the difference of assemblage composition based on the abundance distribution of the OTUs or the degree of similarity among the seasons (**Figure 5**).

## Relationship With Environmental Factors

The relationships between the environmental factors and the detected species were clarified in a CCA ordination diagram using data from the 20 species and the set of eight environmental factors (**Table 4**). A Monte-Carlo test indicated that T was the key environmental factor affecting the fish assemblages ($P < 0.05$); as shown in the plot, the first axis is strongly correlated with T, S, and DO.

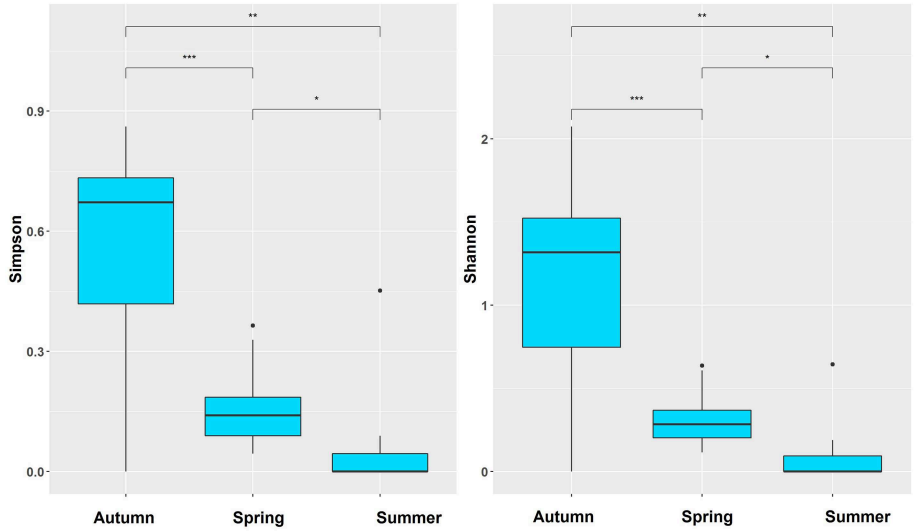


**FIGURE 3 |** The seasonal difference based on Simpson and Shannon-Wiener index; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

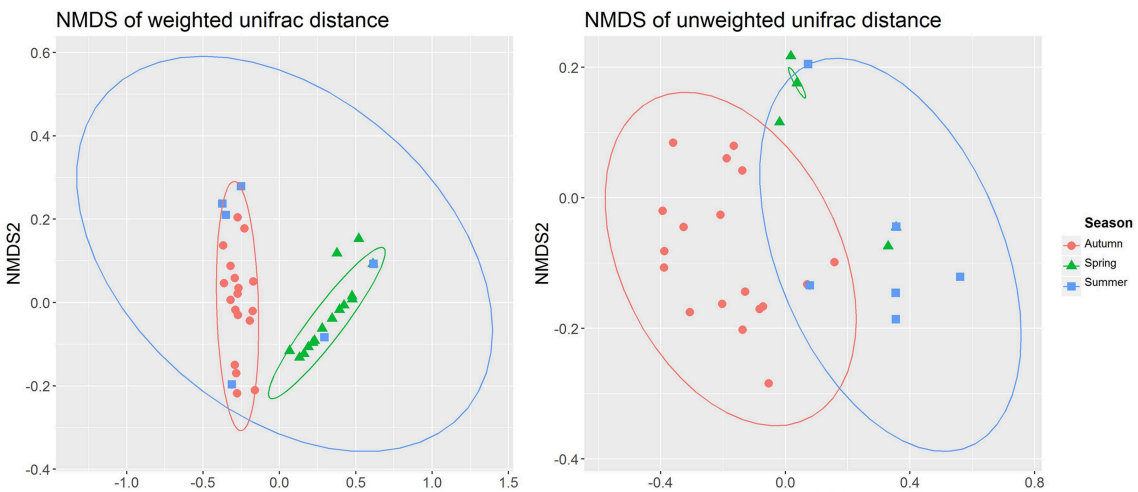


**FIGURE 4 |** Two-dimensional sorting graph of NMDS analysis.

As shown in the CCA ordination plot of fish species (**Figure 6**), the correlation between environmental factors and distributions of the different species was inconsistent, although most of the species showed a strong relationship with T and DO.

## DISCUSSION

The results of the present study demonstrate that eDNA can likewise be an effective tool to estimate fish species diversity, abundances, biomass, and spatial distributions in the Yangtze Estuary and its adjacent waters. As a survey method in aquatic environments, eDNA technology is less disruptive to target species and less destructive to ecosystems, and it can be used to monitor species in different vertebrate categories. For example, eDNA has been effectively used to detect the Japanese giant salamander *Andrias japonicus* (Fukumoto et al., 2015), the loach *Misgurnus fossilis* (Sigsgaard et al., 2015), Alabama sturgeon *Scaphirhynchus suttkusi* (Pfleger et al., 2016), and brook trout *Salvelinus fontinalis* (Evans et al., 2017), among other monitored species.

In recent years, studies have shown that by monitoring changes in eDNA concentrations, eDNA can be used to predict functional trends for spawning grounds, feeding grounds, and nursery habitats in a target area, with less effort and research cost and more-efficient ecological monitoring. Spear et al. (2015) used eDNA to investigate the resource of hellbender salamander *Cryptobranchus alleganiensis* and found that the

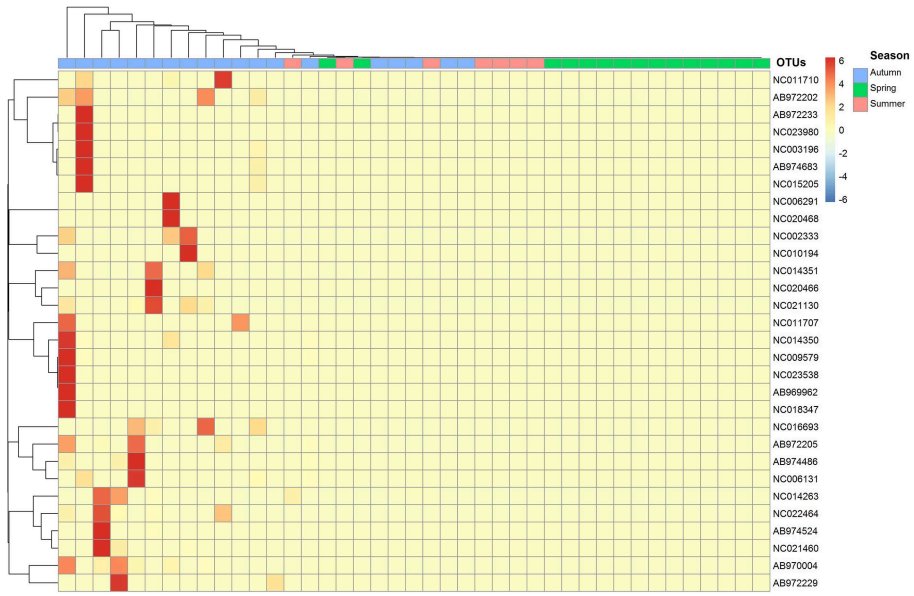


**FIGURE 5 |** Thermal image shows the difference of assemblage composition; OTUs id refers to **Table 2**.

**TABLE 4 |** Conditional effects and correlations of environmental variables by CCA.

| Factor | Explains % | Contribution % | Pseudo-F | *P* |
|---|---|---|---|---|
| T | 8.8 | 27.8 | 3.8 | 0.002 |
| S | 4.3 | 13.6 | 1.9 | 0.03 |
| DO | 3.5 | 11.1 | 1.6 | 0.118 |
| COD | 3.1 | 9.7 | 1.4 | 0.168 |
| pH | 2.5 | 7.9 | 1.1 | 0.362 |
| TN | 5.6 | 17.6 | 2.6 | 0.004 |
| D | 2.2 | 7 | 1.1 | 0.356 |
| TP | 1.7 | 5.4 | 0.8 | 0.598 |

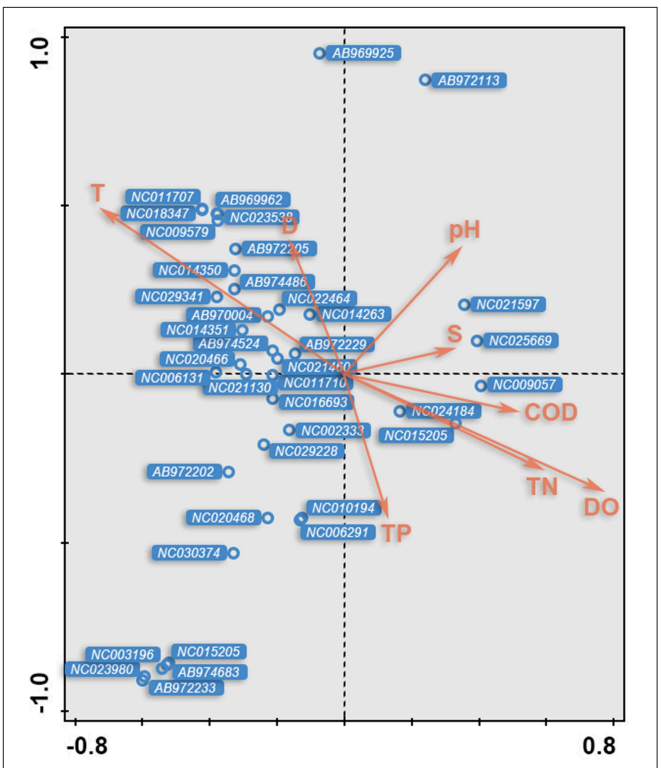


**FIGURE 6 |** The CCA biplot of fish species; OTUs ID were in italic and referred to **Table 2**.

eDNA concentration was at its highest during the animal's concealed breeding period. Erickson et al. (2016) studied the migration path and spawning location of invasive carp using eDNA techniques, and found a correlation between the eDNA concentration and the migration path. Buxton et al. (2017) found that the seasonal concentration of eDNA for northern crested newt *Triturus cristatus* was highest starting in June and exposed an increasing population size for juveniles starting in mid-August. Bylemans et al. (2017) demonstrated that changes in eDNA concentration were an important means of monitoring the oviposition behavior of the endangered freshwater Macquarie perch *Macquaria australasica*, and likewise might be used to discern the spawning periods of other fish species.

Other studies have shown that species abundance and floristic distribution could be determined based on positive correlations between eDNA concentrations and species biomass. Pilliod et al. (2013) used traditional field survey methods and eDNA to sample 13 rivers in Idaho, concluding that the eDNA method was more advantageous and that eDNA concentrations significantly positively correlated with biodensity and biomass. Maruyama et al. (2014) studied the release rate of eDNA into the freshwater environment for various developmental stages of bluegill *Lepomis macrochirus* and found a positive

correlation between eDNA concentration and biomass, based on qPCR. Evans et al. (2016) determined the sequences of six mitochondrial gene fragments for nine freshwater species (eight fishes and one amphibian); the sequence copy number positively correlated with the abundances of the nine species. Doi et al. (2017) undertook a snorkeling survey of the Zopo River in Japan and proved a relationship between the concentration of eDNA and the abundance and biomass of ayu *Plecoglossus altivelis*. Pont et al. (2018) demonstrated of the capacity of eDNA metabarcoding to describe longitudinal fish assemblage patterns in a large river, and metabarcoding appears to be a reliable, cost-effective method for future monitoring. Thus, eDNA has the potential to be used as an indicator of biological abundance. Lacoursière-Roussel et al. (2016) argue that a large number of examples showing a relationship between eDNA concentration and species abundance proves that eDNA might now be feasibly and widely used in fisheries assessments.

In the present study, after quality control, the total number of reads obtained from the run was 2,308,620 for spring, 1,475,304 for summer, and 4,320,556 for autumn, with a total of 41 fish species identified in the 2018 samples across three seasons. This number is a reasonable estimate when compared with traditional surveys. For instance, Zhang (2012) identified 36 species from fish captures in the estuary in 2004 and 2007, 42 species in 2009 and 2010, and 36 species in 2011, though the species composition differed from that found in the present study (**Table 5**). However, the dominant species were different from the present work compared to other studies. The main reason for this should be for the net survey, IRI index is always used for determine dominant species (Zhang, 2012) while we used number of sequence reads in the present study. Other studies also suggest that eDNA analysis is a useful tool to estimate fish abundance/biomass as well as their spatial distribution (such as Doi et al., 2017).

Besides the detection of dominant species, our eDNA analysis also distinguished several rarer fishes, namely large yellow croaker *Larimichthys crocea*, bighead croaker *Collichthys niveatus*, Pacific rudderfish *Psenopsis anomala*, and Gray's grenadier anchovy *Coilia grayii*, which were all regularly captured in the 1980s, before construction of the Three Gorges Dam (Luo and Shen, 1994), but have not been captured in recent years. This finding is particularly useful in relation to conservation of the fisheries resource in the Yangtze Estuary and its adjacent waters.

Strong seasonal changes in the composition of estuarine fish communities are relatively common (Castillo-Rivera et al., 2003), particularly in intermittently open estuaries. Many studies have shown that most estuarine fish assemblages undergo significant seasonal changes in community structure, often related to changes in the estuarine mouth phase and salinity regime (e.g., Vorwerk et al., 2003; Becker and Laurenson, 2008; James et al., 2008; Mendoza et al., 2009). However, some studies have found no seasonal patterns in the fish communities of intermittently closed estuaries, such as in southeastern Australia and along the southeast coast of South Africa; there is no clear seasonality to the mouth-opening events in these systems, which might explain the lack of seasonal patterns in the fish species composition (Griffiths, 2001; Jones and West, 2005; James et al., 2008).

**TABLE 5 |** Species composition in different years.

| Year | Species number | Dominant species | Method |
|---|---|---|---|
| 1998–2001 (Yu and Xian, 2010) | 48 | *Harpodon nehereus, Benthosema ptero tum, Setipinna taty, Pampus argenteus, Johnius belengeri* | Net survey |
| 2004 and 2007 (Zhang, 2012) | 36 | *Trichiurus lepturus*, *Setipinna taty*, *Engraulis japonicus*, *Trachurus japonicus*, *Larimichthys polyactis*, *Scomber japonicus* | Net survey |
| 2009 and 2010 (Zhang, 2012) | 42 | *Larimichthys polyactis*, *Chelidonichthys kumu*, *Harpadon nehereus*, *Setipinna taty*, *Pampus argenteus* | Net survey |
| 2011 (Zhang, 2012) | 36 | *Coilia mystus and Larimichthys polyactis* | Net survey |
| 2018 (This study) | 41 | *Larimichthys crocea*, *Harpadon nehereus*, *Miichthys miiuy*, *Trypauchen vagina*, *Coilia nasus*, *Saurida microlepis*, *Johnius grypotus*, *Setipinna taty*, *Coilia mystus*, *Amblychaeturichthys hexanema* | eDNA |

In the present study, the $D$ and $H'$ indices calculations, NMDS, and the thermal image showed that significant seasonal differences occur in the Yangtze Estuary and its adjacent waters. CCA indicated that T, S, and DO were the main environmental factors affecting the seasonal fish assemblages. In previous work, we attributed significant variations in the fish assemblage from 1999 to 2009 to declines in the number of fish species and biomass, leading to the succession of dominant species. Water temperature, depth, and salinity still have a strong impact on the fish assemblage, while TSM has an important influence on their spatial structure. Freshwater inflow is the critical factor determining abiotic and biotic variability (Morais et al., 2009), and has an important impact on the distribution and abundance of ichthyoplankton within an estuarine ecosystem (Faria et al., 2006). Changes in freshwater flow into an estuary and adjacent areas affect nutrient levels, with consequences for primary productivity and associated trophic chains (Morais et al., 2009). In Yangtze Estuary and its adjacent waters, there are significant differences of freshwater flow in different seasons (Luo and Shen, 1994), thus the environmental factors, especially T, S, and DO change a lot in different season. Variation in environmental factors plays a major role in partitioning fish assemblages in many estuarine systems (e.g., Akin et al., 2003, 2005). The current results suggest that the environmental factors impacting the estuary's fish assemblage have likely changed since the previous study; while it is a common phenomenon that an estuary ecosystem may undergo change, continuous ecological monitoring is needed.

The present primary study is the first case to use eDNA to determine the seasonal fish assemblage structure in the Yangtze Estuary and its adjacent waters, and ongoing research

on this topic is warranted. Continuing investigation will broadly benefit the region's marine resource conservation, sustainable exploitation of fisheries, aquatic products processing, environmental-friendly development, and socioeconomic stability.

## DATA AVAILABILITY

The datasets generated for this study can be found in the **Supplementary Material**.

## AUTHOR CONTRIBUTIONS

HZ wrote the first draft of the manuscript. HZ and WX collected the samples and designed the experiment. SY guided the methods. WI guided the data analysis. All the authors revised the manuscript.

## FUNDING

This work was supported by National Natural Science Foundation of China (Nos. 41976094, U1606404, and 31872568); The International Partnership Program of Chinese Academy of Sciences (CAS-JSPS, GJHZ1885); The Youth Talent Support Program of the Laboratory for Marine Ecology and Environmental Science, Pilot National Laboratory for Marine Science, and Technology (Qingdao) (LMEES-YTSP-2018-01-12). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00515/full#supplementary-material

## REFERENCES

Akin, S., Buhan, E., Winemiller, K. O., and Yilmaz, H. (2005). Fish assemblage structure of Koycegiz Lagoon-Estuary, Turkey: spatial and temporal distribution patterns in relation to environmental variation. *Estuar. Coast. Shelf Sci.* 64, 671–684. doi: 10.1016/j.ecss.2005.03.019

Akin, S., Winemiller, K. O., and Gelwick, F. P. (2003). Seasonal and spatial variations in fish and macrocrustacean assemblage structure in Mad Island Marsh estuary, Texas. *Estuar. Coast. Shelf Sci.* 57, 269–282. doi: 10.1016/S0272-7714(02)00354-2

Andrews, S. (2010). *FastQC: A Quality-Control Tool for High-Throughput Sequence Data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed April 21, 2016).

Barnes, M. A., and Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conserv. Genet.* 17, 1–17. doi: 10.1007/s10592-015-0775-4

Becker, A., and Laurenson, L. J. B. (2008). Presence of fish on the shallow flooded margins of a small intermittently open estuary in southeastern Australia under variable flooding regimes. *Estuar. Coast.* 31, 43–52. doi: 10.1007/s12237-007-9004-9

Bohmann, K., Evans, A., Gilbert, M. T., Carvalho, G. R., Creer, S., Knapp, M., et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol. Evol.* 29, 358–367. doi: 10.1016/j.tree.2014.04.003

Buxton, A. S., Groombridge, J. J., Zakaria, N. B., and Griffiths, R. A. (2017). Seasonal variation in environmental DNA in relation to population size and environmental factors. *Sci. Rep.* 7:46294. doi: 10.1038/srep46294

Bylemans, J., Furlan, E. M., Hardy, C. M., McGuffie, P., Lintermans, M., and Gleeson, D. M. (2017). An environmental DNA (eDNA) based method for monitoring spawning activity: a case study using the endangered Macquarie perch (*Macquaria australasica*). *Methods Ecol. Evol.* 8, 646–655. doi: 10.1111/2041-210X.12709

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Castillo-Rivera, M., Zavala-Hurtado, J. A., and Zárate, R. (2003). Exploration of spatial and temporal patterns of fish diversity and composition in a tropical estuarine system of Mexico. *Rev. Fish Biol. Fisher.* 12, 167–177. doi: 10.1023/A:1025051027676

Davy, C. M., Kidd, A. G., and Wilson, C. C. (2015). Development and validation of environmental DNA (eDNA) markers for detection of freshwater turtles. *PLoS ONE* 10:e0130965. doi: 10.1371/journal.pone.0130965

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., and Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. *J. Appl. Ecol.* 49, 953–959. doi: 10.1111/j.1365-2664.2012.02171.x

Doi, H., Inui, R., Akamatsu,Y., Kanno, K., Yamanaka, H., Takahara, T., et al. (2017). Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshw. Biol.* 62, 30–39. doi: 10.1111/fwb.12846

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Erickson, R. A., Rees, C. B., Coulter, A. A., Merkes, C. M., McCalla, S. G., Touzinsky, K. F., et al. (2016). Detecting the movement and spawning activity of bigheaded carps with environmental DNA. *Mol. Ecol. Resour.* 16, 957–965. doi: 10.1111/1755-0998.12533

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., et al. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Mol. Ecol. Resour.* 16, 29–41. doi: 10.1111/1755-0998.12433

Evans, N. T., Shirey, P. D., Wieringa, J. G., Mahon, A. R., and Lamberti, G. A. (2017). Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. *Fisheries* 42, 90–99. doi: 10.1080/03632415.2017.1276329

Faria, A., Morais, P., and Chicharo, M. (2006). Ichthyoplankton dynamics in the Guadiana estuary and adjacent coastal area, South-East Portugal. *Estuar. Coast. Shelf Sci.* 70, 85–97. doi: 10.1016/j.ecss.2006.05.032

Ficetola, G. F., Miaud, C., Pompanon, F., and Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biol. Lett.* 4, 423–425. doi: 10.1098/rsbl.2008.0118

Fukumoto, S., Ushimaru, A., and Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. *J. Appl. Ecol.* 52, 358–365. doi: 10.1111/1365-2664.12392

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., and Waits, L. P. (2011). Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. *PLoS ONE* 6:e22746. doi: 10.1371/journal.pone.0022746

Goldberg, C. S., Strickler, K. M., and Pilliod, D. S. (2015). Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. *Biol. Conserv.* 183, 1–3. doi: 10.1016/j.biocon.2014.11.040

Griffiths, S. P. (2001). Factors influencing fish composition in an Australian intermittently open estuary. *Is stability salinity dependent. Estuar. Coast. Shelf Sci.* 52, 739–751. doi: 10.1006/ecss.2000.0756

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P., et al. (2013). MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Mol. Biol. Evol.* 30, 2531–2540. doi: 10.1093/molbev/mst141

James, N. C., Whitfield, A. K., and Cowley, P. D. (2008). Long-term stability of the fish assemblages in a warm-temperate South African estuary. *Estuar. Coast. Shelf Sci.* 76, 723–738. doi: 10.1016/j.ecss.2007.07.036

Jerde, C. L., Mahon, A. R., Chadderton, W. L., and Lodge, D. M. (2011). Sight-unseen detection of rare aquatic species using environmental DNA. *Conserv. Lett.* 4, 150–157. doi: 10.1111/j.1755-263X.2010.00158.x

Jones, M. V., and West, R. J. (2005). Spatial and temporal variability of seagrass fishes in intermittently closed and open coastal lakes in southeastern Australia. *Estuar. Coast. Shelf Sci.* 64, 277–288. doi: 10.1016/j.ecss.2005.02.021

Lacoursière-Roussel, A., Côté, G., Leclerc, V., and Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *J. Appl. Ecol.* 53, 1148–1157. doi: 10.1111/1365-2664.12598

Ludwig, J. A., and Reynolds, J. F. (1988). *Statistical Ecology: A Primer in Methods and Computing*. Hoboken, NJ: John Wiley and Sons.

Luo, B., and Shen, H. (1994). *Impacts of the Three Gorges Project on Estuarine Ecology and Environment*. Beijing: Science Press.

Magnuson, J. J., Benson, B. J., and McLain, A. S. (1994). Insights on species richness and turnover from long-term ecological research: fishes in north temperate lakes. *Am. Zool.* 34, 437–451. doi: 10.1093/icb/34.3.437

Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., and Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLoS ONE* 9:e114639. doi: 10.1371/journal.pone.0114639

Mendoza, E., Castillo-Rivera, M., Zárate-Hernández, R., and Ortiz Burgos, S. (2009). Seasonal variations in the diversity, abundance, and composition of species in an estuarine fish community in the Tropical Eastern Pacific, Mexico. *Ichthyol. Res.* 56, 330–339. doi: 10.1007/s10228-009-0102-5

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., et al. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Roy. Soc. Open Sci.* 2:150088. doi: 10.1098/rsos.150088

Morais, P., Chícharo, M. A., and Chícharo, L. (2009). Changes in a temperate estuary during the filling of the biggest European dam. *Sci. Total. Environ.* 407:2245–2259. doi: 10.1016/j.scitotenv.2008.11.037

Pfleger, M. O., Rider, S. J., Johnston, C. E., and Janosik, A. M. (2016). Saving the doomed: using eDNA to aid in detection of rare sturgeon for conservation (*Acipenseridae*). *Glob. Ecol. Conserv.* 8, 99–107. doi: 10.1016/j.gecco.2016.08.008

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Can. J. Fish. Aquat. Sci.* 70, 1123–1130. doi: 10.1139/cjfas-2013-0047

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Sci. Rep.* 8:10361. doi: 10.1038/s41598-018-28424-8

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., and Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA: a review of eDNA as a survey tool in ecology. *J. Appl. Ecol.* 51, 1450–1459. doi: 10.1111/1365-2664.12306

Sato, Y., Miya, M., Fukunaga,T., Sado, T., and Iwasaki, W. (2018). MitoFish and MiFish pipeline: a mitochondrial genome database of fish with an analysis pipeline for environmental DNA metabarcoding. *Mol. Biol. Evol.* 35, 1553–1555. doi: 10.1093/molbev/msy074

Sigsgaard, E. E., Carl, H., Møller, P. R., and Thomsen, P. F. (2015). Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biol. Conserv.* 183, 46–52. doi: 10.1016/j.biocon.2014.11.023

Smart, A. S., Weeks, A. R., van Rooyen, A. R., Moore, A., McCarthy, M. A., and Tingley, R. (2016). Assessing the cost-efficiency of environmental DNA sampling. *Methods Ecol. Evol.* 7, 1291–1298. doi: 10.1111/2041-210X.12598

Spear, S. F., Groves, J. D., Williams, L. A., and Waits, L. P. (2015). Using environmental DNA methods to improve detectability in a hellbender (*Cryptobranchus alleganiensis*) monitoring program. *Biol. Conserv.* 183, 38–45. doi: 10.1016/j.biocon.2014.11.016

Taberlet, P., Coissac, E., Hajibabaei, M., and Rieseberg, L. H. (2012). Environmental DNA. *Mol. Ecol.* 21, 1789–1793. doi: 10.1111/j.1365-294X.2012.05542.x

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., and Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE* 7:e35868. doi: 10.1371/journal.pone.0035868

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., et al. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Mol. Ecol.* 21, 2565–2573. doi: 10.1111/j.1365-294X.2011.05418.x

Vorwerk, P. D., Whitfield, A. K., Cowley, P. D., and Paterson, A. W. (2003). The influence of selected environmental variables on fish assemblage structure in a range of southeast African estuaries. *Environ. Biol. Fish.* 66, 237–247. doi: 10.1023/A:1023922521835

Yang, D. L., Wu, G. Z., and Sun, J. R. (1990). The investigation of pelagic eggs, larvae and juveniles of fishes at the mouth of the Changjiang River and adjacent areas. *Oceanol. Limnol. Sin.* 4, 346–355.

Yu, H. C., and Xian, W. W. (2010). Fish assemblage structure in adjacent Sea of Yangtze Estuary and its relationship with environmental factors. *J. Yangtze River Sci. Res. Inst.* 27, 88–92.

Zhang, H., Xian, W., and Liu, S. (2015). Ichthyoplankton assemblage structure of springs in the Yangtze Estuary revealed by biological and environmental visions. *PeerJ* 3:e1186. doi: 10.7717/peerj.1186

Zhang, H., Xian, W., and Liu, S. (2016). Autumn ichthyoplankton assemblage in the Yangtze Estuary shaped by environmental factors. *PeerJ* 4:e1922. doi: 10.7717/peerj.1922

Zhang, H., Xian, W. W., and Liu, S. D. (2019). Seasonal variations of the ichthyoplankton assemblage in the Yangtze Estuary and its relationship with environmental factors. *PeerJ* 7:e6482. doi: 10.7717/peerj.6482

Zhang, Y. Q. (2012). *Environmental impact on the fish assemblage structure in adjacent sea area of the Yangtze River Estuary* (Master's thesis). Graduate University of Chinese Academy of Sciences, Beijing, China.

Zhong, J., Wu, M., and Lian, Q. (2007). Composition of fish larvae and juveniles in spring and summer in surf zone of Yangtze Estuary. *J. Fish. Sci. China.* 3, 436–443. doi: 10.3321/j.issn:1005-8737.2007.03.014

Zhu, X., Liu, D., and Sha, X. (2002). Relationship between community structure of ichthyoplankton and environmental factors in the estuary of Changjiang River in spring. *Stud. Mar. Sin.* 44, 169–179.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 04 September 2019
doi: 10.3389/fmars.2019.00545

# Phylogeography and Historical Demography of Two Sympatric Atlantic Snappers: *Lutjanus analis* and *L. jocu*

*Allyson Santos de Souza[1]*, Eurico Azevedo Dias Júnior[1], Manolo Fernandez Perez[2], Marcelo de Bello Cioffi[2], Luiz Antonio Carlos Bertollo[2], Erik Garcia-Machado[3], Marcelo Nazareno Souza Vallinoto[4], Pedro Manoel Galetti Jr.[2] and Wagner Franco Molina[1]*

[1] Departamento de Biologia Celular e Genética, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [2] Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [3] Centro de Investigaciones Marinas, Universidad de la Habana, Havana, Cuba, [4] Departamento de Genética e Biologia Molecular, Universidade Federal do Pará, Belém, Brazil

*Lutjanus analis* (mutton snapper) and *Lutjanus jocu* (dog snapper) are mesopredator species with extensive geographic distribution in the Atlantic Ocean. Although historically overfished, their genetic diversity, population structuring, and historical demography along the Brazilian coast are unknown. Here, we present genetic data for the hypervariable region 1 (HVR1) of the mtDNA control region of both *L. jocu* and *L. analis,* and for *cytB* of *L. analis* from distinct geographic regions. Phylogeographic analyses based on HVR1 sequences revealed no geographic structuring of mtDNA lineages for either species. The bimodal mismatch distribution plots of mutton and dog snapper populations implied that they might have experienced secondary contact. Historical demographic parameters estimated that population expansions ranged from 13,631 to 61,171 years before present (ybp) for *L. analis* and 36,783–55,577 ybp for *L. jocu*, associated with events that occurred at the end of the last glaciation period. Estimates of the average effective population size for *L. jocu* were higher than those for *L. analis*, with the largest population occupying the Brazilian northeastern region coast. High migration rates are maintained among the three northernmost locations, with a substantial decrease to the further southern region. Our study suggests that a tripartite interaction of larval dispersion and interregional adult movement (aggregate spawning), allied to historical contingencies, contributed to contemporary population genetic patterns of these species, and adds relevant information for conservation management of such vulnerable – and valuable – marine resources.

**Keywords: Lutjanidae, Brazilian coast, genetic diversity, HVR1 control region, cytochrome B, Pleistocene**

OPEN ACCESS

***Edited by:***
*Dongdong Xu,*
*Marine Fisheries Research Institute of Zhejiang, China*

***Reviewed by:***
*Jin Sun,*
*Hong Kong University of Science and Technology, Hong Kong*
*Bingjian Liu,*
*Zhejiang Ocean University, China*

****Correspondence:***
*Allyson Santos de Souza*
*souzaas@yahoo.com.br*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *21 June 2019*
***Accepted:*** *19 August 2019*
***Published:*** *04 September 2019*

***Citation:***
*Souza AS, Dias Júnior EA, Perez MF, Cioffi MB, Bertollo LAC, Garcia-Machado E, Vallinoto MNS, Galetti PM Jr and Molina WF (2019) Phylogeography and Historical Demography of Two Sympatric Atlantic Snappers: Lutjanus analis and L. jocu. Front. Mar. Sci. 6:545. doi: 10.3389/fmars.2019.00545*

## INTRODUCTION

Genetic data are critical for developing strategies to delimit populations and conserve marine species. Identifying geographic borders or diagnostic characteristics for distinct genetic subunits play decisive roles in defining threatened species and proposing conservation units. This information has a direct impact on conservation and management policies (Morin et al., 2010). Nevertheless, spatial distribution patterns of genetic diversity remain scarce for most commercially

exploited species (Reiss et al., 2009). These include important marine resources, such as fishes of the family Lutjanidae, which are captured over the entire area of its geographic distribution (Allen, 1985; Burton, 2002; Graham et al., 2008).

Lutjanidae species, known as snappers, occur in the Eastern and Indo-Western Pacific, Eastern, and Western Atlantic. Many species have an extensive geographical distribution, although a few species have restricted ones (Allen, 1985; Lindeman et al., 2001; Moura and Lindeman, 2007). They are demersal fishes of medium to large size, with some species reaching more than one meter, occupying very shallow environments up to 500 m deep (Allen, 1985). Lutjanid fishes constitute one of the most valuable fisheries categories, which contributes to their considerable exploitation. Additionally, they also represent an important food resource for communities dependent on artisanal fishing (Rezende et al., 2003).

Twelve species of the genus *Lutjanus* have been intensely exploited along the Brazilian coast (Rezende et al., 2003; Klippel et al., 2005; Frédou et al., 2009). Among these fishing targets are *Lutjanus analis* (Cuvier, 1828) and *Lutjanus jocu* (Bloch and Schneider, 1801), mainly found in coral reefs, and on the sandy bottoms of bays and estuaries, from Massachusetts (United States) to southeastern Brazil, including the Caribbean and Gulf of Mexico (Allen, 1985).

Although several biological aspects of the *Lutjanus* species are well known (e.g., Rezende et al., 2003; Burton et al., 2005; Mattos and Maynou, 2009; Teixeira et al., 2010; Preveiro et al., 2011), including initiatives toward fish farming (Watanabe et al., 1998), population genetic studies are generally scarce (Garber et al., 2004; Zhang et al., 2006; Karlsson et al., 2009). Among the latter, some data are available for certain North Atlantic species (Heist and Gold, 2000; Gold et al., 2011) and, to a lesser extent, for Western Atlantic species (Salles et al., 2006; Gomes et al., 2008).

Accordingly, the Brazilian *Lutjanus*' genetic patterns are still unknown, hindering appropriate conservation strategies for populations of this genus. This is a particularly necessary and urgent task, since it is estimated that the capture level of the species exceeds nearly 90% of that considered adequate (Frédou et al., 2009). In fact, *L. analis* is being overfished on the northeast coast of Brazil and is considered a near-threatened species (Frédou et al., 2009; IUCN, 2019).

In order to ascertain the levels of connectivity between populations of *L. analis* and *L. jocu* along the Brazilian coast, sequences of the hypervariable region 1 (HVR1) of the mtDNA control region of both species, and also of the cytochrome b (*cytB*) gene of *L. analis* were analyzed, comprising wide areas from northeastern to southeastern Brazilian coast. These data proved to be useful for assessing the stock and evolutionary dynamics of these particularly vulnerable species.

# MATERIALS AND METHODS

## Biological Material

*Lutjanus analis* and *L. jocu* individuals were collected through artisanal fishing in four areas of the Brazilian coast; three in the NE region (Ceará – CE, Rio Grande do Norte – RN and Bahia – BA) and one in the SE region (Espírito Santo – ES), covering a representative zone of the species' distribution in the South Western Atlantic. Eighty one *L. analis* individuals – the mutton snapper – and one hundred *L. jocu* – the dog snapper – were used in phylogeographic analyses (**Table 1**). Fragments of muscle or liver tissues were placed in 2.0 ml microtubes containing 95% ethyl alcohol and stored at a temperature of −20°C.

**TABLE 1 |** Collection sites of *Lutjanus jocu* and *Lutjanus analis* individuals along the Brazilian coast.

| Species | N | Locality | Geographic coordinates |
|---|---|---|---|
| *Lutjanus analis* | 27 | Fortaleza, Ceará (CE) state | 03°37'22.75"S, 38°19'19.13"W |
| | 29 | Natal, Rio Grande do Norte (RN) state | 5°44'5.29"S, 35°4'3.94"W |
| | 17 | Salvador, Bahia (BA) state | 12°57'54.19"S, 38°17'46.77"W |
| | 8 | Vila Velha, Espírito Santo (ES) state | 20°22'53.10"S, 40°15'32.32"W |
| *Lutjanus jocu* | 31 | Fortaleza, Ceará (CE) state | 03°37'22.75"S, 38°19'19.13"W |
| | 39 | Natal, Rio Grande do Norte (RN) state | 5°44'5.29"S, 35°4'3.94"W |
| | 14 | Salvador, Bahia (BA) state | 12°57'54.19"S, 38°17'46.77"W |
| | 16 | Vila Velha, Espírito Santo (ES) state | 20°22'53.10"S, 40°15'32.32"W |

## DNA Extraction and Amplification

DNA was extracted according to Sambrook et al. (1989). PCR primers L1 (5' CCT AAC TCC CAA AGC TAG GTA TTC 3') and H2 (5' CCG GCA GCT CTT AGC TTT AAC TA 3') (Gomes et al., 2008) were used to amplify hypervariable region 1 (HVR1) and *cytB-F* (5' – ACC ACC GTT GTT ATT CAA CTA CAA GAA C– 3') and *cytBI-5R* (5' – GGT CTT TGT AGG AGA AGT ATG GGT GGA A– 3') (Jiménez et al., 2007) were used to amplify the cytochrome b gene. PCR amplifications were composed of 4 μl of dNTP mix (1.25 mM), 2.5 μl of buffer (10×), 1 μl of $MgCl_2$ (50 mM), 1.0 μl of each primer (10 pmol/μl), 1–2 μl of total DNA, 0.2 μl of *Taq* DNA Polymerase (5 U/μl) (Life Technologies Corporation, Carlsbad, United States) and ultra-pure water, for a final volume of 25 μl. The reactions were performed under the following conditions: initial denaturation at 94°C for 2 min; 30 denaturation cycles at 94°C for 30 s; hybridization at 57°C for control region and 52şC for cytB for 1 min; extension at 72°C for 2 min; final extension at 72°C for 5 min. PCR products were purified after ExoSAP-IT® digestion (USB Corporation, Cleveland, United States), according to manufacturer's recommendations.

Samples were sequenced on an Applied Biosystems® 3500xL Genetic Analyzer (Thermo Fisher Scientific®, United Kingdom).

# DATA ANALYSIS

Nucleotide sequences for the *cytB* gene and HVR1 region (GenBank access numbers JQ727916-JQ72996 – *Lutjanus*

*analis*) of mtDNA were checked, manually corrected, and aligned by Muscle (Edgar, 2004), implemented in MEGA6 (Tamura et al., 2013).

Dambe software (Xia and Xie, 2001) was used to determine the possible existence of saturation involving transitions and transversions. Haplotype (h) and nucleotide ($\pi$) diversities (Nei, 1987), as well as genetic differentiation between localities, using pairwise F-statistics were estimated using Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). $F_{st}$ values were tested for significance with 10,000 permutations. Analyses of molecular variance (AMOVA) (Excoffier et al., 1992) were also implemented to determine possible structuring between the different geographic regions analyzed. In addition, Spatial Analysis of Molecular Variance (SAMOVA) was used as an approach to define groups of populations that were geographically homogeneous and also maximally differentiated from each other (Dupanloup et al., 2002).

The distribution of haplotype diversity in each population was assessed with a haplotype network for each DNA marker built in TCS (Clement et al., 2000). We used TCSbu (Santos et al., 2015) for graphical representation of the obtained networks.

Mismatch analyses (frequency of pairwise nucleotide site differences between sequences) were carried out to estimate the historical demographic parameters $\tau$, $\theta_0$ and $\theta_1$, according to the sudden population expansion model (Rogers and Harpending, 1992). Adjustments to an unimodal distribution of pairwise nucleotide differences among sequences, as expected for expanding populations, was tested using Harpending's raggedness index ($r$). Non-significant $r$ indices suggest a good fit to the growth-decline model, while significant indexes are indicative of a stable population (Harpending, 1994). The sum of square deviations (SSD) between the observed and expected distributions was also considered (Schneider and Excoffier, 1999). Additionally, Fu's $F_S$ (Fu, 1997) and Tajima's $D$ neutrality tests (Tajima, 1989), with 10,000 permutations, were used to infer historical demographic variations.

Expansion times expressed in units of mutational time $\tau$ were translated to time since expansion as $\tau = 2ut$, where 2u is the mutational rate per site per million years. Similarly, $\theta$ estimates ($\theta_0 = 2N_0\ \mu$ and $\theta_1 = 2N_1\mu$ were used to estimate the female effective population sizes before ($\theta_0$) and after ($\theta_1$) expansion, where $\mu$ is the mutation rate (Rogers and Harpending, 1992; Rogers, 1995). Confidence intervals of parameters were calculated by applying a parametric bootstrap approach (Schneider and Excoffier, 1999). Values of $\tau$ were transformed in real-time by using the estimated expansion time, through the equation $u = \mu mt$, where u is the mutational rate by generation, $\mu$ (mi) the estimated mutation rate for the sequence analyzed, and mt the number of nucleotide bases involved. The expansion time interval was calculated as $\tau = 2ut$, where $t$ is the estimated time from the occurrence of expansion. To that end, a generation time of 4 years was used for *L. analis* (Burton, 2002) and 4.6 years for *L. jocu*, established as means of different estimates ranging from 2 to 6 years (Ault et al., 1998, Froese and Binohlan, 2000).

Estimates of contemporary migration rates among populations for each species were obtained using Migrate-n 4.2.14 (Beerli, 2006; Beerli and Palczewski, 2010). Based on preliminary runs, all priors were set as default, except for theta in *L. analis*, which was draw from a uniform distribution with mean 0.01, minimum 0.00 and maximum 0.10; and migration rate in *L. jocu*, that was set with mean 500.0, minimum 0.0 and maximum 5000.0. The analyses were conducted with two independent runs of three long chains over 30,000,000 iterations each, after discarding the first 50,000 steps as burn-in. Posterior samples were recorded every 500 steps. We used a static heating scheme with temperatures 1.0; 1.5; 3.0 and 10,000.0. Convergence for each parameter was evaluated with effective sample size (ESS) values.

# RESULTS

## Molecular Diversity – *cytB* and HVR1

Sequencing of the *cytB* gene from 73 *L. analis* individuals resulted in the production of a 667 bp sequence. We identified 19 distinct haplotypes differentiated by 20 polymorphic sites, all involving DNA sequence transitions (**Table 2**). The sequences showed no evidence of saturation between transitions and transversions (ts/tv). The haplotype and nucleotide diversity varied between 0.88 and 0.61, and 0.001 and 0.005, respectively (**Table 2**).

PCR amplification of the HVR1 of *L. analis* and *L. jocu* produced 493 bp and 394 bp, respectively, analyzable sequences. For *L. analis,* a total of 245 polymorphic sites and 262 nucleotide substitutions (39 transversions and 223 transitions) were detected among the 72 haplotypes identified. Samples of *L. jocu* exhibited 91 haplotypes, with 272 polymorphic sites and 278 nucleotide substitutions (27 transversions and 251 transitions).

This marker in both species also showed no evidence of saturation involving $t_s/t_v$. Only five *L. analis* and nine *L. jocu* haplotypes were shared within and between the different sites. The amount of nucleotide bases followed the A > T > C > G pattern (Sbisà et al., 1997). Geographic samples showed high genetic variability in terms of haplotype ($h$) and nucleotide ($\pi$) diversity, whose values varied from 0.985 to 1.0, and 0.029 to 0.033, respectively, in *L. analis,* and from 0.991 to 1.0, and 0.035 to 0.038, respectively, in *L. jocu* (**Table 2**). The *L. analis* populations of CE, BA, and RN were more polymorphic than the ES population, while nucleotide diversity was similar in all *L. jocu* populations.

The reduced number of haplotypes shared within and between populations reflects the high $h$ and $\pi$ values observed, which is expected for populations without genetic depression.

## Population Genetic Structure

$F_{st}$ indexes based on HVR1 sequences indicated panmixia between the distribution areas for both *L. analis* and *L. jocu* (**Table 3**). However, the *cytB* sequences suggested a possible population structure of *L. analis* in relation to RN and ES ($F_{st} = 0.19$, $p = 0.005$). Given the absence of population structure suggested by $F_{st}$ estimates concerning HVR1 sequences, AMOVA of samples from different localities were assessed as a single population. AMOVA analyses in both species indicated that genetic variation is majorly found within populations, with no

**TABLE 2 |** Molecular and genetic diversity characteristics among mtDNA sequences of *Lutjanus analis* and *Lutjanus jocu* by locality.

| Site/species | Code | N | H | Lp | $t_s$ | $t_v$ | *h* | π |
|---|---|---|---|---|---|---|---|---|
| *Lutjanus analis* | | | | | | | | |
| **HVR-1** | | | | | | | | |
| Ceará | CE | 27 | 23 | 64 | 61 | 9 | 0.985 ± 0.012 | 0.032 ± 0.015 |
| Rio Grande do Norte | RN | 29 | 29 | 76 | 69 | 11 | 1.000 ± 0.009 | 0.032 ± 0.016 |
| Bahia | BA | 17 | 17 | 63 | 55 | 13 | 1.000 ± 0.020 | 0.033 ± 0.016 |
| Espírito Santo | ES | 8 | 8 | 42 | 38 | 6 | 1.000 ± 0.062 | 0.029 ± 0.016 |
| **cytB** | | | | | | | | |
| Ceará | CE | 23 | 11 | 11 | 11 | 0 | 0.806 ± 0.079 | 0.003 ± 0.002 |
| Rio Grande do Norte | RN | 20 | 10 | 12 | 12 | 0 | 0.889 ± 0.050 | 0.005 ± 0.003 |
| Bahia | BA | 14 | 8 | 11 | 11 | 0 | 0.824 ± 0.097 | 0.004 ± 0.002 |
| Espírito Santo | ES | 16 | 6 | 7 | 7 | 0 | 0.616 ± 0.134 | 0.001 ± 0.001 |
| *Lutjanus jocu* | | | | | | | | |
| **HVR-1** | | | | | | | | |
| Ceará | CE | 31 | 30 | 78 | 74 | 4 | 0.997 ± 0.008 | 0.038 ± 0.019 |
| Rio Grande do Norte | RN | 39 | 37 | 84 | 78 | 9 | 0.997 ± 0.006 | 0.035 ± 0.017 |
| Bahia | BA | 14 | 14 | 55 | 49 | 7 | 1.000 ± 0.027 | 0.036 ± 0.019 |
| Espírito Santo | ES | 16 | 15 | 55 | 50 | 7 | 0.991 ± 0.025 | 0.036 ± 0.019 |

*N – number of individuals, H – number of haplotypes, Lp – number of polymorphic loci, $t_s$ – number of transitions, $t_v$ – number of transversions, h – haplotype diversity and π – nucleotide diversity.*

variance among groups (**Table 4**). The absence of valid clusters was also supported by SAMOVA.

## Haplotype Networks

Haplotype networks for HVR1 in both species were similar with several singletons present in all populations, thus reflecting the high variability of this sequence and the absence of population differentiation. *L. analis cytB* network showed a lower number of haplotypes, with a high-frequency central haplotype present in all populations, and with other haplotypes with intermediate frequencies also appearing in several populations and in a few singletons located at the tips of the network (**Figure 1**).

## Migration Rates

Estimated θ values (an estimator of effective population sizes, scaled by mutation) for *L. analis*, showed higher values in CE and BA populations, while RN and ES showed much lower estimates. In *L. jocu*, higher values were found in CE, RN and BA and a smaller sizes in ES (**Figure 2**). Migration rates among *L. analis* populations presented fewer immigrating individuals per generation in ES, and higher immigration in BA and CE. *L. jocu* populations presented similar gene flow patterns, with lower immigration rates observed in ES, while a high number of migrants per generation was detected from BA to CE, from RN to BA, and between CE and RN (**Figure 2**).

## Demographic History

Analysis of deviations from neutrality based on Tajima D and Fu's FS, presented negative values for all populations of *L. analis* and *L. jocu*. Indeed, Fu's $F_S$ (Fu, 1997), a sensitive index to identify sudden population expansions (Ramos-Onsins and Rozas, 2002), exhibited negative and significant values ($P < 0.05$), indicating a demographic expansion for all sampled populations, except for *L. analis* from ES (**Table 4**). Raggedness (*r*) indices and the sum of square deviations (SSD) performed on *L. analis* data were not significant ($P > 0.05$), also supporting population expansion for

**TABLE 3 |** Pairwise $F_{ST}$ values among populations of *Lutjanus analis* and *Lutjanus jocu* based on HVR1 sequence dataset (below the diagonal), and *p* values (above the diagonal).

| Species/site | CE | RN | BA | ES |
|---|---|---|---|---|
| *L. analis* | | | | |
| CE | – | 0.722 | 0.762 | 0.167 |
| RN | −0.013 | – | 0.908 | 0.218 |
| BA | −0.019 | −0.023 | – | 0.116 |
| ES | 0.034 | 0.014 | 0.053 | – |
| *L. jocu* | | | | |
| CE | – | 0.852 | 0.916 | 0.763 |
| RN | −0.010 | – | 0.863 | 0.652 |
| BA | −0.022 | −0.017 | – | 0.553 |
| ES | −0.013 | −0.009 | −0.009 | – |

*CE, Ceará state; RN, Rio Grande do Norte state; BA, Bahia state; ES, Espírito Santo state; *p < 0.05.*

**TABLE 4 |** Analysis of molecular variance (AMOVA) of *Lutjanus analis* and *Lutjanus jocu,* based on $F_{st}$ indices among the different sample sites.

| Variation source | df | SD | Variance | Variation % |
|---|---|---|---|---|
| *Lutjanus analis – cytB* | | | | |
| Among populations | 3 | 7.072 | 0.059 Va | 4.43* |
| Within populations | 69 | 88.600 | 1.284 Vb | 95.57 |
| *Lutjanus analis – HVR1* | | | | |
| Among populations | 3 | 20.718 | −0.037 Va | −0.50 |
| Within populations | 77 | 587.208 | 7.620 Vb | 100.50 |
| *Lutjanus jocu – HVR1* | | | | |
| Among populations | 3 | 14.478 | −0.096 Va | −1.38 |
| Within populations | 96 | 680.842 | 7.092 Vb | 101.38 |

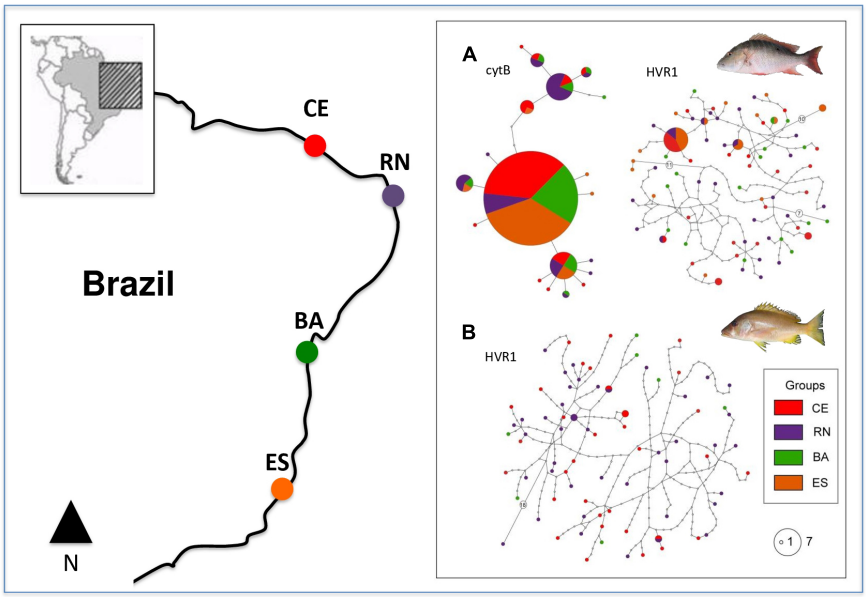


**FIGURE 1 |** Map of collection points and haplotype networks of *Lutjanus analis* **(A)** and *Lutjanus jocu* **(B)**. Each haplotype is represented by a circle, color coded according to the population of origin. Circle sizes are related to their respective frequency. Unsampled haplotypes are represented as empty circles, and multi-step differences between haplotypes are represented with an empty circle containing the number of mutation steps.

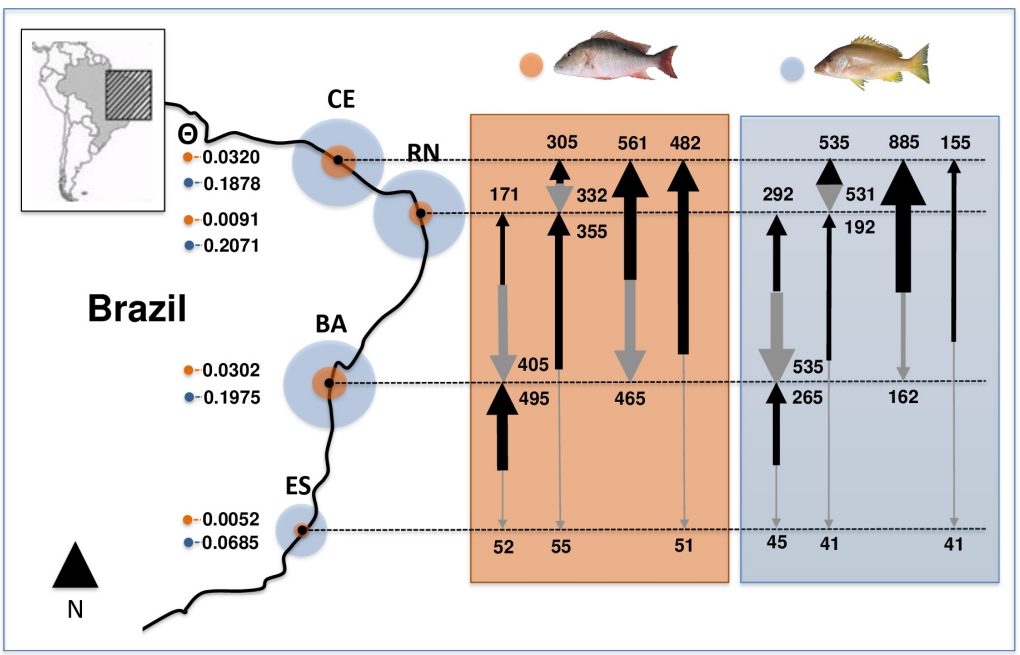


**FIGURE 2 |** Effective population sizes (median estimates of θ; proportional to circles size) and migration rates among populations of *Lutjanus analis* (orange circle) and *Lutjanus jocu* (blue circle). Arrows represent migration rates, with arrow boldness being relative to the estimate of M.

this species. Mismatch distribution graphs (Rogers, 1995) showed a typical bimodal pattern in both species. Associated with other indicators, the bimodal mismatch distribution can be indicative of secondary contact during the period of population expansion of the species (**Figure 3**).

Coalescence data based on *t* indicated a period of population expansion between 13,795 and 61,100 ybp, for *L. analis* and between 36,783 and 55,577 ybp for *L. jocu* (**Table 5**). This parameter was not calculated for ES due to no indication of expansion.

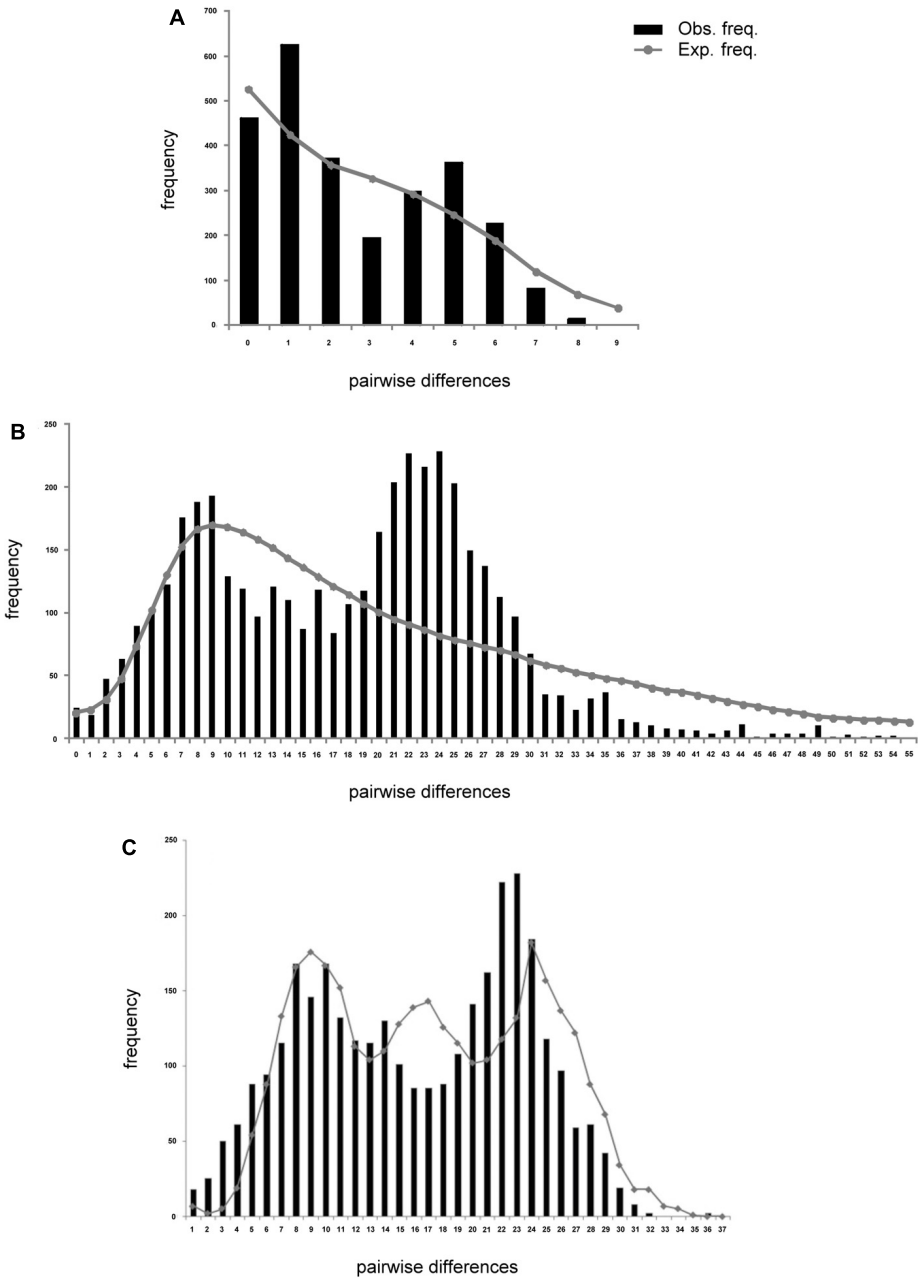


**FIGURE 3 |** Mismatch distribution plots of *Lutjanus analis* **(A,B)** and *Lutjanus jocu* **(C)**. mtDNA *cytB* **(A)**, and HVR1 control region sequences **(B,C)**. The observed distribution is represented by vertical bars. The expected distribution, as predicted by the population expansion model, is represented by a solid line.

# DISCUSSION

## Population Structure of the Mutton and Dog Snappers in Brazilian Coastal Areas

Genetic congeneric comparisons provide an adequate phylogenetic model for estimating contemporary and historical events on population dynamics and genetic population structure. On the Brazilian coast, the species *L. analis* and *L. jocu*, present sympatric distribution, and share several biological and ecological characteristics (Allen, 1985), offering an excellent model for comparisons of spatial genetic patterns from historical and phylogenetic perspectives.

Mutton and dog snappers showed high haplotype (h) and nucleotide diversity ($\pi$) for mtDNA sequences, a pattern similarly presented by other lutjanid species (Ovenden and Street, 2003; Garber et al., 2004; Blaber et al., 2005; Zhang et al., 2006; Gomes et al., 2008), and compatible with the occurrence of secondary contacts among previously differentiated strains (Grant and Bowen, 1998).

**TABLE 5 |** Genetic demographic history for populations of *Lutjanus analis* and *Lutjanus jocu*, based on mtDNA HVR1 control region sequences variation.

| Sites | *Demography* | | | | | | |
|---|---|---|---|---|---|---|---|
| | D | Fs | W | τ | Rag | SSD | Expansion age (ty) |
| *L. analis* | | | | | | | |
| CE | −0.05 | −8.32*** | 16.34 | 5.06 | 0.012 | 0.017 | 13,631–15,385 |
| RN | −0.58 | −16.78*** | 18.84 | 9.04 | 0.005 | 0.010 | 24,646–27,816 |
| BA | −0.31 | −6.85*** | 18.63 | 19.88 | 0.009 | 0.009 | 54,200–61,171 |
| ES | −0.62 | −1.47 | 16.19 | 2.58 | 0.056 | 0.051 | – |
| Total | −0.82 | −24.16*** | 21.34 | 6.30 | – | – | 17,175–19,385 |
| *L. jocu* | | | | | | | |
| CE | −0.84 | −16.52*** | 19.02 | 12.81 | – | – | 37,970–42,854 |
| RN | −1.10 | −23.38*** | 19.86 | 16.96 | – | – | 50,132–55,577 |
| BA | −0.76 | −4.96** | 16.98 | 15.66 | – | – | 42,276–52,229 |
| ES | −0.60 | −4.14* | 16.57 | 12.40 | – | – | 36,783–41,515 |
| Total | −1.38 | −24.13*** | 24.33 | 14.55 | – | – | 43,161–48,713 |

*In all statistical tests, significance was assessed using 10,000 permutations or bootstrap replicates. Tajima's D, Tajima's neutrality test (Tajima, 1989); Fu's $F_S$ (Fu, 1997); W, Watterson's theta; τ = unit of population growth; SSD, sums of squared deviations; ty, thousands of years. * < 0.05, ** < 0.01, *** < 0.001.*

Along the Brazilian coast, *L. analis* and *L. jocu* can be considered as a unique panmictic unit, suggesting shared evolutionary responses to similar environmental pressures and historical events to which they were subjected. In fact, the $F_{st}$ values, based on the comparative analysis of the mitochondrial DNA HVR1 control region sequences, do not indicated genetic structuring in *L. analis* and *L. jocu*. Analyzes based on single locus markers can be prone to peculiarities of the selected marker (Hurst and Jiggins, 2005; Allio et al., 2017), and some patterns need be assessed by using other unlinked markers (Godinho et al., 2008). Nevertheless, analyzes with *cytB* sequences suggested some genetic structuring of the ES population of *L. analis*, located in the southernmost sampled site.

The level of genetic structuring in the marine environments seems to be the result of complex historical interactions due to interregional connectivity, physical barriers, restricted migratory behavior, short pelagic larval phase, $N_{ef}$ size, species ecology, and ocean current actions (Planes et al., 2001; Stepien et al., 2001). In contrast, panmixia is the predominant pattern challenging these factors in a large number of marine species (Liu et al., 2016, 2017), such as other lutjanids of the Brazilian province as *L. purpureus* (Salles et al., 2006; Gomes et al., 2008) and *Ocyurus chrysurus* (Vasconcellos et al., 2008).

Ocean currents have important biogeographic implications in the process of genetic differentiation of marine populations. In the Western Atlantic, two main ocean currents derived from the South Equatorial Current, the North Brazil one flowing in the northeastern/northern direction (RN – CE) and the Brazil Current flowing in a northeastern/southeastern direction (RN – BA – ES) (Lumpkin and Garzoli, 2005), dominate the areas of *L. analis* and *L. jocu* occurrence. In particular, the Brazil Current, with its northern and southern branches and extensive spatial trajectory, covers vast areas of both species' distributions, favoring gene flow among populations.

Besides the physical factors acting in the Western Atlantic, the absence of genetic structuring is associated with biological and ecological characteristics of the species involved (Lecchini and Galzin, 2003). In this respect, *L. analis* and *L. jocu* share biological features that favor long-distance dispersal with others lutjanid species. Among these are included the relatively long pelagic larval phase of approximately one month (Lindeman et al., 2001), high dispersive potential of adult individuals, ontogenetic habitat variation with a direct relationship between body size and depth (Frédou and Ferreira, 2005), formation of reproductive aggregations (Paris et al., 2005, França and Olavo, 2015), long life, and delayed sexual development (Claro, 1981; Allen, 1985; Graham et al., 2008).

Several marine organisms synchronize the release of eggs and larvae by forming spawning aggregations during specific periods. Like others, lutjanids, *L. analis* and *L. jocu* are r-strategists that form reproductive aggregations with large densities of individuals in spawning areas (Carter and Perrine, 1994; Lindeman et al., 2001; Paris et al., 2005). This reproductive strategy, although presenting considerable levels of self-recruitment in some areas, contributes larvae to distant populations, promoting increases in the genetic variability and homogeneity of the populations (Paris et al., 2005) presented by both species.

Ecological abilities to explore and colonize habitats can also increase genetic cohesion among *L. analis* e *L. jocu* populations along the Brazilian coast. In this respect, continuous environments (environmental corridors) favorable for the maintenance of different ontogenetic phases of the species are found throughout their distribution areas, such as rocky and sandy bottoms and reefs which extend for approximately 3000 km along the NE coast (Maida and Ferreira, 1997), as well as mangroves and estuaries (Cocheret de la Morinière et al., 2003), generating favorable habitats for both species. Additionally, the role of deep reef environments in the formation of effective corridors cannot be ruled out when establishing large population interchanges, even during periods of pronounced marine regression such as during the last glacial maximum (LGM).

The indication of population structure for *L. analis* between RN-ES based on *cytB* sequences may likely be associated with historical-geographic conditions (Pinheiro et al., 2015), and low contemporary gene flow. These regions have particular geological and geographic aspects that, under historical ocean level changes, could be responsible for particular demographic and genetic lineage divergences. The Brazilian continental shelf adjacent to the Rio Grande do Norte State has a reduced width and shallow depths as compared with other parts of the Brazilian shelf (Vital et al., 2010). In front of the RN coast extends a submarine volcanic chain, paralleling the northern equatorial coast (east-west direction), and known as the Fernando de Noronha Ridge. This geomorphologic formation consists of an alignment of at least, six volcanic seamounts, which extends in a direction toward emersed areas of the Rocas Atoll and Fernando de Noronha Archipelago. Historical asynchronic demographic patterns involving reef fishes related to these continental and insular areas have been attributed

to changes in eustatic sea levels (Souza et al., 2015). During transgression periods, the insular areas were reduced, while demographic expansion in the continental populations had occurred. In contrast, during historical regressions, the insular areas were increased, favoring demographic expansions. In fact, despite the limited information on genetic structuring of reef fish groups in these insular and continental areas, population genetic data for some reef species have shown evidence of differentiated genetic structures among them (Cunha et al., 2014; Souza et al., 2015).

The higher haplotype diversity and the occurrence of a more frequent haplotype in RN and, to a minor degree, to northeastern populations support the possible role of Fernando de Noronha Ridge as an historical functional refuge zone for genetic lineages of *L. analis* during the sea level changes that occurred in the Pleistocene epoch. Similarly, the continental shelf in the ES region is narrow, also with a sequential chain of underwater seamounts denominating the Vitória-Trindade Seamount Chain (VTC), where the Trindade Island and Martin Vaz Archipelago are the unique emersed areas. During the last oceanic transgression, the VTC also acted as refuge areas, promoting genetic diversification of evolutionary lineages (Pinheiro et al., 2015). Likely, temporary isolation of *L. analis* lineages in both aforementioned north and south seamount regions created conditions contributing to the genetic divergences detected by *cytB* analysis.

Other lutjanids species, such the Atlantic species *Ocyurus chrysurus*, represent a large genetic unit (Vasconcellos et al., 2008), similarly to *L. campechanus* in the Caribbean, Gulf of Mexico and Florida coast (Garber et al., 2004), or *L. kasmira* in Indo-Pacific areas, where no signs of genetic structuring along the more than 12,000 km distribution exist, except for the Marquesas Archipelago (Gaither et al., 2010). In fact, panmictic genetic populations are mainly recurrent in lutjanids from other marine regions in the absence of apparent physical barriers.

## DEMOGRAPHIC HISTORY

Both species presented values of $N_{ef}$ high enough to maintain a balance between the loss of variability by genetic drift and its replacement by new mutations. However, was evidenced a marked discrepancy between the species, in which *L. jocu* presented higher $N_{ef}$ values than *L. analis*.

Differences in contemporary or historically available habitats of *L. analis* and *L. jocu*, affected by recurrent cycles of retraction and sea level rise, may be linked to this demographic divergence pattern. Indeed, the widespread effects of such events have been perceived in several marine fishes (Grant and Bowen, 1998; Benzie, 1999; Lourie and Vincent, 2004), including lutjanid species such as *L. kasmira*, *L. fulvus* (Gaither et al., 2010) and *L. erythropterus* (Zhang et al., 2006). The demographic history of *L. analis* and *L. jocu* seems to be related to changes during the Pleistocene involving marine environments. Both species present evidence of population expansion along the Brazilian coast, as indicated by results of the Tajima (D) and Fu (Fs) tests. The measure of Fu's Fs statistic, based on the HVR1 control region sequences, revealed negative values significantly different from zero in all localities, thus supporting recent sudden population expansions (Ramos-Onsins and Rozas, 2002). Raggedness indices and SSD applied to searches for additional demographic signs of population expansion in *L. analis* were also in concordance with a historically expanding pattern.

The *L. analis* and *L. jocu* populations presented bimodal mismatch distributions that are usually associated with constant population size. However, this may also indicate the presence of two distinct lineages (e.g., Alvarado-Bremer et al., 2005). In this sense, the networks indicated more than one distinct lineage for bimodal mismatches (*L. analis* and *L. jocu*). The data suggested that past fragmentation was followed by population expansions resulting in secondary contacts of genetically diverged populations that had been separated by historical sea level changes during the Pleistocene. The population expansion period was significantly asynchronous between the two species (*L. analis* – ≅17.1–19.4 Kybp; *L. jocu* – ≅43.1–48.7 Kybp), but consistent with changes which occurred during the last glacial period between 120 and 12 Kybp (Martinson et al., 1987). Events in this period are considered important agents in the demographic patterns of several other lutjanids (Zhang et al., 2006; Gaither et al., 2010).

Due to exploration of different environment depths, the rise and retraction of sea level certainly promoted differential historical habitat conditions for *L. analis* and *L. jocu*, respectively. The population expansion of *L. analis* occurred during the beginning of the LGM period, in which ice sheets reached their maximum volume, between 19 and 26.5 Kybp (Martinson et al., 1987; Clark et al., 2009). This event possibly supplied a geographic expansion of *L. analis* by providing access to new available ocean areas. In contrast, as *L. jocu* occupies varied shallow habitats (mangroves, estuaries, and reefs) during its ontogeny, with eventual migration offshore with age (Moura et al., 2011), it must have undergone a differentiated demographic impact by changes in these ecologically complex habitats.

High $N_{ef}$ values were found in *L. analis* and especially in *L. jocu* populations. This population size discrepancy is an indication of considerably different niches available for these species. In fact, the vast and structurally complex shallow coastal habitats occupied by *L. jocu* have been shown to favor greater effective population sizes. Both species have their larger population effectives in northeast populations (CE, RN, and BA), with a drastic reduction in the ES region, near the southern limit of their geographic distribution.

The estimation of a large female contingent connected with reproduction, in association with high haplotype diversity, suggested potential resilience to present over-exploitation of both species. In fact, a high $N_{ef}$ value contributes to maintaining the balance between loss of adaptive genetic variation due to genetic drift and its restoration by mutation (Schultz and Lynch, 1997), and this is linked to the evolutionary feasibility of species. In this way, it is a useful indicator of population response to evolutionary and ecological forces (Waples, 2010), and environmental disturbances (Anderson, 2005). Besides providing useful genetic comparative indications for marine populations (Hare et al., 2011), the present $N_{ef}$ estimates are similar to those of the Indo-Pacific lutjanids (Gaither et al., 2010).

## CONCLUSION

Genetic patterns indicate that the mutton and dog snappers exhibit panmictic populations along the Brazilian coast with a high level of genetic variability and large effective population sizes. Despite the absence of genetic structuring, as revealed by HVR1 region sequence analysis, for all populations of both species, there are indications of genetic structuring concerning *L. analis* populations of the RN and ES related to *cytB* sequences. This is likely associated to genetic divergence in refuge regions during Pleistocene glaciations. Historical contingencies in this period were also responsible for demographic expansions events in these species. In summary, the importance of spawning aggregations in maintaining high genetic variability along large areas of distribution, particular attention should be given to protecting reproductive areas against overfishing and other anthropogenic disturbance. The present data provide useful information for the sustainable exploitation of these overfished species, and create parameters for future monitoring of natural stocks.

## DATA AVAILABILITY

The datasets generated for this study can be found in the GenBank access numbers JQ727916–JQ72996.

## ETHICS STATEMENT

The animal study was reviewed and approved by Committee of Ethics in the Use of Animals of the Federal University of Rio Grande (# 044/2015).

## AUTHOR CONTRIBUTIONS

ED and WM contributed to the conception and design of the study. AS, ED, and MP organized the database and performed the statistical analysis. AS, ED, and WM wrote the first draft of the manuscript. MP, EG-M, MV, MC, LB, and PG wrote sections of the manuscript. All authors contributed to the manuscript revision, and read and approved the submitted version.

## ACKNOWLEDGMENTS

The authors wish to thank the Brazilian artisanal fishing communities involved in the study area, and the biologists Priscilla Malafaia and Flávio P. Filho for their help with sample collection, as well as the Brazilian research institutes CNPq and CAPES.

## REFERENCES

Allen, G. R. (1985). FAO Species Catalogue, vol. 6, Snappers of the world. An annotated and illustrated catalogue of lutjanid species known to date. *FAO Fish. Synop.* 6:208. doi: 10.1371/journal.pone.0087359

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

Alvarado-Bremer, J. R., Viñas, J., Mejuto, J., Ely, B., and Pla, C. (2005). Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Mol. Phylogenet. Evol.* 36, 169–187. doi: 10.1016/j.ympev.2004.12.011

Anderson, E. C. (2005). An efficient Monte Carlo method for estimating Ne from temporally spaced samples using a coalescent-based likelihood. *Genetics* 170, 955–967. doi: 10.1534/genetics.104.038349

Ault, J. S., Bohnsack, J. A., and Meester, G. A. (1998). A retrospective (1979-1996) multispecies assessment of coral reef fish stocks in the Florida Keys. *Fish. Bull.* 96, 395–414.

Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22, 341–345. doi: 10.1093/bioinformatics/bti803

Beerli, P., and Palczewski, M. (2010). Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185, 313–326. doi: 10.1534/genetics.109.112532

Benzie, J. A. H. (1999). Genetic structure of coral reef organisms: ghosts of dispersal past. *Am. Zool.* 39, 131–145. doi: 10.1093/icb/39.1.131

Blaber, S. J. M., Dichmont, C. M., Buckworth, R. C., Badrudin, H., Sumiono, B., Nurhakim, S., et al. (2005). Shared Stocks of snappers (Lutjanidae) in Australia and Indonesia: integrating biology, population dynamics and socio-economics to examine management scenarios. *Rev. Fish Biol. Fishe.* 15:111. doi: 10.1007/s11160-005-3887-y

Burton, M. L. (2002). Age, growth and mortality of mutton snapper, *Lutjanus analis*, from the east coast of Florida, with a brief discussion of management implications. *Fish. Res.* 59, 31–41. doi: 10.1016/s0165-7836(02)00007-3

Burton, M. L., Brennan, K. J., Munoz, R. C., and Parker, R. O. (2005). Preliminary evidence of increased spawning aggregations of mutton snapper (*Lutjanus analis*) at Riley's Hump two years after establishment of the Tortugas South ecological reserve. *Fish. Bull.* 103, 404–410.

Carter, J., and Perrine, D. (1994). A spawning aggregation of dog Snapper, *Lutjanus jocu* (Pisces, Lutjanidae) in Belize, Central-America. *Bull. Mar. Sci.* 55, 228–234.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., et al. (2009). The last glacial maximum. *Science* 325, 710–714. doi: 10.1126/science.1172873

Claro, R. (1981). Ecologia y ciclo de vida del pargo criollo, *Lutjanus analis*, en la plataforma cubana. *Academia de Ciencias de Cuba.* 186, 1–83.

Clement, M., Posada, D., and Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659. doi: 10.1046/j.1365-294x.2000.01020.x

Cocheret de la Morinière, E., Pollux, B. Y. A., Nagelkerken, I., and Van der Velde, G. (2003). Diet shifts Caribbean grunts (Haemulidae) and snappers (*Lutjanidae*) and the relation with nursery-to-coral reef migrations. *Estuar. Coast. Shelf Sci.* 57, 1079–1089. doi: 10.1016/s0272-7714(03)00011-8

Cunha, M. I, Souza, A. S., Dias, E. A. Jr., Amorim, K. D., Soares, R. X., Costa, G. W. W. F., et al. (2014). Genetic multipartitions based on D-Loop sequences and chromosomal patterns in brown chromis, Chromis multilineata (Pomacentridae), in the Western Atlantic. *Biomed Res. Int.* 2014:254698. doi: 10.1155/2014/254698

Dupanloup, I., Schneider, S., and Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11, 2571–2581. doi: 10.1046/j.1365-294x.2002.01650.x

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Excoffier, L., and Lischer, H. E. L. (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. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

França, A. R., and Olavo, G. (2015). Indirect signals of spawning aggregations of three commercial reef fish species on the continental shelf of Bahia, east coast of Brazil. *Braz. J. Oceanogr.* 63, 289–302.

Frédou, T., and Ferreira, B. P. (2005). Bathymetric trends of Northeastern Brazilian snappers (Pisces, *Lutjanidae*): implications for the reef fishery dynamic. *Braz. Arch. Biol. Technol.* 48, 787–800. doi: 10.1590/s1516-89132005000600015

Frédou, T., Ferreira, B. P., and Letourneur, Y. (2009). Assessing the stocks of the primary snappers caught in Northeastern Brazilian Reef Systems. 2-A multi-fleet age-structured approach. *Fish. Res.* 99, 97–105. doi: 10.1016/j.fishres.2009.05.009

Froese, R., and Binohlan, C. (2000). Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. *J. Fish Biol.* 56, 758–773. doi: 10.1006/jfbi.1999.1194

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Gaither, M. R., Bowen, B. W., Toonen, R. J., Planes, S., Messmer, V., Earle, J., et al. (2010). Genetic consequences of introducing allopatric lineages of *Bluestriped snapper* (*Lutjanus kasmira*) to Hawaii. *Mol. Ecol.* 19, 1107–1121. doi: 10.1111/j.1365-294X.2010.04535.x

Garber, A., Tringali, M., and Stuck, K. (2004). Population structure and variation in Red snapper (Lutjanus campechanus) from the Gulf of Mexico and Atlantic coast of Florida as determined from mitochondrial DNA control region sequence. *Mar. Biotechnol.* 6, 175–185.

Godinho, R., Crespo, E. G., and Ferrand, N. (2008). The limits of mtDNA phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. *Mol. Ecol.* 17, 4670–4683. doi: 10.1111/j.1365-294X.2008.03929.x

Gold, J. R., Voelker, G., and Renshaw, M. A. (2011). Phylogenetic relationships of tropical western Atlantic snappers in subfamily *Lutjaninae* (*Lutjanidae*: Perciformes) inferred from mitochondrial DNA sequences. *Biol. J. Linn Soc.* 102, 915–929. doi: 10.1111/j.1095-8312.2011.01621.x

Gomes, G., Schneider, H., Vallinoto, M., Santos, S., Orti, G., and Sampaio, I. (2008). Can *Lutjanus purpureus* (South red snapper) be "legally" considered a Red Snapper (Lutjanus campeghanus). *Genet. Mol. Biol.* 31, 372–376. doi: 10.1590/s1415-47572008000200035

Graham, R. T., Carcamo, R., Rhodes, K. L., Roberts, C. M., and Requena, N. (2008). Historical and contemporary evidence of a mutton snapper (*Lutjanus analis* Cuvier, 1828) spawning aggregation fishery in decline. *Coral Reefs.* 27, 311–319. doi: 10.1007/s00338-007-0329-4

Grant, W. S., and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Am. Genetic Assoc.* 89, 415–426. doi: 10.1093/jhered/89.5.415

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., et al. (2011). Understanding and estimating effective population size for practical application in marine species management. *Conserv. Biol.* 25, 438–449. doi: 10.1111/j.1523-1739.2010.01637.x

Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.

Heist, E. J., and Gold, J. R. (2000). DNA microsatellites and genetic structure of red snapper (Lutjanus campechanus) in the Gulf of Mexico. *Trans. Am. Fish Soc.* 129, 469–475.

Hurst, G. D., and Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R Soc. B Biol. Sci.* 272, 1525–1534. doi: 10.1098/rspb.2005.3056

IUCN, (2019). *IUCN Red List of Threatened Species. Version 2011.2*. Available at: www.iucnredlist.org (accessed May 14, 2019).

Jiménez, S., Schönhuth, S., Lozano, I. J., González, J. A., Sevilla, R. G., Diez, A., et al. (2007). Morphological, ecological, and molecular analyses separate *Muraena augusti* from *Muraena helena* as a valid species. *Copeia.* 2007, 101–113. doi: 10.1643/0045-8511(2007)7%5B101:meamas%5D2.0.co;2

Karlsson, S., Saillant, E., and Gold, J. R. (2009). Population structure and genetic variation of lane snapper (Lutjanus synagris) in the northern Gulf of Mexico. *Mar. Biol.* 156, 1841–1855. doi: 10.1007/s00227-009-1217-y

Klippel, S., Costa, P. A. S., Martins, A. S., and Peres, M. B. (2005). "Avaliação dos estoques de lutjanídeos da costa central do Brasil: análise de coortes e modelo preditivo de Thompson e Bell para comprimentos," in *Pesca e Potenciais de Exploração de Recursos Vivos na Região Central da Zona Econômica Exclusiva Brasileira*, eds P. A. S. Costa, A. S. Martins, and G. Olavo, (Rio de Janeiro: Museu Nacional do Rio de Janeiro), 83–98.

Lecchini, D., and Galzin, R. (2003). Influence of pelagic and benthic, biotic and abiotic, stochastic and deterministic processes on the dynamics of auto-recruitment of coral reef fish. *Cybium* 27, 167–184.

Lindeman, K. C., Lee, T. N., Wilson, W. D., Claro, R., and Ault, J. S. (2001). Transport of larvae originating in southwest Cuba and the Dry Tortugas: evidence for partial retention in grunts and snappers. *Proc. Gulf Caribb. Fish. Inst.* 52, 732–747.

Liu, B. J., Zhang, B. D., Gao, T. X., and Liu, J. X. (2017). Signatures of temporal genetic differentiation of the small yellow croaker (Larimichthys polyactis) revealed by microsatellite markers. *Fish. Res.* 194, 50–54. doi: 10.1016/j.fishres.2017.05.005

Liu, B. J., Zhang, B. D., Xue, D. X., Gao, T. X., and Liu, J. X. (2016). Population structure and adaptive divergence in a high gene flow marine fish: the small yellow croaker (*Larimichthys polyactis*). *PLoS One* 11:e0154020. doi: 10.1371/journal.pone.0154020

Lourie, S. A., and Vincent, A. C. J. (2004). A marine fish follows Wallace's Line: the phylogeography of the three-spot seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in South-east Asia. *J. Biogr.* 31, 1975–1985. doi: 10.1111/j.1365-2699.2004.01153.x

Lumpkin, R., and Garzoli, S. L. (2005). Near-surface circulation in the Tropical Atlantic ocean. *Deep sea Res.* 52, 495–518. doi: 10.1038/s41598-019-39392-y

Maida, M., and Ferreira, B. P. (1997). "Coral reefs of Brazil: Overview and field guide," in *Proceedings of the 8th International Coral Reef Symposium*, Panama, FL, 263–274.

Martinson, D. G., Pisias, N. G., Hays, J. D., Imbrie, J., Moore, T. C., and Shackleton, N. J. Jr. (1987). Age dating and the orbital theory of the Ice Ages: development of a high-resolution 0–300000-year chronostratigraphy. *Quat. Res.* 27, 1–29. doi: 10.1016/0033-5894(87)90046-9

Mattos, S. M. G., and Maynou, F. (2009). Virtual population analysis of two snapper species, *lutjanus analis* and *lutjanus chrysurus*, caught off Pernambuco state, North-Eastern Brazil. *Braz. J. Oceanogr.* 57, 229–242. doi: 10.1590/s1679-87592009000300006

Morin, P. A., Martien, K. K., Archer, F. I., Cipriano, F., Steel, D., Jackson, J., et al. (2010). Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. *J. Hered.* 101, 1–10. doi: 10.1093/jhered/esp107

Moura, R. L., Francini-Filho, R. B., Chaves, E. M., Minte-Vera, C. V., and Lindeman, K. C. (2011). Use of riverine through reef habitat systems by *Dog snapper* (Lutjanus jocu) in eastern Brazil. *Estuar. Coast. Shelf S.* 95, 274–278. doi: 10.1016/j.ecss.2011.08.010

Moura, R. L., and Lindeman, K. C. (2007). A new species of snapper (Perciformes: Lutjanidae) from Brazil, with comments on the distribution of *Lutjanus griseus* and *L. apodus*. *Zootaxa* 1422, 31–43. doi: 10.11646/zootaxa.1422.1.2

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

Ovenden, R. J., and Street, R. (2003). Genetic population structure of mangrove jack, *Lutjanus argentimaculatus* (Forsskall). *Mar. Freshwater Res.* 54, 127–137.

Paris, C. B., Cowen, R. K., Claro, R., and Lindeman, K. C. (2005). Larval transport pathways from *Cuban snapper* (Lutjanidae) spawning aggregations based on biophysical modeling. *Mar. Ecol. Prog. Ser.* 296, 93–106. doi: 10.3354/meps296093

Pinheiro, H. T., Mazzei, E., Moura, R. L., Amado-Filho, G. M., Carvalho-Filho, A., and Braga, A. C. (2015). Fish biodiversity of the Vitória-Trindade seamount chain, Southwestern Atlantic: an updated database. *PLoS One* 10:e0118180. doi: 10.1371/journal.pone.0118180

Planes, S., Doherty, P., and Bernardi, G. (2001). Unusual case of extreme genetic divergence in a marine fish, *Acanthochromis polyacanthus*, within the Great Barrier Reef and the Coral Sea. *Evolution* 55, 2263–2273. doi: 10.1111/j.0014-3820.2001.tb00741.x

Preveiro, M., Minte-Vera, C. V., Freitas, M. O., Moura, R. L., and Dei Tos, C. (2011). Age and growth of the *Dog snapper Lutjanus jocu* (Bloch & Schneider,

1801) in Abrolhos Bank, Northeastern Brazil. *Neotrop. Ichthyol.* 9, 393–401. doi: 10.7717/peerj.3167

Ramos-Onsins, S. E., and Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19, 2092–2100. doi: 10.1093/oxfordjournals.molbev.a004034

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W. J. (2009). Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish Fish.* 10, 361–395. doi: 10.1111/j.1467-2979.2008.00324.x

Rezende, S. M., Ferreira, B. P., and Frédou, T. (2003). A pesca de lutjanídeos no nordeste do Brasil: Histórico das pescarias, características das espécies e relevância para o manejo. *Boletim Técnico Científico CEPENE.* 11, 257–270.

Rogers, A. R. (1995). Genetic evidence for a Pleistocene population expansion. *Evolution* 49, 608–615. doi: 10.1111/j.1558-5646.1995.tb02297.x

Rogers, A. R., and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–569.

Salles, R., Fonteles-Filho, A. A., Furtado-Neto, M. A. A., Carr, S. M., and Freitas, S. M. (2006). Morphometric and mitochondrial DNA analyses of the Caribbean red snapper, *Lutjanus purpureus* (Teleostei, Lutjanidae), in western Atlantic off Northern Brazil. *Bol. Inst. Pesca.* 32, 115–125.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. New York, NY: Cold-Springer Harbor Laboratory Press.

Santos, A. M., Cabezas, M. P., Tavares, A. I., Xavier, R., and Branco, M. (2015). tcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics* 32, 627–628. doi: 10.1093/bioinformatics/btv636

Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., and Saccone, C. (1997). Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205, 125–140. doi: 10.1016/s0378-1119(97)00404-6

Schneider, S., and Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152, 1079–1089.

Schultz, S. T., and Lynch, M. (1997). Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* 51, 1363–1371. doi: 10.1111/j.1558-5646.1997.tb01459.x

Souza, A. S., Dias, E. A. D. Jr., Galetti, P. M. Jr., Machado, E. G., Pichorim, M., and Molina, W. F. (2015). Wide-range genetic connectivity of Coney, *Cephalopholis fulva* (Epinephelidae), through oceanic islands and continental Brazilian coast. *An. Acad. Bras. Cienc.* 87, 121–136. doi: 10.1590/0001-3765201520130411

Stepien, C. A., Rosenblatt, R. H., and Bargmeyer, B. A. (2001). Phylogeography of the spotted sand bass, Paralabrax maculatofasciatus: divergence of Gulf of California and Pacific coast populations. *Evolution* 55, 1852–1862. doi: 10.1111/j.0014-3820.2001.tb00834.x

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Teixeira, S. F., Duarte, Y. F., and Ferreira, B. P. (2010). Reproduction of the fish *Lutjanus analis* (mutton snapper; Perciformes: Lutjanidae) from Northeastern Brazil. *Rev. Biol. Trop.* 58, 791–800.

Vasconcellos, A. V., Vianna, P., Paiva, P. C., Schama, R., and Solé-Cava, A. (2008). Genetic and morphometric differences between *Yellowtail snapper* (*Ocyurus chrysurus*, Lutjanidae) populations of the tropical West Atlantic. *Genet. Mol. Biol.* 31, 308–316. doi: 10.1590/s1415-47572008000200026

Vital, H., Gomes, M. P., Tabosa, W. F., Frazão, E. P., Santos, C. L. A., and Plácido Júnior, J. S. (2010). Characterization of the Brazilian continental shelf adjacent to Rio Grande do Norte state, NE Brazil. *Braz. J. Oceanogr.* 58, 43–54. doi: 10.1590/s1679-87592010000500005

Waples, R. S. (2010). Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Mol. Ecol. Resour.* 10, 785–796. doi: 10.1111/j.1755-0998.2010.02876.x

Watanabe, W. O., Ellis, E. P., Ellis, S. C., Chaves, J., and Manfredi, C. (1998). Artificial propagation of mutton snapper *Lutjanus analis*, a new candidate marine fish species for aquaculture. *J. World Aquacult. Soc.* 29, 176–187. doi: 10.1111/j.1749-7345.1998.tb00977.x

Xia, X., and Xie, Z. (2001). DAMBE: data analysis in molecular biology and evolution. *J. Hered.* 92, 371–373. doi: 10.1093/jhered/92.4.371

Zhang, J., Cai, Z., and Huang, L. (2006). Population genetic structure of crimson snapper *Lutjanus erythropterus* in East Asia, revealed by analysis of the mitochondrial control region. *ICES J. Mar. Sci.* 63, 693–704. doi: 10.1016/j.icesjms.2006.01.004

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 01 October 2019
doi: 10.3389/fmars.2019.00615

# Cryptic Diversity and Database Errors Challenge Non-indigenous Species Surveys: An Illustration With *Botrylloides* spp. in the English Channel and Mediterranean Sea

*Frédérique Viard[1]*, Charlotte Roby[1], Xavier Turon[2], Sarah Bouchemousse[1] and John Bishop[3]*

*[1] Laboratory AD2M (UMR7144), Station Biologique de Roscoff, Sorbonne Université, CNRS, Roscoff, France, [2] Department of Marine Ecology, Centre for Advanced Studies of Blanes (CEAB, CSIC), Blanes, Spain, [3] Citadel Hill Laboratory, Marine Biological Association of the United Kingdom, Plymouth, United Kingdom*

Molecular tools have been extensively used in recent decades to examine biological invasion processes, and are increasingly being adopted as efficient tools to support non-indigenous species surveys, notably through barcoding approaches, i.e., the use of a reference sequence specific to a given species to validate its identification. The technique is easy to use but requires reliable reference sequences to be available in public databases. In addition, the increasing discovery of cryptic species in marine taxa may complicate taxonomic assignment. We illustrate these two issues in the ascidian genus *Botrylloides*, in which at least three global marine invaders have been recognized, including *B. violaceus* and *B. diegensis*. We obtained COI sequences from >750 colonies of *Botrylloides* spp. sampled in W Europe or provided by expert colleagues from other regions. Phylogenetic trees clearly distinguished our targeted taxa [i.e., *B. violaceus*, *B. diegensis* and *B. leachii* (native)]. They also revealed another discrete lineage apparently related to a recently described eastern Mediterranean species. By examining public databases, we found sequences of *B. diegensis* erroneously assigned to *B. leachii*. This observation has major implications as the introduced *B. diegensis* can be misidentified as a putatively native species. We also checked published sequences of the genus *Botrylloides* in the Mediterranean Sea, complemented with new samples. Based on our custom reference database, all published sequences of *B. leachii* corresponded to *B. diegensis*, although this NIS has hardly been reported at all in the Mediterranean region. Such database errors are unfortunate, as the barcoding approach is a powerful tool to identify the recognized *Botrylloides* species currently present in European seas. This is of particular importance because a trait often used during field assessment, i.e., single-color vs. two-color colonies, is misleading to distinguish *B. violaceus* and *B. diegensis* respectively: a substantial proportion of the single-color morph are actually *B. diegensis* in both the Mediterranean Sea and the English Channel. Altogether, this study exemplifies the advantages and disadvantages of molecular barcoding in NIS surveys and studies. The limitations that were identified are all easy to resolve once proper vouchers and collections are set up.

**Keywords: non-native species, cryptic species, barcoding, COI, marine, tunicates**

OPEN ACCESS

***Edited by:***
*Tifeng Shan,*
*Institute of Oceanology, Chinese Academy of Sciences, China*

***Reviewed by:***
*Néstor E. Ardila,*
*EcoMar s.a.s, Colombia*
*Alfonso Angel Ramos-Esplá,*
*University of Alicante, Spain*

****Correspondence:***
*Frédérique Viard*
*viard@sb-roscoff.fr*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *22 June 2019*
***Accepted:*** *18 September 2019*
***Published:*** *01 October 2019*

***Citation:***
*Viard F, Roby C, Turon X, Bouchemousse S and Bishop J (2019) Cryptic Diversity and Database Errors Challenge Non-indigenous Species Surveys: An Illustration With Botrylloides spp. in the English Channel and Mediterranean Sea. Front. Mar. Sci. 6:615. doi: 10.3389/fmars.2019.00615*

## INTRODUCTION

In marine environments, biological introductions exert important pressure on natural ecosystems. Introduction of non-indigenous species (NIS) has occurred at an increasing rate since the 20th century, in pace with the increasing range and intensity of vectors (Nunes et al., 2014). Because successful eradications of invasive species in marine systems are very rare (Sambrook et al., 2014; Ojaveer et al., 2015), both pre-border prevention measures (e.g., limiting the introduction of NIS by ballast waters) and early detection focused on introduction hotspots (e.g., marinas and harbors) have been recommended (Ojaveer et al., 2014). Accurate identification is a pre-requisite for the implementation of policies and regulations such as the Marine Framework Strategy Directive (MSFD), in which Descriptor 2 is dedicated to NIS, with one criterion involving the detection and identification of new NIS, and another concerning the changing spatial extent of already-reported NIS.

However, identifying new or recently introduced NIS can be challenging when based only on traditional methods (Darling et al., 2017), such as rapid assessment surveys (RAS), which rely heavily on field recognition of species and have been commonly used to survey NIS in marinas and harbors (e.g., Cohen et al., 2005; Arenas et al., 2006; Campbell et al., 2007; Bishop et al., 2015a; Lehtiniemi et al., 2015). Many marine animal NIS in introduction hotspots belong to taxonomic groups, such as bryozoans, hydrozoans and tunicates, that require substantial taxonomic expertise. In addition, the development of molecular studies in the last three decades has revealed a large number of cryptic species (i.e., species that are not distinguishable based on morphological traits) in these taxonomic groups (Knowlton, 2000; Appeltans et al., 2012). Consequently, an increasing number of studies have recommended the use of molecular tools to complement the traditional methods, and thus achieve reliable taxonomic identification of marine NIS (Comtet et al., 2015; Darling et al., 2017; Dias et al., 2017).

The usefulness of the molecular barcoding approach is well established. The benefits of the approach have been demonstrated by numerous studies (for a review, see Comtet et al., 2015). Such an approach can be used to ascertain the presence of a new NIS (e.g., *Asterocarpa humilis*: Bishop et al., 2013), to reveal false morphology-based NIS identification (e.g., *Crepidula* spp.: McGlashan et al., 2008) or to determine the taxonomic status of previously unrecognized NIS (Ordóñez et al., 2016). However, the robustness of the molecular barcoding approach relies on some important pre-requisites, notably the availability of reliable reference sequence data for the targeted species. Thus, the prompt updating of databases is essential following taxonomic revision and discovery of unrecognized cryptic species. The sequences delivered in public databases such as GenBank (Benson et al., 2013) or BOLD [Barcode of Life Data System; (Ratnasingham and Hebert, 2007)] should thus be error-free and validated by taxonomic experts. BOLD is expected to provide such a framework. However, the public portal also retrieves non-curated sequences from Genbank that can propagate errors. There is so far a limited number of initiatives specifically aimed at delivering reference data for marine NIS. Although rare, such initiatives should be strongly encouraged, notably to support international regulations and policies such as the MSFD in Europe (Darling et al., 2017). Such an initiative has recently been undertaken in Western Australia, where a reference collection (vouchers, DNA and sequences for one marker (COI)) was assembled for 75 species (out of 79) of the "Western Australian Prevention List for Introduced Marine Pests" (Dias et al., 2017). Similarly, facing the issue of cryptic diversity, integrative taxonomy coupling molecular phylogenetics with morphology and other species attributes (e.g., ecology, life-history traits) is needed (Pante et al., 2015).

We here illustrate the issues described above using the colonial tunicate genus *Botrylloides* as a particularly relevant case study. At least two *Botrylloides* species, *B. violaceus* Oka, 1927 and *B. diegensis* Ritter and Forsyth, 1927, with the recent addition of *B. giganteus* (Pérés, 1949) (Rocha et al., 2019) are globally invasive. *B. violaceus* and *B. diegensis* are both native to the N Pacific. While the former is native to the NW Pacific, there is more uncertainty regarding the native range of the latter: although *B. diegensis* was originally described from the NE Pacific (southern California), it is likely that the species had been introduced from the western or southern Pacific (Carlton, 2009). The two species are important members of the fouling community colonizing artificial substrates on the Pacific coast of the United States, for instance in harbors and marinas (e.g., Nydam and Stachowicz, 2007). The same two species have also been introduced in Europe, notably in the English Channel (EC), where they are well established in marinas of the United Kingdom and N France (Bishop et al., 2015a,b). In Europe, one putatively native species, *B. leachii* (Savigny, 1816), is also recognized, often showing coloration somewhat similar to the two-tone color pattern seen in *B. diegensis* (**Supplementary Figure S3**, panel 6). The original description (Savigny, 1816) of *B. leachii* was based on material "*Communiqué par M*[onsieur]. *Leach, directeur du Musée britannique*" (i.e., William Elford Leach, then working in the Department of Natural History of the British Museum) and the specimen was thought by Savigny to have originated probably from the English coast ("*Habite les côtes de l'Angleterre?*"). Conversely to the introduced species, the native taxon seems uncommon in artificial habitats in some parts of the EC (J. Bishop, L. Lévêque, and F. Viard, pers. obs.).

The first reports of non-indigenous *Botrylloides* spp. in the English Channel were made after RAS in 2004 (Arenas et al., 2006). A single species, *B. violaceus*, was reported, recognized by its single-colored colonies and large and morphologically distinctive brooded larvae. Colonies of *B. violaceus* could occur in a range of single colors, including violet, cream, yellow (**Supplementary Figure S3**, panel 4), brick-red and (commonly) orange. A second species was encountered during the 2004 surveys, but was not reported by Arenas et al. (2006) since its identity was uncertain (Bishop et al., 2015a,b). This second species showed a strong and distinctive two-colored pattern featuring a broad ring of solid pigmentation (commonly orange or cream) surrounding each buccal orifice, against a contrasting darker background (**Supplementary Figure S3**, panel 2), and was subsequently identified as *B. diegensis*. This distinction, based on color morphs, was used by Lambert and Lambert (2003) to

distinguish *B. violaceus* and *B. diegensis* in southern California, United States. In 2011, the present authors and colleagues collected extensive samples on both sides of the English Channel for an intended population-genetic study of *B. violaceus*; single-color colonies were thus collected, with or without confirmation of the distinctive but seasonally brooded larvae of that species, under the assumption that these would all be the target species. Preliminary molecular analyses revealed that these single-color colonies included specimens attributable to *B. diegensis*, so that *B. diegensis* had been misidentified in the field as *B. violaceus* on the criterion of possessing single-colored, rather than two-colored, colonies. *B. diegensis* in the English Channel could thus occur in a single-color morphotype (e.g., **Supplementary Figure S3**, panel 1) easily confused with *B. violaceus*, even by those supposedly familiar with the taxa. In addition, during our surveys some color morphs believed to belong to *B. leachii* were encountered that had a two-color pattern similar to the two-color pattern of *B. diegensis*, although differing in detail (**Supplementary Figure S3**, panel 6).

Following from these preliminary observations and molecular results, in this study we examined a large number of *Botrylloides* colonies displaying a single-color pattern, using sequencing data to make an in-depth assessment of the association between color-morph and species. We chose to analyze a fragment of the mitochondrial gene region cytochrome oxidase I (COI) which has been previously investigated to study *B. violaceus* in its N American introduced range (Bock et al., 2011; Lejeusne et al., 2011), besides being a marker commonly used in barcoding studies of marine invertebrates (Comtet et al., 2015) and a standard for BOLD (Hebert et al., 2016). This dataset was completed with data from specimens kindly provided by taxonomic experts (see acknowledgment section); this allowed us to get reliable reference sequences for the three *Botrylloides* species, *B. violaceus*, *B. diegensis* and *B. leachii*, putatively present in the EC. We further analyzed if similar confusion existed in the Mediterranean, where six *Botrylloides* species have been reported: *B. leachii* (native), *B. violaceus* (introduced), *B. pizoni* Brunetti and Mastrototaro, 2012 (likely a synonym of the invasive *B. giganteus*; Reem et al., 2018), *B. niger* Herdman, 1886 (likely a Lessepsian migrant), *B. anceps* (Herdman, 1891) (likely a Lessepsian migrant), and *B. israeliense* Brunetti (2009) (uncertain status). To this end, we used all existing Mediterranean sequences of the genus in GenBank, complemented with new samples, for comparison with our reference sequences. The objectives of our work were four-fold: (i) ascertaining the distribution of the two NIS in the study EC area, (ii) providing further molecular reference data, specific to European seas, (iii) assessing if field identification, based on the typical single-color pattern of *B. violaceus*, is truly reliable, and (iv) updating Mediterranean records of *Botrylloides* spp. backed by sequence data.

# MATERIALS AND METHODS

## Specimen Collection

Colonies ($N = 627$) displaying a single color, as typically described for *B. violaceus*, were randomly sampled along floating pontoons, from April to September 2011, in 20 marinas along the southern coastline of England from Plymouth to Brighton (9 marinas) and along the coastlines of Brittany from Quiberon to Saint-Malo (11 marinas) (**Table 1** and **Figure 1**). RAS done in these same locations have shown the presence of non-indigenous *Botrylloides* spp. (Bishop et al., 2015a; FV and L. Lévêque, unpublished data). All the sampling was undertaken by SCUBA diving or from the surface by removing colonies from the pilings and pontoons in marinas. Samples were preserved in 96% ethanol prior to genetic analysis.

A reference collection of 128 specimens was gathered for analysis (**Supplementary Table S1**). It included a total of 84 specimens of *B. diegensis*, of which 39 were kindly provided by expert colleagues from the W coast of United States, Italy and New Zealand. The remainder were sampled in four marinas in Brittany. In every case, these colonies displayed the typical two-color pattern of *B. diegensis*. We also included 17 reference specimens for *B. leachii* from United Kingdom, Ireland and France. These latter specimens mostly displayed the color pattern of *Botrylloides radiata* Alder and Hancock, 1848, as illustrated by Alder and Hancock (1912), pl. 64, Figures 10 and 11), a form synonymized with *B. leachii* by Berrill (1950) and Millar (1970) (**Supplementary Figure S3**, panels 5 and 6). These specimens will be hereafter referred to as the 'radiata' morph of *B. leachii*. Nine typical violet *B. violaceus* colonies, also displaying the distinctive large larvae of this species, were obtained in two French localities. Finally, to be used as an outgroup, 18 specimens of *Botryllus schlosseri* were collected in United Kingdom and France. All these samples were preserved in 96% ethanol after their collection.

In order to update Mediterranean records, a third dataset was constructed, using samples obtained from the NE of the Iberian Peninsula (oyster cultures in the Ebro Delta, 40°46′27.43″N, 00°44′27.11″E), the S of France (marina of Sète, 43°23′44.2″N, 3°41′51.0″E), the SW of Italy (Miseno Lagoon, Naples, 40°47′34.19″N, 14°04′39.37″E) and the Adriatic (Venetian Lagoon, 45°25′50.54″N, 12°22′07.77″E), as well as a compilation of all Mediterranean *Botrylloides* COI sequences available in GenBank (**Supplementary Table S2**).

## Mitochondrial DNA Amplifications and Sequencing

Genomic DNA was extracted from a single zooid (dissected under the microscope), using the Nucleospin 96 Tissue Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's protocol and using a final elution volume of 80 μl.

A fragment of the cytochrome c oxidase subunit I (COI) gene was amplified using the LCO1490 and HCO2198 primers (Folmer et al., 1994). Amplifications were carried out in a 30 μl reaction volume with 4 μl of genomic DNA, 1X PCR buffer (Thermoprime AbGene), 2 mM of $MgCl_2$ (Thermoprime AbGene), 0.05 mM of each dNTP, 0.4 μM of each primer and 0.33 U of Taq DNA polymerase (Thermoprime AbGene). PCR was performed following the protocol of Lejeusne et al. (2011): initial denaturing step at 94°C for 3 min, followed by 5 amplification cycles (94°C for 50 s, 45°C for 50 s, 72°C for 60 s), 30 further cycles (94°C for 50 s, 50°C for 50 s, 72°C for 60 s), and a final

**TABLE 1 |** Population sampling, proportion of mis-identification (based on COI sequencing) and haplotype counts.

| Population | Collection date | *N* | % mis-identification | *B. violaceus* haplotypes | *B. diegensis* haplotypes |
|---|---|---|---|---|---|
| (1) Quiberon | 27/09/2011 | 22 | 73 | Bv-H1 (6) | Bd-H1 (14), Bd-H3 (2) |
| (2) Lorient | 26/05/2011 | 26 | 8 | Bv-H1 (17), Bv-H2 (2), Bv-H4 (5) | Bd-H1 (2) |
| (3) Concarneau | 12/05/2011 | 27 | 78 | Bv-H1 (6) | Bd-H1 (19), Bd-H3 (2) |
| (4) Camaret | 11/08/2011 | 34 | 100 | Bd-H1 (34) | none |
| (5) Brest - Moulin Blanc | 18/04/2011 | 32 | 9 | Bv-H1 (29) | Bd-H1 (3) |
| (6) Aber Wrac'h | 28/04/2011 | 38 | 24 | Bv-H1 (13), Bv-H4 (16) | Bd-H1 (8), Bd-H3 (1) |
| (7) Bloscon[1] | 09/08/2011 | 4 | 100 | none | Bd-H6 (4) |
| (8) Trebeurden | 30/06/2011 | 31 | 100 | none | Bd-H1 (22), Bd-H3 (9) |
| (9) Perros-Guirec | 13/05/2011 | 12 | 100 | none | Bd-H1 (11), Bd-H6 (1) |
| (10) St-Quay Portrieux | 25/05–7/07/2011 | 62 | 98 | Bv-H3 (1) | Bd-H1 (56), Bd-H3 (4), Bd-H6 (1) |
| (11) St-Malo | 01/06–7/09/2011 | 104 | 100 | Bv-H3 (1) | Bd-H1 (86), Bd-H3 (16), Bd-H5 (2) |
| (12) Queen Anne's Battery | 07/04/2011 | 27 | 0 | Bv-H1 (1), Bv-H4 (17), Bv-H5 (9) | none |
| (13) Weymouth | 30/06/2011 | 27 | 0 | Bv-H1 (13), Bv-H4 (9), Bv-H5 (5) | none |
| (14) Poole Town Quay | 21/06/2011 | 11 | 55 | Bv-H1 (1), Bv-H4 (4) | Bd-H1 (5), Bd-H4 (1) |
| (15) Lymington | 20/06/2011 | 26 | 0 | Bv-H1 (4), Bv-H4 (22) | none |
| (16) Southampton Ocean Village | 19/06/2011 | 20 | 0 | Bv-H1 (15), Bv-H4 (4), Bv-H5 (1) | none |
| (17) Hamble | 19/06/2011 | 30 | 0 | Bv-H1 (15), Bv-H4 (22) | none |
| (18) Gosport Premier | 18/06/2011 | 31 | 87 | BvX-H6 (4) | Bd-H1 (27) |
| (19) Southsea | 17/06/2011 | 32 | 44 | Bv-H1 (9), Bv-H4 (8), BvX-H6 (1) | Bd-H1 (14) |
| (20) Brighton | 16/09/2011 | 31 | 94 | Bv-H1 (1), BvX-H6 (1) | Bd-H1 (27), Bd-H2 (2) |

*Population numbers refer to those indicated in* ***Figure 1****. Accession numbers for the COI haplotypes are provided in* ***Supplementary Table S3*** *as* ***Supplementary Material****. [1] The Bloscon marina was under construction at the collection date (i.e., no floating pontoons; colonies were sampled on pilings).*

elongation step at 72°C for 5 min. Sequencing reactions were performed at the LGC Genomics platform (Berlin, Germany) on purified (ExoSAP®-it) PCR products using the reverse primer (HCO2198). Mediterranean amplicons were obtained with the same primers and protocols, and sequenced at Macrogen Company (Korea).

## Mitochondrial DNA Analyses

The sequences were edited using CodonCode Aligner v. 3.7.1 (CodonCode Corporation, Dedham, MA). They were aligned using BioEdit v.7.1.981 33 (Hall, 1999). After alignment, a 455 base-pair fragment was retained, further reduced to 435 base pairs for the EC dataset and to 395 base pairs for the Mediterranean dataset, for comparison with sequences obtained for our reference samples or available in public databases. DnaSP version 5 (Rozas et al., 2003) was used to compute the number and frequency of haplotypes (i.e., unique sequences) per taxon or locality and estimate polymorphism. The number of base substitutions per site across sequences was computed with MEGA7 (Tamura et al., 2011), using the Maximum Composite Likelihood model (Tamura et al., 2004) with positions with less than 95% site coverage eliminated.

Phylogenetic trees were built with MEGA7 (Tamura et al., 2011) using 32 unique sequences for the analyses of the EC dataset and 54 sequences with the Mediterranean dataset (details in Results). Goodness of fit with various evolution models was first tested: based on both BIC criteria and AICc value, the best model explaining the data was the Hasegawa-Kishino-Yano model with 54–61% of the sites evolutionarily invariable, according to the trees. Rooted phylogenetic trees were constructed using a maximum likelihood method with heuristic search (Subtree-Pruning-Regrafting method). To assess the reliability of the inferred trees, bootstrap tests were carried out (1000 bootstraps). Note that similar topologies were obtained using other tree construction methods, in particular a simple distance-based method (i.e., Neighbor-joining).

To start investigating the status of a new reported lineage (see results), GMYC (Pons et al., 2006) and mPTP (Kapli et al., 2017) species delimitation analyses were carried out on the EC and Mediterranean unique sequences of *Botrylloides* spp. These two approaches differ in their properties (e.g., GYMC requires ultrametric trees), and one may be more confident in the inferred species delimitation when similar results are obtained with both. GMYC analysis was carried out with the R library *splits*, using an ultrametric tree built using the Beast2 software (Bouckaert et al., 2019), with a Yule tree prior and site model as obtained with MEGA analysis (see above). The mPTP analysis was carried out with the web-service available at http://mptp.h-its.org, with the tree produced with MEGA (note that the same results were obtained with the tree produced by the Beast2 Bayesian phylogenetic inference).

# RESULTS

## Haplotypic Diversity and Comparison to Reference Samples

Altogether, over the 627 colonies sampled in 20 marinas in 2011, we obtained 12 mitochondrial haplotypes (**Table 1**) over 455 base pairs, with no missing data. The polymorphism

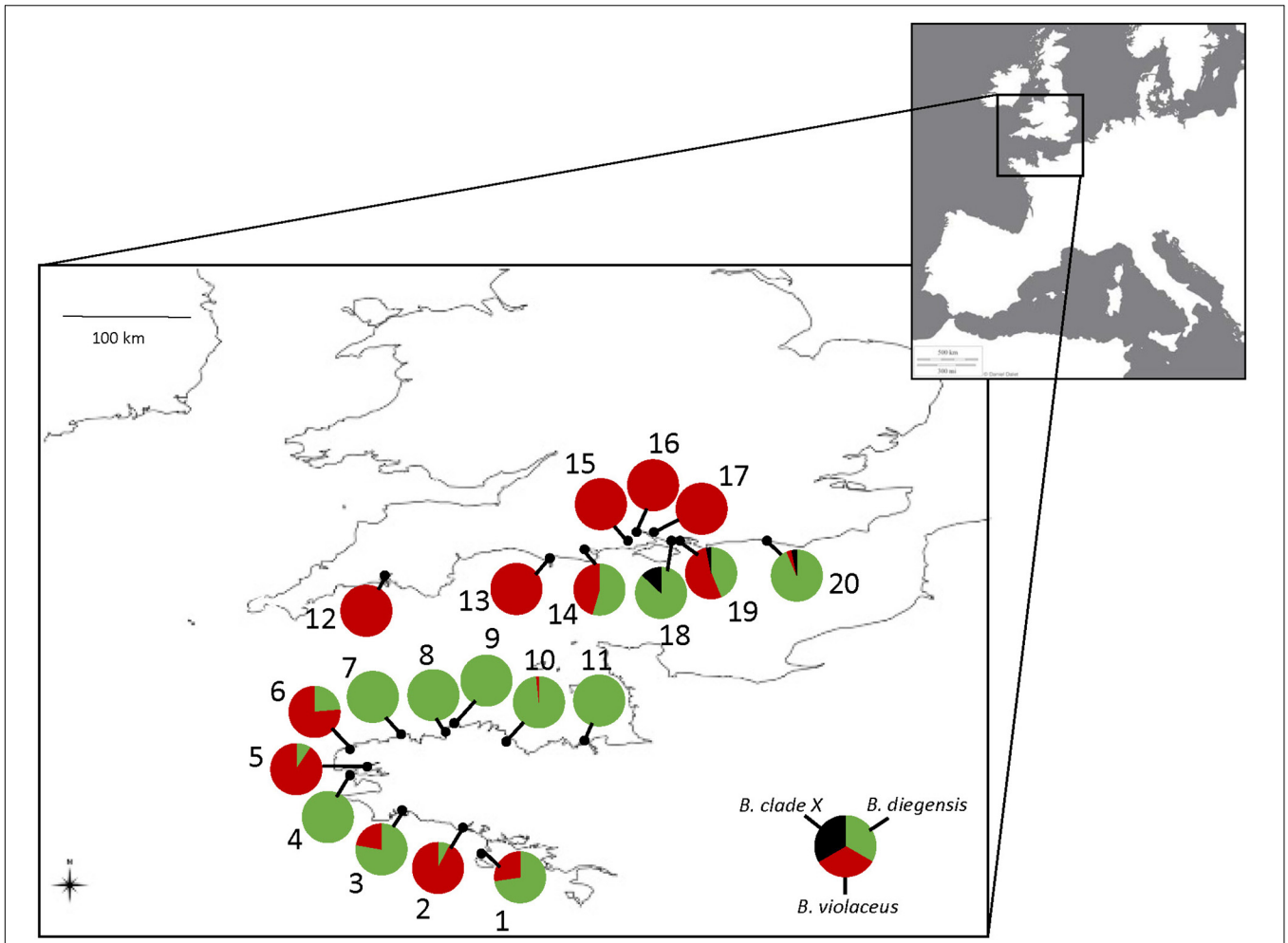


**FIGURE 1 |** Study area and sampling localities of the 627 single-color colonies sampled in the English Channel and morphologically assigned to *B. violaceus*, with pie charts indicating the percentage of the three *Botrylloides* taxa, after being identified with molecular barcoding (see **Table 1**). Numbers refer to population name as indicated in **Table 1**. The source for the European map in the inset is ©Terre Ouverte.

was very high with 122 sites showing substitutions, for a total number of 140 mutations. In addition, although these colonies were all assigned to *B. violaceus* based on their color, evolutionary divergence over all sequence pairs was substantial, with a number of base substitutions per site of 0.175 over the 12 sequence pairs. Marked variations were observed among pairwise sequence divergence estimates, ranging from 0.002 to 0.301. On the other hand, the analysis of the 127 reference samples including the three *Botrylloides* spp. putatively present and *Botryllus schlosseri* (outgroup) yielded 20 mitochondrial haplotypes (**Supplementary Table S1**).

A phylogenetic tree was built to map the 12 haplotypes obtained over the 627 colonies with the 20 sequences obtained for the 127 reference samples (**Figure 2**). This tree displayed three monophyletic groups associated with 1) *B. diegensis*, 2) *B. leachii* ('radiata' morph) and 3) *B. violaceus* reference sequences. It showed that 11 haplotypes obtained from the colonies sampled actually belonged to two groups, namely those clustering with *B. violaceus* and *B. diegensis* references. As these colonies were all originally thought to be *B. violaceus*, based on the single-color pattern, this result thus clearly indicates mis-identification during collection in the field. No sequence clustered with our *B. leachii* references. Among the 11 haplotypes, seven were 100% identical to one of the 20 reference sequences: four (out of six assigned to the *B. diegensis* clade) were identical to *B. diegensis* references (Bd-H1.Ref, Bd-H2.Ref, Bd-H3.Ref, and Bd-H6.Ref in **Figure 2**) and three (out of five) to *B. violaceus* references (Bv-H1.Ref, Bv-H3.Ref, Bv-H4.Ref).

One haplotype ('BvX-H6.EC sampling' in **Figure 2**) clustered with the five *B. violaceus* haplotypes but the topology was poorly supported by bootstrap values. In addition, the evolutionary divergence of this haplotype (hereafter named BvX-H6) from the other haplotypes of *B. violaceus* was particularly high (19.6%) (**Table 2**). This strong divergence is clearly reflected in the mismatch distribution shown in **Supplementary Figure S1** provided as **Supplementary Material**. When removing this haplotype from the analyses, the three targeted species (i.e., *B. diegensis, B. violaceus* and *B. leachii*) were characterized by a

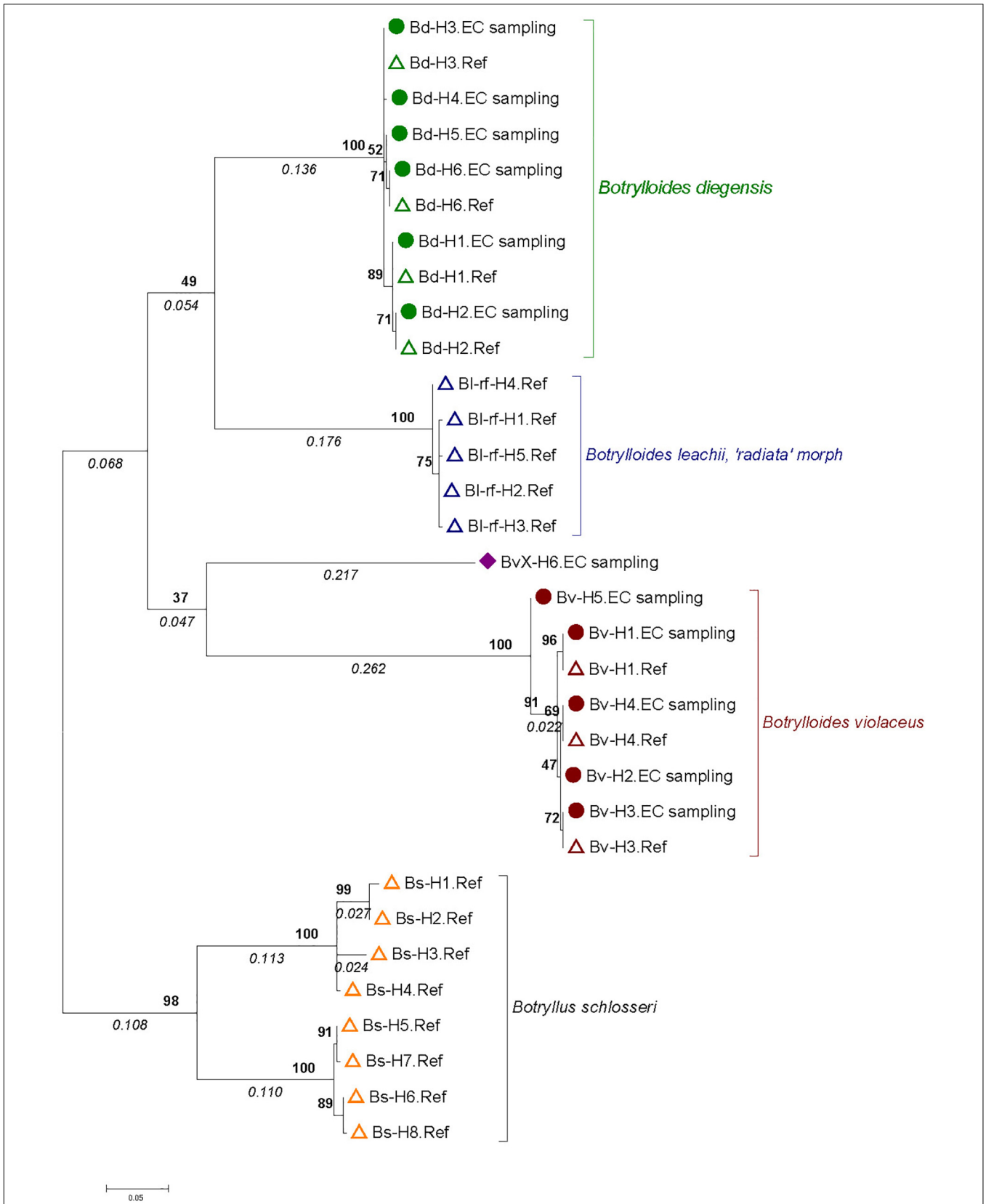


**FIGURE 2 |** Continued

**FIGURE 2 |** Molecular phylogenetic tree based on 30 *Botrylloides* spp. and *Botryllus schlosseri* haplotypes. The tree was built by the Maximum Likelihood Method based on the Hasegawa-Kishino-Yano model with 61% of sites evolutionarily invariable. It is made of 12 haplotypes obtained from 627 single-color *Botrylloides* spp. colonies sampled in the English Channel (solid circles and diamond, name ending with "EC sampling") and 20 haplotypes from reference samples (open triangles; name ending with 'Ref') from *Botrylloides* spp. and *Botryllus schlosseri* colonies (see text and **Table 1** and **Supplementary Table S1** for details). A total of 435 positions were analyzed. The tree is drawn to scale, with branch lengths measured as number of substitutions per site indicated in italics below the branch. Numbers in bold on nodes indicate percent bootstrap support values (1000 bootstraps).

**TABLE 2 |** Estimates of average molecular divergence over haplotype pairs within and between taxonomic groups (i.e., the three accepted species, examining separately the haplotype BvX-H6).

| | ***B. violaceus*, excluding BvX-H6** | **BvX-H6 alone** | ***B. diegensis*** | ***B. leachii*** |
|---|---|---|---|---|
| *B. violaceus*, excluding BvX-H6 | 0.0129/5.6 | 85.200 | 82.733 | 81.700 |
| BvX-H6 alone | 0.196 | N/A | 73.833 | 78.500 |
| *B. diegensis* | 0.191 | 0.170 | 0.007/2.87 | 67.167 |
| *B. leachii* | 0.188 | 0.181 | 0.155 | 0.005/2 |

*Values on the diagonal are comparing haplotype within group, with distance (computed over 434 base pairs) and absolute number separated by a slash. For comparison between groups, distance and absolute number are indicated below and above the diagonal, respectively.*

low level of within-species divergence (0.5 to 1.29%) and much higher level (15.5–19.6%) of divergence among them (**Table 2**), highlighting that the COI marker used is a robust barcoding marker to distinguish them, in particular *B. diegensis* from *B. violaceus*, and that BvX-H6 corresponds to an evolutionarily divergent lineage as compared to the three targeted species. Colonies with the BvX-H6 haplotype displayed a single-color pattern (**Supplementary Figure S3**, panels 7 and 8) broadly similar to *B. violaceus*.

## Mis-Identification Rates and Haplotype Distribution Across Populations

Based on the results of the phylogenetic analysis, each of the 627 colonies could be assigned to either *B. diegensis, B. violaceus* or the divergent BvX-H6 lineage. Altogether 373 (59%) of the 627 single-color colonies sampled in the United Kingdom and France as *B. violaceus* in 2011 were unambiguously assigned to *B. diegensis* on molecular criteria, as shown above. We observed large variations across regions and populations (**Figure 1**), with mis-identification ranging from 0 to 100%, and a mean mis-identification rate of 53% per population. The mis-identification was not associated with a particular haplotype, with six haplotypes found among these *B. diegensis* single-color colonies. The number of reference specimens is not large enough to encompass the whole diversity of *B. diegensis* but it is noteworthy that haplotypes Bd-H1, Bd-H2, Bd-H3 and Bd-H6 characterizing our *B. diegensis* two-color reference samples (**Supplementary Table S1**), were also found in single-color colonies of *B. diegensis* mis-identified as *B. violaceus* (**Table 1**).

## Comparison With Sequences in Public Databases

We completed our analysis with a species identification search on the BOLD portal, using as a query each of the distinct haplotypes obtained in this study. The five *B. violaceus* haplotypes were found to match with reference sequences in the BOLD system (**Supplementary Table S3**): they were assigned with high similarity (97.59–97.62%) to '*Botrylloides violaceum*' (note that this neuter form of the specific name is no longer accepted according to WoRMS, being replaced by the masculine form used here; see Ryland, 2015). These haplotypes (Bv-H1 to Bv-H5) matched with four haplotypes (named as Bv2, Bv8, Bv9, and Bv11) obtained as part of two population genetics studies of *B. violaceus* in N. America (Bock et al., 2011; Lejeusne et al., 2011).

The haplotype BvX-H6 did not match any sequences in BOLD, and the best match obtained using a BLAST search on the NCBI portal was with *Botrylloides israeliense* (Reem et al., 2018) but with only 93.27% similarity (for a 91% query cover over 461 bp), which indicates that the assignment is not robust.

None of the *B. diegensis* haplotypes matched with any sequences registered under this name in BOLD (**Supplementary Table S3**): they were all assigned at 99.76–100% similarity to references registered under the name *B. leachii* (or the incorrect variant spelling *B. leachi*) with different published (Griggio et al., 2014; López-Legentil et al., 2015) or unpublished sources. It is noteworthy that one of these sources (Griggio et al., 2014) is a full mitochondrial genome sequence.

The five haplotypes of the 'radiata' morph of *B. leachii* did not match with any references registered in BOLD (**Supplementary Table S4**). When searching in the NCBI portal, the best match was with a sequence (acc. no. KY235402) provided under the name *B. leachii*, but with an extremely low level of similarity (85.15–85.68%), thus well below the threshold used for taxonomic assignment. Interestingly this sequence matches, with high similarities (99.76–100%), our *B. diegensis* haplotypes in BOLD (not shown). Thus, we can conclude that there are no reference sequences corresponding to our specimens of the 'radiata' morph of *B. leachii* either in BOLD or NCBI.

## Comparison With Mediterranean Sequences

A total of 22 sequences were retrieved from GenBank of Mediterranean *Botrylloides* (six of them from whole mitochondrial mtDNA data), and 17 new sequences were obtained from the samples collected in Spain, France and

Italy (**Supplementary Table S2**). These samples were mapped onto a phylogenetic tree with the 25 haplotypes obtained with the EC dataset (EC sampling and reference sequences; **Supplementary Tables S3**, **S4**).

The results (**Figure 3**) showed that all sequences in GenBank previously assigned to *B. leachii* clustered unambiguously with *B. diegensis*. The only previous *B. violaceus* sequence (i.e., complete mitochondrial genome HF548552; Griggio et al., 2014) grouped with our reference sequences of *B. violaceus*. *B. pizoni* branched within the [*B. diegensis* – *B. leachii*] clade, with a higher similarity to our *B. leachii* references, whereas *B. niger* branched at the root of this clade. In both cases, very low bootstrap values indicated that the topology was poorly supported. The tree, however, indicates that they are distinct from the *B. violaceus* clade.

As for the newly generated sequences, two colonies from Venice (with a clear two-color pattern), all colonies from Sète (with a clear two-color pattern), and one from the Ebro Delta (with a uniform reddish color) grouped with *B. diegensis*. Another three colonies from Venice (of uniform orange or red coloration) were in the *B. violaceus* clade. The *Botrylloides* colony from Miseno Lagoon (NAP1 in **Figure 3**) clustered with *B. israeliense* and somewhat less closely with the EC haplotype BvX-H6.

Using GMYC, nine entities were identified among the *Botrylloides* sequences (**Supplementary Table S5**), with BvX-H6 assigned to a different entity ('species') than *B. israeliense* plus the sample collected in NAP – both assigned to the same species with a maximum likelihood support of 1. Interestingly the haplotype Bv-H5 was also distinguished from the other *B. violaceus* haplotypes suggesting a putative additional cryptic species. However, the mPTP analysis distinguished only seven entities (**Supplementary Table S5**), with BvX-H6, NAP and *B. israeliense* clustered into a single entity, and Bv-H5 clustered with the other haplotypes of *B. violaceus*. The other delimited entities were the same in the two analyses (**Supplementary Table S5**).

## DISCUSSION

We used COI sequencing to examine >750 colonies of the genus *Botrylloides* and *Botryllus* with a special interest in three species, two introduced species, namely *B. violaceus* and *B. diegensis*, and one species putatively native to Europe, namely *B. leachii*. Combining phylogenetic analyses and barcoding, we showed that COI is a robust barcode for distinguishing the three species. We also confirmed the presence of the two NIS in the EC. In addition, this study went well beyond our expectations: it pointed out (1) the risk of mis-identification between the two NIS when rapidly identified in the field based on simple external criteria (colony color patterns), (2) the presence of errors in databases, with reference data available for *B. leachii* (native) at the time of the analysis (last checked in May 2019) actually being sequences corresponding to *B. diegensis* (introduced in the study regions), (3) the presence of *B. diegensis*, probably previously mistaken as *B. leachii*, in the Mediterranean Sea, based on our custom reference database, thus pointing to the need to revise all previous reports of *B. leachii*, as it has likely been confused with *B. diegensis* and (4) the existence of a cryptic divergent lineage, displaying a single-color pattern similar to *B. violaceus*.

Thus the molecular findings indicate that single-color colonies lacking larvae cannot be identified to species level with confidence from external appearance alone (see brief discussion in Bishop et al., 2015a), at least until additional distinguishing characters are recognized. Preliminary morphological observations based on external features of the species investigated are presented in **Supplementary Figure S3**. Identification by internal morphology, i.e., microscopical examination of the zooids and brooded larvae, requires laboratory facilities and training, and would be very time-consuming if large numbers of specimens were to be processed, for instance to estimate the relative abundance of the different species in multiple samples. In addition, some key morphological features, such as gonads and larvae, may be present only seasonally. Recourse to molecular means of identification appears a fruitful option, particularly for colorless preserved specimens that may not be reproductive. Furthermore, clear distinctions based on internal anatomy are not yet very consistently reported in the literature, given the current level of taxonomic confusion. Morphological characterization of specimens that have been pre-sorted into putative taxonomic categories based on DNA sequence information may help to achieve the full clarification of any differences in internal anatomy. However, discovery of reliable differences in zooidal morphology can prove challenging in this genus. Saito et al. (1981) distinguished *B. simodensis* Saito and Watanabe, 1981 from *B. violaceus* on the basis of reproductive characteristics and the color of live colonies, while preserved, non-reproductive colonies were very hard to separate on the basis of zooidal morphology. Atsumi and Saito (2011) were subsequently able to distinguish two additional species from *B. simodensis* based on differences in reproductive seasonality and the color of living colonies, backed up by mtDNA sequence information, despite them sharing very similar zooidal features.

As shown by the phylogenetic tree (**Figure 2**), COI is reliable to distinguish the three accepted species with a clear barcoding gap among them (**Supplementary Figure S1**). This is in line with previous studies showing that COI can be a relevant barcoding marker for metazoans (Bucklin et al., 2011), notably considering its properties as a rapidly evolving marker (as compared to the nuclear genome). This study also highlights a clear benefit of the barcoding approach, i.e., to enable species identification in groups that lack easy morphological traits to use in taxonomic determination. We carried out extensive sampling of *Botrylloides* spp. colonies that externally looked like *B. violaceus*, based on one simple field criterion, namely the single-color patterns of the colonies. The results of COI sequencing and comparison with a newly generated reference dataset clearly indicated that we made a large number of errors in assigning species name in the genus *Botrylloides*. Error rates varied 0 to 100% (**Figure 1**) according to the population. We could compare the results of the molecular identifications of the supposedly '*B. violaceus*' population samples with results of RAS carried by one of us (JB) while sampling for the intended molecular study in the nine populations from the United Kingdom. During the

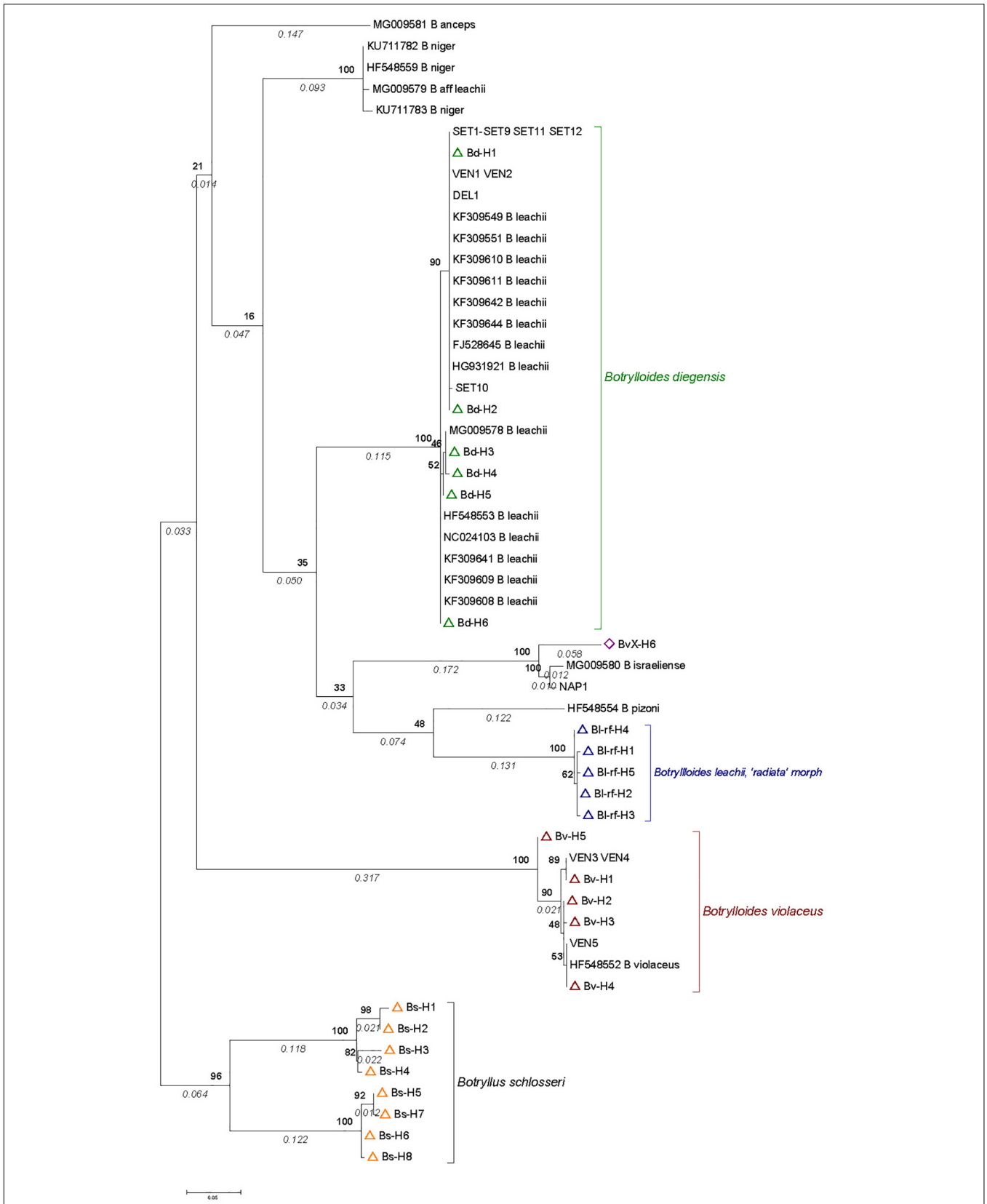


**FIGURE 3 |** Continued

**FIGURE 3 |** COI phylogenetic tree based on *Botrylloides* sequences obtained from samples in the English Channel, reference sequences (see **Figure 2**), and sequences obtained from samples from the Mediterranean Sea (published data or from this study; see details in **Supplementary Table S2**). The tree was built with MEGA 7 by the Maximum Likelihood Method based on the Hasegawa-Kishino-Yano model with 54% of sites evolutionarily invariable. It is made of 52 sequences over 395 base pairs. They include sequences obtained from Mediterranean *Botrylloides* either newly collected or obtained from GenBank (see **Supplementary Table S2** for codes and details). In addition, we included 6, 5, and 5 haplotypes from our reference sequences for *B. diegensis*, *B. violaceus* and *B. leachii* 'radiata' morph (open triangle symbols in green, red and blue, respectively), together with the haplotype named as BvX-H6 (violet diamond) (see details in **Supplementary Tables S3**, **S4**). The tree is drawn to scale, with branch lengths measured as number of substitutions per site indicated in italics below the branch. Numbers in bold on nodes indicate percent bootstrap support values (1000 bootstraps).

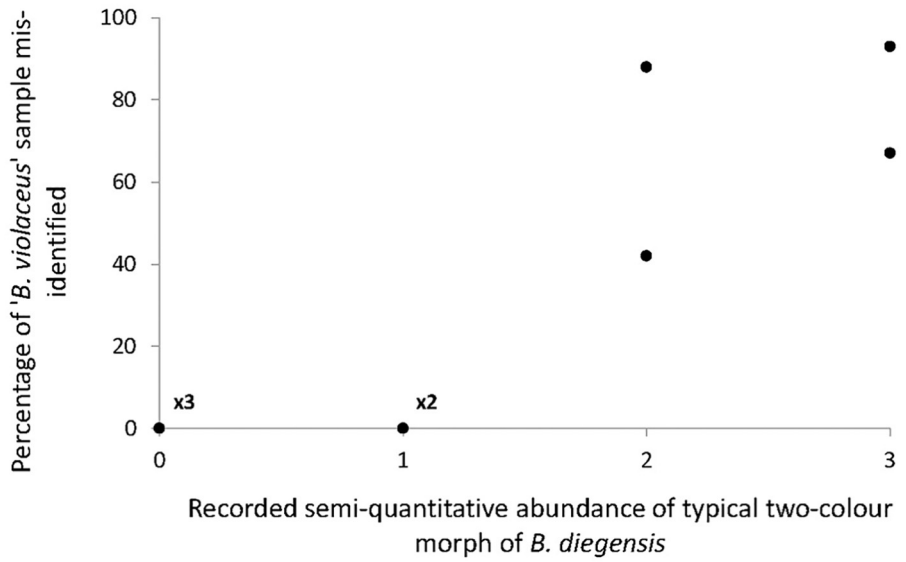


**FIGURE 4 |** Relationship between the percentage of *B. violaceus* mis-identification (i.e., individuals that were assigned to *B. diegensis* with barcoding) and the abundance of typical two-color morph of *B. diegensis* in the nine study populations on the English coast. The four-point abundance scale was obtained for *B. diegensis* during rapid assessment surveys carried out at the same time as sampling for molecular analyses.

RAS, a semi-quantitative abundance of the distinctive two-color morph *B. diegensis* was recorded on a four-point scale of abundance. Plotting the percentage of mis-identification as a function of the abundance of two-color *B. diegensis* shows that the highest mis-identification rates were observed where two-color *B. diegensis* colonies were relatively common (**Figure 4**). From this plot, it is clear that the typical two-color morph is easily identified as compared to *B. violaceus*. It also confirms that the distinctive '*B. diegensis*' two-color morph is associated with a conspecific uniformly colored morph. Although the number of sites where the two morphs were both identified is low, the correlation further suggests that these two morphs within *B. diegensis* occur in relatively constant proportions from place to place. Interestingly, the single-color morph of *B. diegensis* shares haplotypes with the two-color morph (**Supplementary Table S1** compared to **Supplementary Table S3**). In addition, colonies of the same color can be found in both single-color *B. diegensis* and in *B. violaceus*.

Even before becoming the barcoding marker of choice for many groups, COI had been extensively used in phylogeographic studies, thus to investigate intra-specific diversity and evolutionary history (Avise, 2000; Beheregaray, 2008; Hickerson et al., 2010). Such population-based investigations often revealed divergent lineages, which, through an integrative taxonomy approach, were sometimes assigned to new species (Pante et al., 2015). *Botryllus* and *Botrylloides* are complex taxa in which species status has already been debated (Brunetti, 2009; Reem et al., 2018). For instance, Bock et al. (2012) showed that at least three divergent clades that might be cryptic species characterize *B. schlosseri*. Griggio et al. (2014) further indicated that the clade A (as defined in Bock et al., 2012) may in turn be undergoing speciation. Our eight reference sequences for this accepted species were obtained from specimens collected in Brittany and south-western United Kingdom. They are also indicative of the presence of two divergent clades in the W English Channel (**Figure 2**), which are corresponding to clade A and E described by Bock et al. (2012). These authors also showed they co-occur in three populations (namely Brest, Falmouth and Brixham) from EC. Similarly, we found one haplotype, which we called BvX-H6, from colonies externally somewhat similar to *B. violaceus,* that showed a very large sequence divergence from the other studied *Botrylloides* species, with a divergence similar (ca. 20%) to the one measured between reference sequences of *B. violaceus* and *B. diegensis* (**Table 2** and **Supplementary Figure S1**).

Two populations from United Kingdom, namely Southsea and Gosport, where the first dataset revealed the presence of

BvX-H6 were sampled a second time in June 2012. Out of the 13 and 35 specimens examined, 77 and 24% specimens showed this particular haplotype in Southsea and Gosport, respectively. In Gosport, the other specimens were all *B. diegensis* and in Southsea they were all *B. violaceus* (**Supplementary Figure S2**). As part of another population-genetics study, we used the microsatellite markers specifically developed to target *B. violaceus* (Molecular Ecology Resources Primer Development Constorium et al., 2010) on the colonies assigned here to *B. violaceus* (with haplotypes Bv-H1 to Bv-H5) and to this new clade. Interestingly, whereas all the colonies with COI haplotypes Bv-H1 to Bv-H5 were easily genotyped with these microsatellites, no PCR product could be obtained for any of the BvX-H6 colonies (data not shown, FV, CR, JB unpublished data). This amplification failure strongly suggests that the colonies with haplotype BvX-H6 represent a distinct cryptic species in the EC. The clustering of this form with the *Botrylloides* colony from Miseno Lagoon (Italy) and with *Botrylloides israeliense* (**Figure 3**) indicates that the same divergent clade is represented in the Mediterranean Sea, but the reported divergence between the two geographical regions (ca. 6%), based on comparison of just 395 base pairs, leaves it unclear whether the EC colonies with haplotype BvX-H6 should be regarded as a somewhat divergent lineage of *B. israeliense* or as a sister species which is possibly undescribed. Discrepancies between the outcome of the GYMC and mPTP species delimitation methods left us with uncertainties regarding the status of this new lineage. The external appearance of *B. israeliense* shown by Reem et al. (2018: Figure 1b) does not suggest close kinship with the entity represented by BvX-H6 (**Supplementary Figure S3**, panels 7 and 8), but further molecular studies, including new species delimitation analyses with other markers, ideally coupled with detailed morphological comparison, will be necessary to resolve the status of this lineage and the exact relationship between the EC and Mediterranean populations.

Another facet of this work has been the documentation of some errors in public databases. Such errors have been pointed out previously (e.g., Harris, 2003; Vilgalys, 2003, and references herein). In our case, this issue was assessed by using simple identification searches in the BOLD system and in the NCBI portal. The outcome is illustrated in the Mediterranean Sea with a phylogenetic tree in **Figure 3**, where it is clear that all our *B. diegensis* references are mixed with sequences named *B. leachii* in Genbank. Further, new sequences obtained confirmed that *B. diegensis* in the Mediterranean can also have a two-color pattern (Venice or Sète specimens) or a single reddish coloration (Ebro Delta). We failed to find any sequence, either published or new, of Mediterranean *B. leachii*. Considering that this is the *Botrylloides* species most commonly mentioned in this sea, our results suggest that the presence of *B. diegensis* might have been strongly underestimated in the Mediterranean region. We also confirm the presence of *B. violaceus* in the Venice Lagoon, with external aspect identical to single-color colonies of *B. diegensis*.

The case of *B. diegensis* and *B. violaceus* is a particularly acute issue as these are two NIS, widespread in the northern hemisphere. The lack of reliability of public databases for notorious NIS is an important shortcoming for their early detection and thus effective management (Darling et al., 2017). In this context and for the specific case examined here, it is noteworthy that this mis-identification of *B. diegensis* occurred during a survey of Mediterranean harbors with the specific aim of examining the distribution of NIS in these well-known introduction hotspots (López-Legentil et al., 2015), and during collections by JB and FV in the EC, specifically targeting *B. violaceus* in marinas for population-genetic studies. We now have evidence for the presence of *B. diegensis* in Venice (Italy), Delta of Ebro (Spain), along the Catalan coast (Spain) and in Sète (France). Nevertheless, *B. diegensis* has not been formally reported in the Mediterranean, although Brunetti and Mastrototaro (2017) mention personal communication by A. and W. Bay-Nouailhat indicating its presence in the W Mediterranean. *B. diegensis* might be not only present but common and widespread along the Mediterranean coast, although commonly mis-identified as *B. leachii*. Confusion between *B. diegensis* and *B. leachii* might have contributed to statements that these species share very similar morphology (e.g., Van Name, 1945), and Brunetti (2009) synonymized them, although Brunetti and Mastrototaro (2017) treat them as separate. There is little doubt that the distinctive two-color morph now attributed to *B. diegensis* is a relatively recent arrival in NW Europe, rather than a long-established, potentially native, species such as *B. leachii*. The very striking two-color morph is not illustrated or mentioned in accounts of *Botrylloides* in less recent identification guides to the region (e.g., Eales, 1939; Berrill, 1950; Barrett and Yonge, 1958; Millar, 1970; Picton, 1985; Hayward and Ryland, 1990; Gibson et al., 2001). Furthermore, the continuing spread of the distinctive two-color morph to new regions of the English coast from an apparently recent origin on the south coast has been documented in repeated field surveys since 2004 (e.g., Wood et al., 2016). RAS rely substantially on clearly defined field characteristics, which may be elusive in some groups; detailed molecular and morphological studies are called for before field identification can be deemed reliable for the important ascidian genus *Botrylloides*.

## DATA AVAILABILITY STATEMENT

The haplotypes generated by Sanger sequencing have been deposited in GenBank (accession numbers MK978800–MK978824 for the EC sampling and reference sequences and MN076465–MN076483 for the Mediterranean sampling).

## AUTHOR CONTRIBUTIONS

FV and JB contributed to the conception and design of the study. FV, JB, SB, and XT provided the samples and/or carried out the sampling. CR, SB, and XT performed the molecular work or provided the sequence data. FV and CR made the analyses. FV wrote the first draft of the manuscript with substantial contributions of JB for some sections and XT for the Mediterranean part of the manuscript. All authors contributed to the manuscript revision, and read and approved the submitted version.

## FUNDING

This work was part of the Interreg IVA project 'Marinexus' financed under the European Regional Development Fund. And also received further support from the AquaNIS2.0 project funded by the Fondation TOTAL, and from project CTM2017-88080 (MCIU/AEI/FEDER/UE) of the Spanish Government.

## ACKNOWLEDGMENTS

We are grateful to Anna Yunnie, Christine Wood, Laurent Lévêque, and the divers of the Marine Operations department at the Roscoff Biological Station for their help in sampling and preliminary sorting of the samples in EC. We also thank the marina operators for their permission to conduct the experiment. We are grateful to the following people for helping with the Mediterranean samples or sending specimens to be used as reference: Maria Casso, Franck Gentil, Gretchen Lambert, Adam Langenbacher, Julia Nunn, Mari Carmen Pineda Mark Rius, Kirsty Smith Paolo Sordino, and Davide Tagliapietra. Susanna López-Legentil sequenced the specimen from Miseno Lagoon.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00615/full#supplementary-material

## REFERENCES

Alder, J., and Hancock, A. (1912). The British tunicata: an unfinished monograph. Vol. 3. Aggregata (Ascidiae Compositae). *Ray Soc. Publications* 93, 1–113.

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., et al. (2012). The magnitude of global marine species diversity. *Curr. Biol.* 22, 2189–2202. doi: 10.1016/j.cub.2012.09.036

Arenas, F., Bishop, J. D. D., Carlton, J. T., Dyrynda, P. J., Farnham, W. F., Gonzalez, D. J., et al. (2006). Alien species and other notable records from a rapid assessment survey of marinas on the south coast of England. *J. Mar. Biol. Assoc. U. K.* 86, 1329–1337. doi: 10.1017/s0025315406014354

Atsumi, M. O., and Saito, Y. (2011). Studies on Japanese *botryllid* ascidians. V. A new species of the genus *Botrylloides* very similar to *Botrylloides simodensis* in morphology. *Zool. Sci.* 28, 532–542. doi: 10.2108/zsj.28.532

Avise, J. C. (2000). *Phylogeography: the History and Formation of Species*. Cambridge, MA: Harvard University Press.

Barrett, J. H., and Yonge, C. M. (1958). *Collins Pocket Guide to the Sea Shore*. London: Collins.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Mol. Ecol.* 17, 3754–3774. doi: 10.1111/j.1365-294X.2008.03857.x

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2013). GenBank. *Nucleic Acids Res.* 41, D36–D42. doi: 10.1093/nar/gks1195

Berrill, N. J. (1950). The tunicata, with an account of the British species. *Ray Soc. Publications* 133, 1–354.

Bishop, J. D. D., Roby, C., Yunnie, A. L. E., Wood, C. W., Lévêque, L., Turon, X., et al. (2013). The Southern Hemisphere ascidian *Asterocarpa humilis* is unrecognised but widely established in NW France and Great Britain. *Biol. Invasions* 15, 253–260. doi: 10.1007/s10530-012-0286-x

Bishop, J. D. D., Wood, C. A., Lévêque, L., Yunnie, A. L. E., and Viard, F. (2015a). Repeated rapid assessment surveys reveal contrasting trends in occupancy of marinas by non-indigenous species on opposite sides of the western English Channel. *Mar. Pollut. Bull.* 95, 699–706. doi: 10.1016/j.marpolbul.2014.11.043

Bishop, J. D. D., Wood, C. A., Yunnie, A. L., and Griffiths, C. A. (2015b). Unheralded arrivals: non-native sessile invertebrates in marinas on the English coast. *Aquat. Invasions* 10, 249–264. doi: 10.3391/ai.2015.10.3.01

Bock, D. G., MacIsaac, H. J., and Cristescu, M. E. (2012). Multilocus genetic analyses differentiate between widespread and spatially restricted cryptic species in a model ascidian. *Proc. R. Soc. B: Biol. Sci.* 279, 2377–2385. doi: 10.1098/rspb.2011.2610

Bock, D. G., Zhan, A., Lejeusne, C., MacIsaac, H. J., and Cristescu, M. E. (2011). Looking at both sides of the invasion: patterns of colonization in the violet tunicate *Botrylloides violaceus*. *Mol. Ecol.* 20, 503–516. doi: 10.1111/j.1365-294X.2010.04971.x

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: An advanced software platform for bayesian evolutionary analysis. *PLoS Comput. Biol.* 15:e1006650. doi: 10.1371/journal.pcbi.1006650

Brunetti, R. (2009). Botryllid species (Tunicata, Ascidiacea) from the Mediterranean coast of Israel, with some considerations on the systematics of Botryllinae. *Zootaxa* 2289, 18–32. doi: 10.11646/zootaxa.2289.1.2

Brunetti, R., and Mastrototaro, F. (2017). *Ascidiacea of the European waters*. Milano: Edagricole.

Bucklin, A., Steinke, D., and Blanco-Bercial, L. (2011). DNA barcoding of marine metazoa. *Annu. Rev. Mar. Sci.* 3, 471–508. doi: 10.1146/annurev-marine-120308-080950

Campbell, M. L., Gould, B., and Hewitt, C. L. (2007). Survey evaluations to assess marine bioinvasions. *Mar. Pollut. Bull.* 55, 360–378. doi: 10.1016/j.marpolbul.2007.01.015

Carlton, J. T. (2009). "Deep invasion ecology and the assembly of communities in historical time," in *Biological Invasions in Marine Ecosystems*, eds G. Rilov, and J. A. Crooks, (Berlin: Springer), 13–56. doi: 10.1007/978-3-540-79236-9_2

Cohen, A. N., Harris, L. H., Bingham, B. L., Carlton, J. T., Chapman, J. W., Lambert, C. C., et al. (2005). Rapid assessment survey for exotic organisms in southern California bays and harbors, and abundance in port and non-port areas. *Biol. Invasions* 7, 995–1002. doi: 10.1007/s10530-004-3121-1

Comtet, T., Sandionigi, A., Viard, F., and Casiraghi, M. (2015). DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. *Biol. Invasions* 17, 905–922. doi: 10.1007/s10530-015-0854-y

Darling, J. A., Galil, B. S., Carvalho, G. R., Rius, M., Viard, F., and Piraino, S. (2017). Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. *Mar. Policy* 85, 54–64. doi: 10.1016/j.marpol.2017.08.014

Dias, J. P., Fotedar, S., Munoz, J., Hewitt, M. J., Lukehurst, S., Hourston, M., et al. (2017). Establishment of a taxonomic and molecular reference collection to support the identification of species regulated by the Western Australian prevention list for introduced marine pests. *Manag. Biol. Invasions* 8, 215–225. doi: 10.3391/mbi.2017.8.2.09

Eales, N. B. (1939). *The Littoral Fauna of Great Britain. A Handbook for Collectors*. Cambridge: Cambridge University Press.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294–299.

Gibson, R., Hextall, B., and Rogers, A. (2001). *Photographic Guide to the Sea and Shore Life of Britain and North-West Europe*. Oxford: Oxford University Press.

Griggio, F., Voskoboynik, A., Iannelli, F., Justy, F., Tilak, M.-K., Xavier, T., et al. (2014). Ascidian mitogenomics: comparison of evolutionary rates in closely related taxa provides evidence of ongoing speciation events. *Genome Biol. Evol.* 6, 591–605. doi: 10.1093/gbe/evu041

Hall, T. A. (1999). BioEdit: a user friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Harris, J. D. (2003). Can you bank on GenBank? *Trends Ecol. Evol.* 18, 317–319. doi: 10.1016/S0169-5347(03)00150-2

Hayward, P. J., and Ryland, J. S. (1990). *The Marine Fauna of the British Isles and North-West Europe. Volume. 2: Molluscs to Chordates*. Oxford: Oxford University Press.

Hebert, P. D., Hollingsworth, P. M., and Hajibabaei, M. (2016). From writing to reading the encyclopedia of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150321. doi: 10.1098/rstb.2015.0321

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., et al. (2010). Phylogeography's past, present, and future: 10 years after. *Mol. Phylogenet. Evol.* 54, 291–301. doi: 10.1016/j.ympev.2009.09.016

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., et al. (2017). Multi-rate poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 33, 1630–1638. doi: 10.1093/bioinformatics/btx025

Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* 420, 73–90. doi: 10.1007/978-94-017-2184-4_8

Lambert, C. C., and Lambert, G. (2003). Persistence and differential distribution of nonindigenous ascidians in harbors of the Southern California Bight. *Mar. Ecol. Prog. Ser.* 259, 145–161. doi: 10.3354/meps259145

Lehtiniemi, M., Ojaveer, H., David, M., Galil, B., Gollasch, S., McKenzie, C., et al. (2015). Dose of truth—Monitoring marine non-indigenous species to serve legislative requirements. *Mar. Policy* 54, 26–35. doi: 10.1016/j.marpol.2014.12.015

Lejeusne, C., Bock, D. G., Therriault, T. W., MacIsaac, H. J., and Cristescu, M. E. (2011). Comparative phylogeography of two colonial ascidians reveals contrasting invasion histories in North America. *Biol. Invasions* 13, 635–650. doi: 10.1007/s10530-010-9854-0

López-Legentil, S., Legentil, M., Erwin, P., and Turon, X. (2015). Harbor networks as introduction gateways: contrasting distribution patterns of native and introduced ascidians. *Biol. Invasions* 17, 1623–1638. doi: 10.1007/s10530-014-0821-z

McGlashan, D. J., Ponniah, M., Cassey, P., and Viard, F. (2008). Clarifying marine invasions with molecular markers: an illustration based on mtDNA from mistaken calyptraeid gastropod identifications. *Biol. Invasions* 10, 51–57. doi: 10.1007/s10530-007-9106-0

Millar, R. H. (1970). *British Ascidians (Synopses of the British Fauna, New Series, No. 1)*. London: Academic Press.

Molecular Ecology Resources Primer Development Constorium, Abdoullaye, D., Acevedo, I., Adebayo, A. A., Behrmann-Godel, J., Benjamin, R. C., et al. (2010). Permanent genetic resources added to molecular ecology resources database 1 August 2009–30 September 2009. *Mol. Ecol. Resour.* 10, 232–236. doi: 10.1111/j.1755-0998.2009.02796.x

Nunes, A. L., Katsanevakis, S., Zenetos, A., and Cardoso, A. C. (2014). Gateways to alien invasions in the European seas. *Aquat. Invasions* 9, 133–144. doi: 10.3391/ai.2014.9.2.02

Nydam, M., and Stachowicz, J. J. (2007). Predator effects on fouling community development. *Mar. Ecol. Prog. Ser.* 337, 93–101. doi: 10.3354/meps337093

Ojaveer, H., Galil, B. S., Campbell, M. L., Carlton, J. T., Canning-Clode, J., Cook, E. J., et al. (2015). Classification of non-indigenous species based on their impacts: considerations for application in marine management. *PLoS Biol.* 13:e1002130. doi: 10.1371/journal.pbio.1002130

Ojaveer, H., Galil, B. S., Minchin, D., Olenin, S., Amorim, A., Canning-Clode, J., et al. (2014). Ten recommendations for advancing the assessment and management of non-indigenous species in marine ecosystems. *Mar. Policy* 44, 160–165. doi: 10.1016/j.marpol.2013.08.019

Ordóñez, V., Pascual, M., Fernández-Tejedor, M., and Turon, X. (2016). When invasion biology meets taxonomy: *Clavelina oblonga* (Ascidiacea) is an old invader in the Mediterranean Sea. *Biol. Invasions* 18, 1203–1215. doi: 10.1007/s10530-016-1062-0

Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., et al. (2015). Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Mol. Ecol.* 24, 525–544. doi: 10.1111/mec.13048

Picton, B. E. (1985). *Ascidians of the British Isles. A Colour Guide*. Ross-on-Wye: Marine Conservation Society.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55:595609.

Ratnasingham, S., and Hebert, P. D. N. (2007). BOLD: the barcode of life data system: barcoding. *Mol. Ecol. Notes* 7, 355–364. doi: 10.1111/j.1471-8286.2007.01678.x

Reem, E., Douek, J., and Rinkevich, B. (2018). Ambiguities in the taxonomic assignment and species delineation of botryllid ascidians from the israeli mediterranean and other coastlines. *Mitochondrial DNA Part A* 29, 1073–1080. doi: 10.1080/24701394.2017.1404047

Rocha, R. M., Salonna, M., Griggio, F., Ekins, M., Lambert, G., Mastrototaro, F., et al. (2019). The power of combined molecular and morphological analyses for the genus *Botrylloides*: identification of a potentially global invasive ascidian and description of a new species. *Syst. Biodivers.* 1–18. doi: 10.1080/14772000.2019.1649738

Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359

Ryland, J. S. (2015). Gender of the genus *Botrylloides* Milne Edwards. *Zootaxa* 3973, 398–400. doi: 10.11646/zootaxa.3973.2.13

Saito, Y., Mukai, H., and Watanabe, H. (1981). Studies on Japanese compound styelid ascidians II. A new species of the genus *Botrylloides* and redescription of *B. violaceus* Oka. *Publications Seto Mar. Biol. Lab.* 26, 357–368. doi: 10.5134/176032

Sambrook, K., Holt, R. H. F., Sharp, R., Griffith, K., Roche, R. C., Newstead, R. G., et al. (2014). Capacity, capability and cross-border challenges associated with marine eradication programmes in Europe: the attempted eradication of an invasive non-native ascidian, *Didemnum vexillum* in Wales, United Kingdom. *Mar. Policy* 48, 51–58. doi: 10.1016/j.marpol.2014.03.018

Savigny, J.-C. (1816). *Mémoires Sur Les Animaux sans Vertèbres, Pt 2, Fasc. 1*. Paris: G. Dufour.

Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Van Name, W. G. (1945). The north and south American ascidians. *Bull. Am. Museum Nat. Hist.* 84, 1–476.

Vilgalys, R. (2003). Taxonomic misidentification in public DNA databases. *New Phytol.* 160, 4–5. doi: 10.1046/j.1469-8137.2003.00894.x

Wood, C., Bishop, J. D. D., Rennocks, L., and Crundwell, R. (2016). *RAS 2015 Non-Native Species Rapid Assessment Surveys in English Marinas (E Anglia & W coast). Report to Natural England and the Bromley Trust*. London: The Bromley Trust.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 11 October 2019
doi: 10.3389/fmars.2019.00628

# Overview on Karyotype Stasis in Atlantic Grunts (Eupercaria, Haemulidae) and the Evolutionary Extensions for Other Marine Fish Groups

*Clóvis Coutinho da Motta-Neto[1]*, Marcelo de Bello Cioffi[2], Gideão Wagner Werneck Félix da Costa[1], Karlla Danielle Jorge Amorim[1], Luiz Antonio Carlos Bertollo[2], Roberto Ferreira Artoni[3] and Wagner Franco Molina[1]*

[1] Departamento de Biologia Celular e Genética, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [2] Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Paulo, Brazil, [3] Departamento de Biologia Molecular e Estrutural, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil

OPEN ACCESS

**Edited by:**
*Dongdong Xu, Marine Fisheries Research Institute of Zhejiang, China*

**Reviewed by:**
*Jose Julian Tavera, Universidad del Valle, Colombia*
*Xiaoyan Wang, Zhejiang Ocean University, China*

***Correspondence:**
*Clóvis Coutinho da Motta-Neto*
*mottaneto.cc@gmail.com*

**Specialty section:**
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

**Received:** *25 June 2019*
**Accepted:** *24 September 2019*
**Published:** *11 October 2019*

**Citation:**
*Motta-Neto CCd, Cioffi MdB, Costa GWWFd, Amorim KDJ, Bertollo LAC, Artoni RF and Molina WF (2019) Overview on Karyotype Stasis in Atlantic Grunts (Eupercaria, Haemulidae) and the Evolutionary Extensions for Other Marine Fish Groups. Front. Mar. Sci. 6:628. doi: 10.3389/fmars.2019.00628*

Karyotype stasis, characterized by the absence of changes in chromosome number, genomic structure, and ploidy, is a recurrent condition in many biological groups, such as plants, amphibians, birds, and fishes. In fishes, the Percomorpha clade (>17,000 species), especially, its largest series, Eupercaria (>6,000 species and 161 families), shows an extensively shared karyotype with $2n$ = 48 acrocentrics; the phylogenetic extension in this series is still unknown. Haemulidae (grunts) family, which is one of the groups in the Eupercaria series, are effectively distributed across wide geographic areas; these fishes have a huge ecological relevance in reef environments and have been shown to be useful in clarifying the karyotype stasis in marine fish. In the western Atlantic Ocean, the Amazonas/Orinoco biogeographic barrier, which bounds the zoogeographic marine Provinces of the Caribbean and Brazil, has a variable effect on the gene flow in diverse fish groups. Geographically, this barrier is located inside the distribution area of several fish species of Atlantic, allowing the examination of its role on the eventual karyotype differentiations in the populations of grunts. In this sense, cytogenomic analyses were carried out in eight haemulid species by conventional cytogenetic approaches (Giemsa staining, C-banding, and Ag-NORs technique), staining with base-specific fluorochromes, and two-color fluorescent *in situ* hybridization (FISH) with 5S and 18S rDNA probes. Additionally, we performed an exhaustive survey on the cytogenetic data of species from the Eupercaria series to assess the general context of karyotype divergences. Haemulid species and populations showed a pronounced sharing of karyotypes with $2n$ = 48a, simple NORs, and reduced centromeric heterochromatin, apart from recurrent patterns of ribosomal 5S and 18S rDNA sites. Karyotype similarities persisted at intra- and congeneric levels in the Haemulidae family, indicating accentuated syntenic conservatism through large divergence periods. Stable karyotype patterns

were extensively present in the majority of members from Eupercaria (88% of the clades in this series). These results indicate that karyotype stasis, which is found in a large spectrum of marine fishes, is a multifactorial process in terms of phylogenetic, biological, and biogeographic contexts.

**Keywords: grunts, karyotype evolution, chromosomal divergence, syntenic regions, rDNA**

## INTRODUCTION

Karyotype stasis refers to the absence of changes in the chromosome number and structure, ploidy levels, as well as the genomic composition during phylogenetic divergence (Kahl, 2015). This process has been described in groups as diverse as plants (Mandáková et al., 2010; Bomfleur et al., 2014; Samad et al., 2016), amphibians (Sessions and Kezer, 1991; Aprea et al., 2004), birds (Ellegren, 2010), and fish (Molina, 2007; Motta-Neto et al., 2011a,b; Molina et al., 2014).

Extensive cytogenetic analyses of at least one species among about 45% of families belonging to Percomorpha (Arai, 2011), a clade with >17,000 species (Hughes et al., 2018), have highlighted notable karyotype similarities among lineages showing very ancient stages of divergence (Molina et al., 2013; Costa et al., 2016).

Meta-analysis approaches have improved the understanding of karyotype stasis in marine fish (Molina, 2007), evidencing cases of slow rate of karyotype evolution (Molina et al., 2014). Previously published data have revealed an astonishing panel of clades sharing karyotypes with $2n = 48$ acrocentric chromosomes, in addition to other traits such as reduced amount of heterochromatin, and occurrence of a single pair of chromosome-bearing nucleolar organizer regions (Ag-NORs/rDNA 18S sites) (Galetti et al., 2000; Molina, 2007).

The family Haemulidae, a typical group of the Percomorpha clade, comprises 136 species of fishes that are commonly known as grunts (Fricke et al., 2019) and distributed circumglobally in tropical and subtropical oceanic waters. Grunts are small-to-medium-sized carnivorous fishes that predominantly inhabit rocky substrates at a low depth (Nelson et al., 2016). This group has conservative karyotypes (Motta-Neto et al., 2012), a well-known phylogenetic relationship (Rocha et al., 2008; Tavera et al., 2012), and several examples of sibling-species divergence in the New World (Tavera et al., 2018; Bernal et al., 2019; Tavera and Wainwright, 2019), making it a good model for the evaluation of slow chromosome changes in a space-temporal context.

Remarkable phylogenetic reorganizations have occurred in the Percomorpha clade (Near et al., 2012; Betancur-R et al., 2013, 2017; Hughes et al., 2018). Among the nine series proposed in this group, Eupercaria is the largest, with more than 6000 species arranged in 161 families and at least 17 orders (Hughes et al., 2018). Eupercaria encompasses some of the most diverse orders (Betancur-R et al., 2017), which have received considerable attention with regards to their cytogenetic data.

To estimate karyotype stasis at different scales of diversity, we compiled the karyotype data for fishes from Eupercaria, the largest series in the Percomorpha (=Percomorphaceae) clade, and performed cytogenetic analyses of eight Haemulidae species and Atlantic populations by means of conventional techniques, staining with base-specific fluorochromes, and fluorescent *in situ* hybridization (FISH) with 18S and 5S rDNA sequences.

## MATERIALS AND METHODS

### Specimens and Chromosome Preparations

The eight species of the family Haemulidae used in the cytogenetic analysis (**Figure 1**) were collected using fishing nets from the Rio Grande do Norte state coast in northeastern Brazil, the Fernando de Noronha Archipelago in the western Atlantic region, and the Caribbean, along the coast of Key Largo, Florida (United States). Information regarding the species and samples is shown in **Table 1**.

The samples were collected with permission from the Chico Mendes Institute for Biodiversity Conservation (ICMBio/SISBio) (License# 02001.001902/06-82). All the experimental procedures were approved by the Committee of Ethics for the use of Animals of the Federal University of Rio Grande do Norte (# 044/2015).

The individuals were subjected to mitotic stimulation for 18–24 h (Molina et al., 2010). The chromosome preparations were obtained from the cells of the cephalic kidney, in accordance with the method described by Gold et al. (1990). The nucleolar organizer regions (NORs) and heterochromatin were analyzed by the Ag-NOR technique (Howell and Black, 1980) and C-banding (Sumner, 1972), respectively. Additionally, the chromosome spreads were staining with $CMA_3$/DAPI base-specific fluorochromes (Schweizer, 1980).

The 5S and 18S rDNA probes were isolated from the genome of *Rachycentron canadum* (Euteleostei, Perciformes) and amplified using primers A 5′-TAC GCC CGA TCT CGT CG ATC-3′ and B 5′-CGA GCT GGT ATG GCC GTA AGC-3′ (Pendás et al., 1994) and NS1 5′-GTA GTC ATA TGC TTG TCTC-3′ and NS8 5′-TCC GCA GGT TCA CCT ACG GA-3′ (White et al., 1990), respectively. The 5S and 18S rDNA probes, which were obtained by the PCR amplification of nuclear DNA, contained 200 bp and 1400 bp of the rRNA genes, respectively. The mapping of 5S and 18S rDNA sequences was performed by *in situ* hybridization (Pinkel et al., 1986). The 5S rDNA probes were labeled by nick translation with biotin-14-dATP (Roche, Mannheim, Germany), and the 18S rDNA probes were labeled with digoxigenin-11-dUTP (Roche, Mannheim, Germany), according to the manufacturer's specifications.

About thirty metaphases of each individual were analyzed using optical microscopy under a 1000× magnification, and photographed using an Olympus$^{TM}$ BX51 epifluorescence

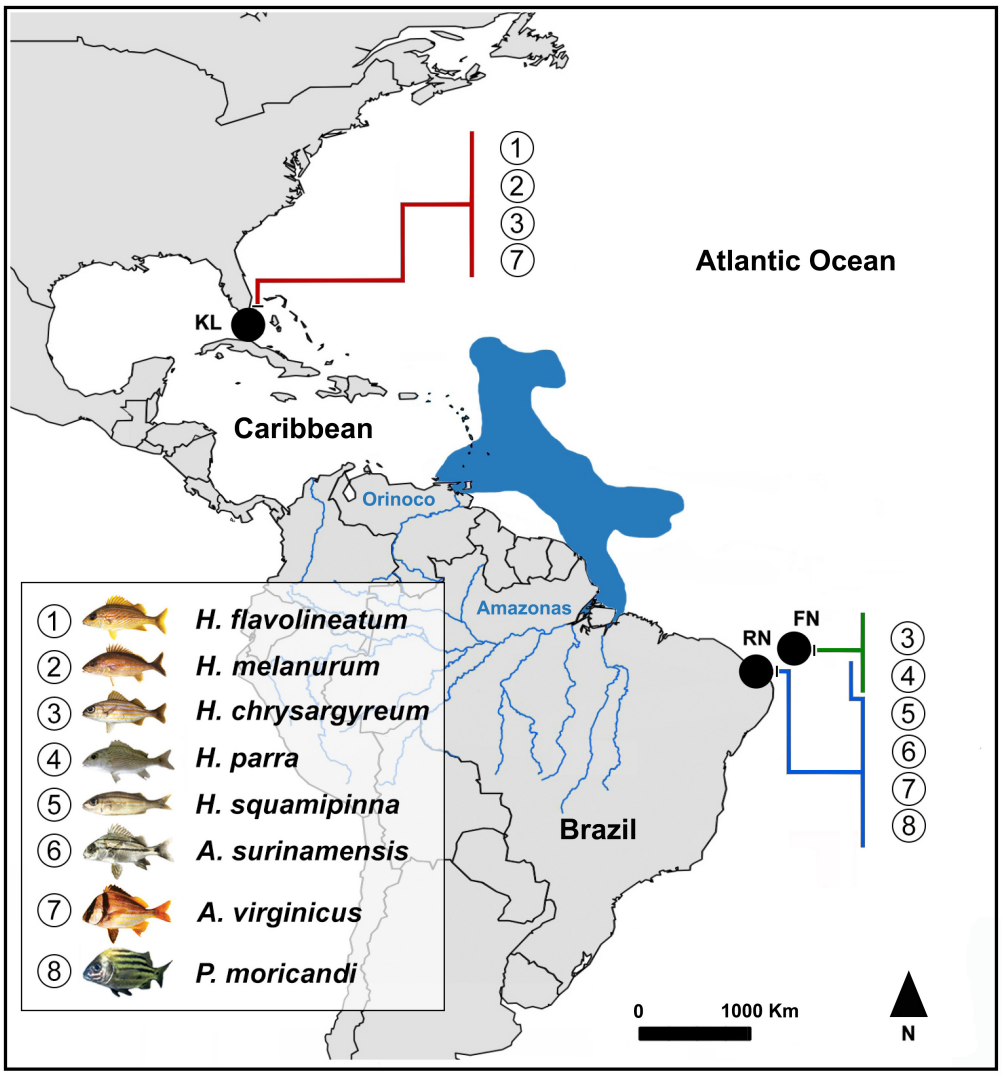


**FIGURE 1 |** Map of the collection sites of the analyzed Haemulidae species. The species of the Brazilian Province were obtained from the Archipelago of Fernando de Noronha (FN) and the Rio Grande do Norte State coast (RN), while the species of the Caribbean Province were obtained from Key Largo, Florida (United States).

photomicroscope, coupled with an Olympus DP-72 digital capture system using the cellSens® Olympus™ software. For the preparation of karyotypes, the chromosomes were classified in terms of their arms ratios (Levan et al., 1964) and organized in the descending order of size.

## Cytogenetic Meta-Analysis in Eupercaria Groups

The survey of karyotype data for members from the Eupercaria series (*sensu* Hughes et al., 2018) was performed by running searches in several scientific web portals.

To ensure adequate phylogenetic representation, karyotype data ($2n$ and NF – chromosome arms) of 671 species (~10% of total species) distributed into 62 families (38.5% of the families) and 15 orders (88% of the orders) of Eupercaria (data summarized as **Supplementary Material**) were analyzed. The karyotypes of species within the Eupercaria series were categorized based on $2n$ and NF values greater than, equal to, or smaller than the basal karyotype of $2n = 48$ acrocentric chromosomes.

NF (chromosome arm number) was determined considering metacentric/submetacentric chromosomes having two arms and subtelocentric/acrocentric chromosomes having one single arm. In species with sex chromosome systems, the karyotypes of the homogametic sex were considered as the standard for the species.

# RESULTS

All the eight haemulid species analyzed had $2n = 48$ acrocentric chromosomes (NF = 48). The karyotypes were remarkably symmetrical, exhibiting a small size variation between the largest

**TABLE 1 |** Data of the Haemulidae species samples utilized for the cytogenetic analyses.

| Species | N | Localities |
|---|---|---|
| *Anisotremus surinamensis* (Bloch, 1791) | 03 | Nísia Floresta, RN – Brazil (n = 2) |
| | | Touros, RN – Brazil (n = 1) |
| *Anisotremus virginicus* (Linnaeus, 1758) | 09 | Touros, RN – Brazil (n = 6) |
| | | Key Largo, FL, United States (n = 3) |
| *Paranisotremus moricandi* (Ranzani, 1842) | 01 | Nísia Floresta, RN – Brazil (n = 1) |
| *Brachygenys chrysargyreum* (Günther, 1859) | 08 | Fernando de Noronha – Brazil (n = 5) |
| | | Key Largo, FL, United States (n = 3) |
| *Haemulon flavolineatum* (Desmarest, 1823) | 11 | Key Largo, FL, United States (n = 11) |
| *Haemulon melanurum* (Linnaeus, 1758) | 02 | Key Largo, FL, United States (n = 2) |
| *Haemulon parra* (Desmarest, 1823) | 07 | Extremoz, RN – Brazil (n = 2) |
| | | Fernando de Noronha – Brazil (n = 4) |
| | | Nísia Floresta, RN – Brazil (n = 1) |
| *Haemulon squamipinna* (Rocha and Rosa, 1999) | 07 | Touros, RN – Brazil (n = 7) |

and the smallest chromosome pairs. Individuals of A. *virginicus*, *H. squamipinna* and *H. parra* presented chromosome pairs with visible secondary constrictions, co-localized with the NORs (**Figures 2**, **3**).

In general, the species exhibited regions with reduced heterochromatin, which was mainly located in the centromeric position, especially in the NORs, which were the only GC-rich sites in the karyotypes (**Figures 2**, **3**).

The Ag-NOR sites were located in a single chromosomal pair in all species. They were mainly positioned in the short arms of the 24th pair, except in case of the species *Brachygenys chrysargyreum* (5th pair, subterminal position of the long arm) and *Haemulon melanurum* (10th pair, pericentromeric position of the long arm) (**Figures 3**, **4**).

Fluorescent *in situ* hybridization showed a complete congruence among the 18S rDNA and Ag-NOR sites. In most species, the 5S rDNA sites were present in the terminal position on the short arm of the 10th pair (**Figures 2**, **3**), except in case of *Anisotremus virginicus* (5th pair, terminal region of the long arm), and *Paranisotremus moricandi* (24th pair, the whole short arm). *H. melanurum* exhibited an exclusive array of ribosomal sites; the 5S and 18S rDNA sites were co-located on the 10th pair of the chromosomes (**Figure 3**).

Population analyses of *A. virginicus* and *B. chrysargyreum* (Brazil/Caribbean) and *Haemulon parra* (oceanic islands and coastal areas of Brazil) did not reveal any structural chromosome divergences or changes in the organization of ribosomal genes.

The estimation of karyotype conservatism in members from the Eupercaria series based on the frequencies of 2*n* and NF values was based on a vast cytogenetic dataset encompassing 88% of the total clades in the Eupercaria series. Samples from the orders Perciformes, Labriformes, Centrarchiformes, Chaetodontiformes, and Spariformes, and the *incertae sedis* groups Lutjanidae, Haemulidae, and Sciaenidae, for whom more cytogenetic information was available, presented a high frequency of karyotypes with 2*n* = 48 (74–100%) (**Figure 5**).

The most notable degree of diversification was observed in case of the well-sampled order Tetraodontiformes (71 spp.), in which karyotypes with 2*n* = 48 were observed in only 5.6% of the species, while 87.3% of the species showed karyotypes with 2*n* < 48, and only 4.3% of the species showed karyotypes with 2*n* = 48a (**Supplementary Tables S1, S2**).

The evolutionary conservatism of the karyotype structure was estimated based on the frequency of karyotypes with NF = 48. The groups with the greatest structural conservatism were the *incertae sedis* families Lutjanidae, Sciaenidae, and Haemulidae (100, 88, and 88% of the total spp., respectively) and the order Chaetodontiformes (75% of the total spp.) (**Figure 5**).

The cytogenetic data of the diverse order Perciformes mainly showed karyotypes with NF > 48 (74% of the spp.). However, this order also includes groups such as Serranidae, which had a high frequency of karyotypes with 2*n* = 48 acrocentric chromosomes (∼61% of spp.) (**Supplementary Tables S1, S2**).

# DISCUSSION

## Chromosomal Conservatism in Haemulidae

The cytogenetic data set for members from the Haemulidae family reveals a remarkable karyotype conservatism. The karyotype stasis in the grunts was characterized by the extensive sharing of a standard karyotype composed of 2*n* = 48 acrocentric chromosomes, with a small variation in size between the largest and smallest chromosomes of the karyotype, distribution of heterochromatin in the centromeric regions of chromosomes, and distribution of Ag-NOR sites recurrently on the same chromosome pair.

In a broader context, stasis has been found to be maintained even in case of comparisons of diverse chromosome regions (Motta-Neto et al., 2012), at a wide phylogenetic spectrum, wide periods of divergence (>25 M.a), and varied biogeographic scales (Caribbean and western Atlantic).

Heterogeneous heterochromatins, when are present in variable amounts among chromosomes could suggest diverse evolutionary origins and dynamics in the karyotype complement. In contrast, the heterochromatin content in the haemulid species showed a regular and reduced distribution in the centromeric regions of the chromosomes, with only the NORs showing a high GC content.

Comparative analyses of replication bands in chromosomes from the haemulid species have shown regions of repetitive DNA with synchronous replication patterns, suggesting that it was likely that they shared functional and compositional similarities (Motta-Neto et al., 2012). In fact, cross-amplification tests for microsatellite sequences have shown a high level of success in *Haemulon* spp. (Quintero et al., 2018), providing support for the evolutionary conservation of repetitive DNA.

Analysis of repetitive DNA classes has served as a very useful approach to assess karyotype changes among groups of fishes. The physical mapping of rDNA in the chromosomes allows the

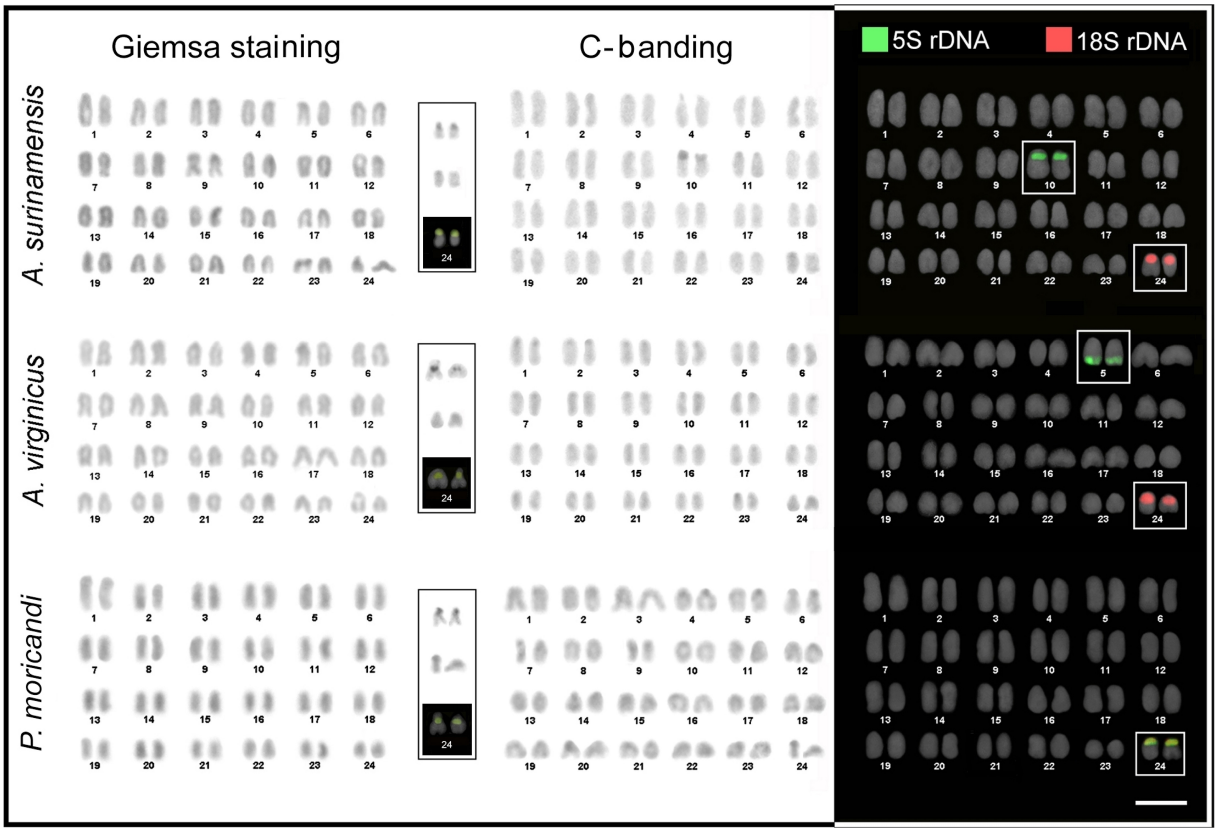


**FIGURE 2 |** Karyotypes of *Anisotremus surinamensis*, *A. virginicus*, and *Paranisotremus moricandi* visualized by conventional staining with Giemsa, C-banding, and double-FISH with18S rDNA (red) and 5S rDNA (green) probes. The NOR-bearing chromosomes are presented in the larger boxes, from the top to the bottom, as observed after silver staining, C-banding, and $CMA_3$/DAPI staining. Scale bar = 5 μm.

identification of cryptic rearrangements in the chromosomes (e.g., Borges et al., 2019) and cytotaxonomic markers (Nakayama et al., 2012; Gornung, 2013; Almeida et al., 2017). However, the interspecific recurrence of rDNA distribution patterns on the chromosomes from the haemulid species restricts their use for this purpose. In fact, in this family, the Ag-NORs/18S rDNA sites are simple, and in most of the species, they were positioned in the short arm of the smallest chromosome pair (24th pair). In some species, they may also occur in other pairs of chromosomes (pair 10 in *H. melanurum*, pair 5 in *B. chrysargyreum*, and pair 18 in *C. nobilis*), but the repertoire of changes is quite restricted.

The 5S rDNA sites were shown to be more evolutionarily diverse among the haemulid species. They were localized in one or two chromosome pairs; however, they also exhibited a considerable regularity of number and position in homeologous chromosomes, usually in the pairs 5 (long arm) and 10 (short arm). Often, these sites were non-syntenic with 18S rDNA; however, co-localized 5S/18S rDNA arrays were found in *H. melanurum* (pair 10) and *P. moricandi* (pair 24) (**Figure 4**).

The mapping of the 5S rDNA sites revealed five basic patterns of the organization of rDNA regions in Haemulidae (**Figure 4**). The first, which was the most frequent (35% of spp.), consisted of karyotypes with a single site in the pericentromeric position of chromosomes in pair 10 and was present in *Anisotremus surinamensis*, *Haemulon flavolineatum*, *H. parra*, *Haemulon squamipinna*, *B. chrysargyreum*, and *H. melanurum*, as well as in species of more basal genera like *Conodon nobilis*. The second (21% of spp.) showed the presence of sites in the terminal arm of chromosomes in pair 5 and was present in *Haemulon aurolineatum*, *Haemulon steindachneri*, *Haemulopsis corvinaeformis*, and *A. virginicus*. The third consisted of sites on chromosomes in pairs 5 and 10 (28% of the spp.) and was present in *Haemulon plumierii*, *Haemulon bonariense* (Nirchio and Oliveira, 2014), possibly in *H. aurolineatum* (Nirchio et al., 2007), and in *C. nobilis*.

The fourth pattern involved the presence of a syntenic 5S/18S rDNA array in the chromosomes of pair 10 (7% of the species) and was exclusive to *H. melanurum*; the fifth pattern, which showed a 5S/18S rDNA array in the chromosomes of pair 24 (7% occurrence), was present in *P. moricandi* (**Figure 4**).

Chromosomal stasis mainly extends to the conservation of the order of genes (Ellegren, 2010). This condition in Haemulidae has been ascertained through similarities of the replication band patterns in the chromosomes (Motta-Neto et al., 2012). Additionally, the number and similarity of the localization

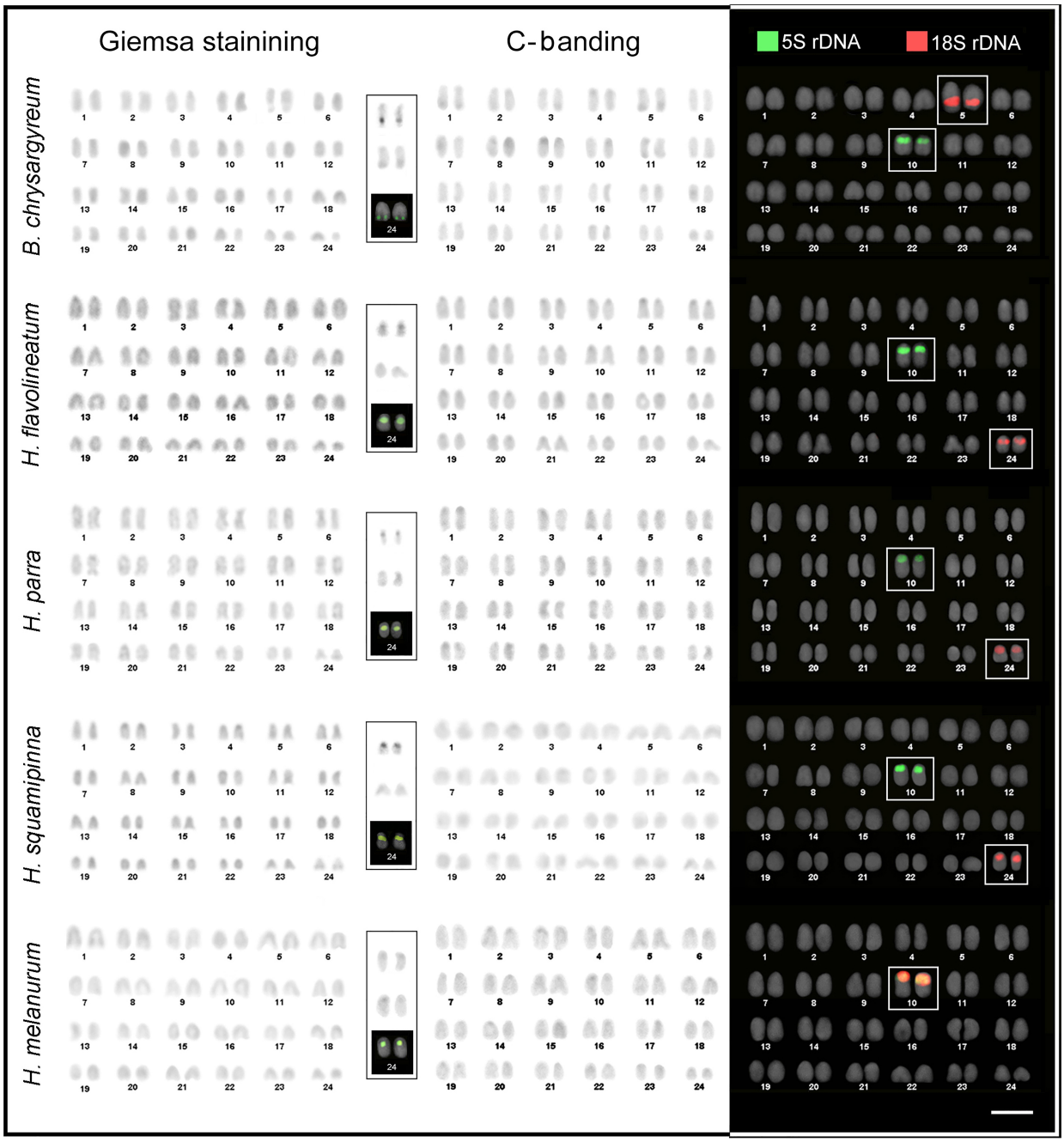


**FIGURE 3 |** Karyotypes of *Brachygenys chrysargyreum*, *Haemulon flavolineatum*, *H. parra*, *H. squamipinna*, and *H. melanurum* visualized by conventional staining with Giemsa, C-banding, and double-FISH with18S rDNA (red) and 5S rDNA (green) probes. The NOR-bearing chromosomes are presented in the larger boxes, from the top to the bottom, as observed after silver staining, C-banding, and $CMA_3$/DAPI staining. Scale bar = 5 μm.

of 5S and 18S rDNA sites have aided the identification of homeolog chromosomes, which have been widely identified in several fish groups (Vicente et al., 2001; Centofante et al., 2002; Costa et al., 2016).

Changes in the 5S rDNA sites in the grunt karyotypes were restricted to only two chromosome pairs (pairs 5 and 10). Similarly, 18S rDNA sites almost invariably occur in pair 24, which is the smallest chromosome pair, evidencing the occurrence of syntenic groups and the low evolutionary dynamics of these repetitive sequences. Thus, the phylogenetic diversification of rDNA sites mainly reveals "variations on the same theme," indicating a slow and recurrent "birth-death" process of these sequences, with sporadic co-localized 5S/18S rDNA

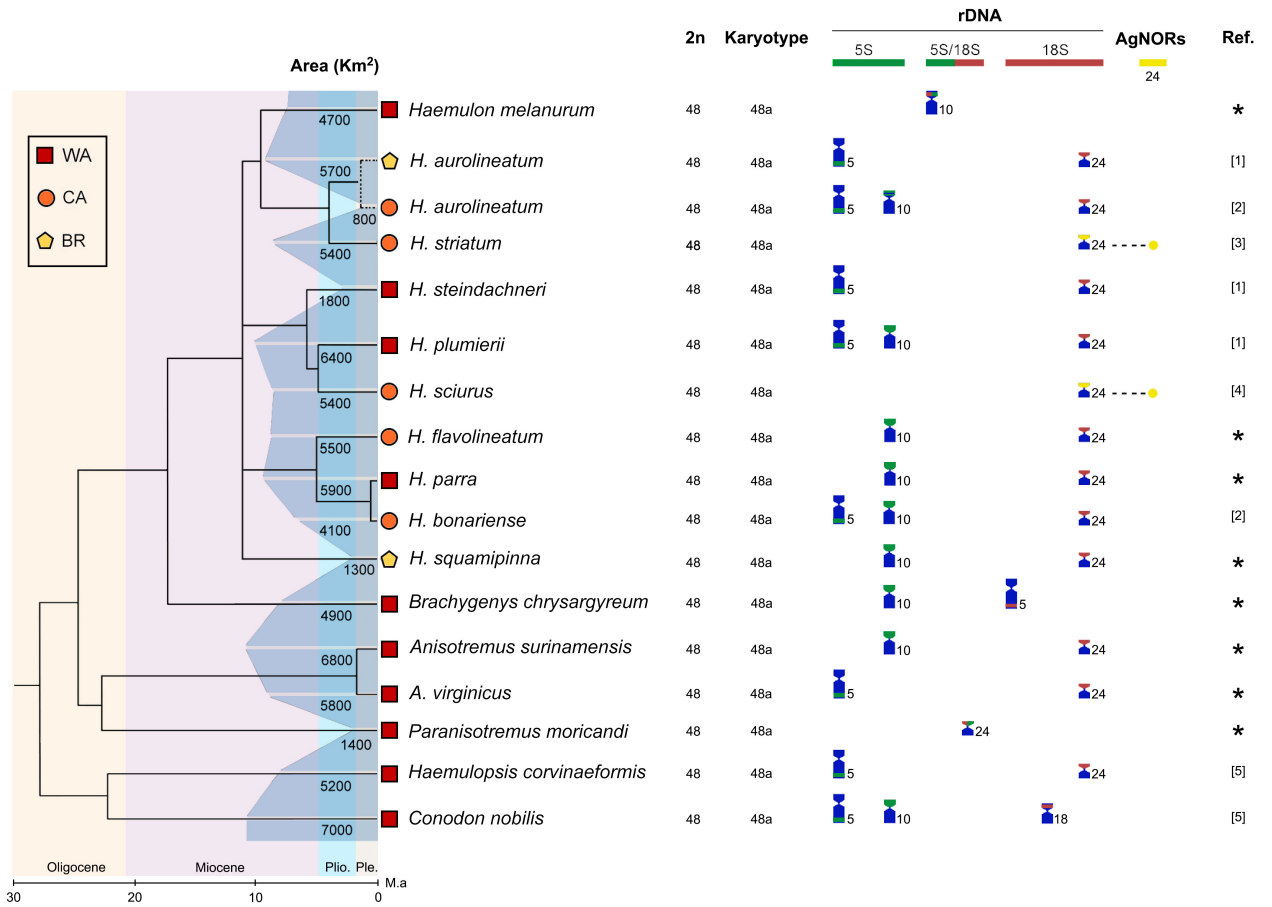


**FIGURE 4 |** Macrostructural aspects of karyotypes and the occurrence of rDNA sites in Western Atlantic haemulid species from the phylogenetic [adapted from Tavera et al. (2018)] and biogeographic perspectives. The distribution ranges of the species were estimated from the databases of the Ocean Biogeographic Information System – OBIS (https://obis.org). The references symbols refer to: * – present study; [1] – Motta-Neto et al. (2011a); [2] – Nirchio et al. (2007); [3] Lima and Molina (2004); [4] – Nirchio and Oliveira (2014); [5] – Motta-Neto et al. (2011b).

arrangements resulting from preferential transpositions between rDNA regions.

## Extrinsic Environmental Factors and Karyotype Conservatism in Haemulidae

Environmental factors and the life history of the species likely contribute to the karyotype patterns of marine fish (Molina and Galetti, 2004; Sena and Molina, 2007).

The roles played by ecological or physical barriers on gene flow and the origin of genetic diversity are well known. The action of physical barriers is evidenced by greater karyotype diversity in freshwater fish species, while in case of marine species, a greater degree of conservation is normally observed (Brum and Galetti, 1997).

The species of Haemulidae have large populations distributed over wide geographic areas, where oceanographic variables (temperature, salinity, phytoplankton biomass, zooplankton biomass, and station depth) play an important role in the distribution patterns of ichthyoplankton (Souza and Mafalda-Junior, 2019).

The western Atlantic haemulid species have broad geographic distribution areas, which range from 800 to 7,000 km$^2$ (**Figure 4**), depicting a geographical context very different from that normally found in case of freshwater fish species. This environmental scenario contributes sufficiently to gene flow, reducing the possibility of the formation of isolated groups, in which, by chance, chromosome rearrangements could be more frequently fixed. Analysis of the extensive geographic distribution, with a marked spatial overlap of sibling species, has revealed that besides vicariant events, ecological speciation also participates in the evolutionary diversification of this family of fishes (Rocha et al., 2008). In fact, in this family, recurrent sympatric speciation processes occurring in short periods of time (∼5 M.a), especially in species of the genus *Haemulon*, have been observed (Tavera et al., 2012; Tavera and Wainwright, 2019); recent events of genetic introgression have also been observed in some cases (Bernal et al., 2017).

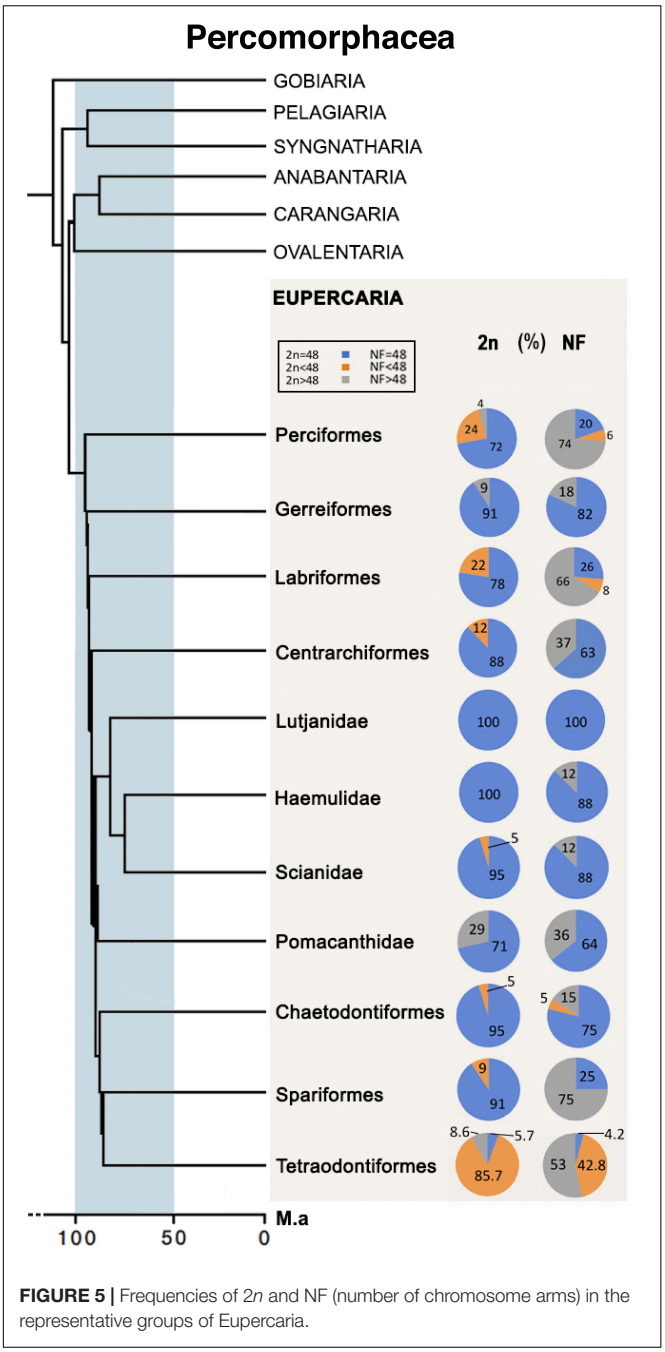


**FIGURE 5 |** Frequencies of 2*n* and NF (number of chromosome arms) in the representative groups of Eupercaria.

In ecological speciation processes, with the fragmentation of ecological niches, in general, the split of the gene pool occurs without remarkable genetic changes or high selection pressures (Rocha and Bowen, 2008), thereby producing less propitious conditions to fix conspicuous chromosome rearrangements.

An extensive sharing of karyotypes in an evolutionary lineage could be caused from a recent diversification burst, during which genetic or chromosome changes have not yet been achieved. However, although some *Haemulon* species have been shown to demonstrate recent diversification (Rocha et al., 2008; Tavera et al., 2018), merely this condition seems insufficient to explain the karyotype stasis in members from this family. In fact, shared cytogenetic patterns are maintained among genera with long periods of divergence, which may reach up to 20 M.a (Tavera et al., 2018). Other factors, such as large populations, environmental continuities, and large distribution areas, seem to contribute to the panmixia of haemulid populations.

A test of the role of biogeographic barriers in the population structure of members from Haemulidae is offered by the plume of the Amazon and Orinoco rivers, which divides the Brazilian and Caribbean Provinces and plays an important role in the diversification of the faunas of these regions (Rocha, 2003; Bowen and Karl, 2008).

Cytogenetic analyses of populations of fishes from the genera *Haemulon* and *Anisotremus* located in the north and south of the Amazonas/Orinoco barrier have revealed a dichotomous pattern of karyotype divergence. Divergences of the rDNA sites were perceptible when the populations of *H. aurolineatum* from the Caribbean (Nirchio et al., 2007) and those from the Brazilian coast were compared (Motta-Neto et al., 2011b). On the other hand, populations of *A. virginicus* from the Caribbean and Brazilian coastal areas and those of *B. chrysargyreum* from the Fernando de Noronha Archipelago (Western Atlantic) region and the Caribbean did not present a detectable karyotype differentiation.

The interpopulational karyotype patterns underscore the chromosome conservatism in this group, even under distinct biogeographical pressures strengthened by action of the Amazonas/Orinoco barrier.

## Karyotype Stasis in Marine Groups of Eupercaria

The dimension of karyotype stasis has been noted in several groups of Percomorpha (=Percomorphaceae) (Galetti et al., 2000; Molina, 2007), the most species-rich clade of modern fishes (Hughes et al., 2018). This clade has been reported to have a controversial taxonomic cohesion; however, the use of comparative genomics databases has promoted a remarkable phylogenetic redefinition, leading to the identification of its major lineages (Near et al., 2012; Betancur-R et al., 2017; Hughes et al., 2018). Its enormous diversity and increasingly robust phylogenetic relationships have made this group an excellent model to investigate karyotype evolution.

Among the 9 Percomorpha series, Eupercaria, the most diverse series, with more than 163 families and 6,000 species (*sensu* Hughes et al., 2018), encompasses orders with marked karyotype conservatism. In fact, members of the clade Percomorpha have been shown to present karyotypes with $2n = 48$ acrocentric chromosomes both in basal and recent groups, clearly indicating that this pattern is symplesiomorphic (Galetti et al., 2000; Motta-Neto et al., 2012; Calado et al., 2014; Molina et al., 2014; Costa et al., 2016). However, thus far, its phylogenetic extent has only been inferred superficially.

Presently, the karyotype conservation estimated in members of the Eupercaria series, allowing a robust view of their karyotype evolution. In ten of the fifteen orders, most species show karyotypes of $2n = 48$. Analysis of groups with greater diversity

and cytogenetic information representativeness (**Figure 5**) revealed a high frequency of this diploid number (74–100%). Several groups, such as Chaetodontiformes (80% of spp.), Serranidae, which belongs to the large order Perciformes, or *incertae sedis* families such as Haemulidae, Sciaenidae, and Lutjanidae (86–100% of spp.), also presented high structural conservatism (NF = 48). The NF maintenance trend has been shown to continue when orders for which lower levels of cytogenetic information are available were considered (Gerreiformes, Ephippiformes, Lobotiformes).

Among the Eupercaria series groups, Labriformes represents one of the most morphologically and ecologically diversified clades in size, shape, and color (Westneat and Alfaro, 2005). Previous cytogenetic studies in this group have showed intermediate rates of chromosomal changes (Molina et al., 2014). Our data reinforce this condition, indicating considerable diversification in the NF (79% of species), despite a more conservative 2*n* variation (61% of species), with regards to a basal karyotype with 2*n* = 48a. On the other hand, the order Tetraodontiformes showed the most accentuated diversification of 2*n* (present only in 5.6% of species) and NF values (4.3% of the species), among the groups. This order, which comprises 438 spp. (Fricke et al., 2019), represents one of the main branches of the teleost diversification; its members present notable morphological and ecological characteristics, and variations in genome size (Brainerd et al., 2001; Santini and Tyler, 2003; Jaillon et al., 2004), which seem to have influenced their rather singular karyotype trends (Sá-Gabriel and Molina, 2005).

Evolutionary stasis is a broad process that involves the absence of conspicuous modifications in different dimensions throughout the evolutionary history of the organisms; its causes have not always been well understood. In some cases, this process may be transient, and has been shown to vary with regards to the occurrence of disruptive factors (Burt, 2001). In some cases, it has been shown to be characterized by the maintenance of a trait that has persisted for millions of years as a possible result of stabilizing selection (Wake et al., 1983). In adaptive terms, evolutionary stasis may reflect developmental homeostasis or static permanence at adaptive peaks (Lerner, 1954; Charlesworth et al., 1982; Kirkpatrick, 1982; Mayr, 1982; Templeton, 1982).

This process could still be related to parameters such as time, inertia, or selection. An explanation of temporal causes would involve the extensive sharing of a karyotype pattern in a group that recently underwent radiation and still did not have enough time to fix karyotype changes (Sola et al., 1981). Another possibility stems from the evolutionary inertia of karyotypes, caused by the absence of environmental destabilizing triggers (extrinsic to chromosomes) or cytogenetic factors (intrinsic to chromosomes). The extrinsic factors favoring gene flow include the absence of geographic (or ecological) barriers, dispersive capacity, population size, and the occurrence of historically stable environments. An additional possibility, but one that is less likely, is that stasis could be caused by the action of stabilizing selection, promoting active restriction to changes in chromosomes. Alternatively, it could be the product of orthoselective processes, where different karyotypes present the tendency to fix certain types of chromosome rearrangements along their evolutionary trajectories, promoting the formation of symmetrical karyotypes, with chromosomes showing approximately the same sizes and morphologies (White, 1973; King, 1981).

Using the available cytogenetic data set from the Eupercaria series, it was possible to analyze the validity of some explanations for the slow karyotype diversification. This clade encompasses evolutionary groups with extremely high diversity (e.g., with regards to natural history, ecological patterns, biogeography, habitat variety) (Nelson et al., 2016) and with very long divergence (70−38 M.a) (Yang et al., 2018); this makes it unlikely that karyotype stasis represents an adaptive condition *stricto sensu* or is a result of a short divergence time. On the other hand, given the evidence of extensive chromosome synteny, one cannot rule out the possibility that chromosome particularities of a group or related groups, such as low and homogeneous heterochromatin contents, contribute to the promotion and establishment of rearrangements in members from the Eupercaria series.

The vast majority of Eupercaria species are marine organisms (Betancur-R et al., 2017), which are present in large populations, and in general, have extensive distribution limits (Nelson et al., 2016). Under these conditions, a set of species of these groups show remarkable karyotype conservatism (Molina, 2007; Motta-Neto et al., 2011a,b, 2012; Molina et al., 2014) suggesting a role of the environment in the slow process of chromosome diversification.

These conditions presented in members from the Eupercaria series strengthen the hypothesis that environmental and intrinsic factors influencing the chromosomes of a group of fish can act synergistically, underscoring the chromosome conservatism in this group. The period of greatest divergence for members from the Eupercaria series (>70 M.a), which coincides with that for groups that share basal karyotypes, still clarifies that processes of karyotype stasis in fish can be maintained for extremely long periods of time.

## Final Remarks

Evolutionary karyotype patterns of large vertebrate groups are estimated based on the phylogenetic spectrum, temporal patterns, and cytogenetic approaches. Cytogenetic data from a representative number of haemulid species, which were used here as an evolutionary model, and the meta-analysis of the karyotype patterns of members of the Eupercaria contributed for a better understanding of the process of karyotype stasis in fishes. Haemulid karyotypes are remarkably conserved with regards to the macro- and microstructure of the chromosomes, suggesting the maintenance of extensive synteny throughout their evolutionary history. The ancient periods of evolutionary divergence among its lineages do not support the hypothesis that karyotype stasis is a byproduct of recent processes of speciation. On the other hand, these data suggest that the causes of this evolutionary process are complex, resulting from the synergistic action of the intrinsic characteristics of the chromosomes, together with their biological characteristics and the geographical context in which the species are inserted. The sharing of a generally homogeneous and reduced content of heterochromatin in members of the Haemulidae family seems to minimize

the destabilizing role promoted by a high and heterogeneous content of repetitive DNAs and a repository of extensive sequence repertoires (e.g., transposable elements, microsatellites, multigene families) capable of generating evolutionary instability in the chromosomes. In addition, life history and biological characteristics, such as large populations, which are found in members in the Eupercaria series, could restrict the fixation of heterozygous variants. In general, cytogenetic analysis of the members from the Eupercaria revealed a broad scenario of karyotype stasis, except for the members from Tetraodontiformes, an order with diversified karyotype evolutionary trends. In summary, karyotype stasis is not an evolutionary process with a single cause, but is rather, a product of the conjunctural and synergistic effects of different factors. These data create a counterpoint for understanding the extreme diversification observed in other groups that do not meet the criteria for determining karyotype stasis.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/**Supplementary Files**.

## ETHICS STATEMENT

The animal study was reviewed and approved by the Committee of Ethics for the use of Animals of the Federal University of Rio Grande do Norte.

## AUTHOR CONTRIBUTIONS

CM-N and WM contributed to the conception and design of the study. CM-N, GC, and KA performed the cytogenetic analyses. CM-N, KA, and WM organized the karyotype database. CM-N and WM wrote the first draft of the manuscript. MC, LB, and RA wrote the sections of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

## FUNDING

This work received national funds through the CNPq – Brazilian National Council for Scientific and Technological Development through project Proc. 442664/2015-0. This research was partly supported by the INCT "Marine Sciences" (565054/2010-4).

## ACKNOWLEDGMENTS

We thank ICMBio SISBIO for permits to collect specimens (licenses 19135-1, 27027-2, and 131360-1). We are also grateful to Jose Garcia Jr. for taxonomic identification of specimens.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00628/full#supplementary-material

## REFERENCES

Almeida, L. A. H., Nunes, L. A., Bitencourt, J. A., Molina, W. F., and Affonso, P. R. (2017). Chromosomal evolution and cytotaxonomy in wrasses (Perciformes; Labridae). *J. Hered.* 108, 239–253. doi: 10.1093/jhered/esx003

Aprea, G., Andreone, F., Capriglione, T., Odierna, G., and Vences, M. (2004). Evidence for a remarkable stasis of chromosome evolution in Malagasy treefrogs (Boophis, Mantellidae). *Ital. J. Zool.* 71, 237–244. doi: 10.1080/11250000409356641

Arai, R. (2011). *Fish Karyotypes: A Check List*. Tokyo: Springer.

Bernal, M., Gaither, M., Simison, W., and Rocha, L. (2017). Introgression and selection shaped the evolutionary history of sympatric sister-species of coral reef fishes (genus: *Haemulon*). *Mol. Ecol.* 26, 639–652. doi: 10.1111/mec.13937

Bernal, M. A., Dixon, G. B., Matz, M. V., and Rocha, L. A. (2019). Comparative transcriptomics of sympatric species of coral reef fishes (genus: *Haemulon*). *PeerJ* 1:e6541. doi: 10.7717/peerj.6541

Betancur-R, R., Broughton, R. E., Wiley, E. O., Carpenter, K., López, J. A., Li, C., et al. (2013). The tree of life and a new classification of bony fishes. *PLoS Curr.* 5:ecurrents.tol.53ba26640df0ccaee75bb165c8c26288. doi: 10.1371/currents.tol.53ba26640df0ccaee75bb165c8c26288

Betancur-R, R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., et al. (2017). Phylogenetic classification of bony fishes. *BMC Evol. Biol.* 17:162. doi: 10.1186/s12862-017-0958-3

Bomfleur, B., McLoughlin, S., and Vajda, V. (2014). Fossilized nuclei and chromosomes reveal 180 million years of genomic stasis in royal ferns. *Science* 343, 1376–1377. doi: 10.1126/science.1249884

Borges, A. T., Cioffi, M. B., Bertollo, L. A. C., Soares, R. X., Costa, G. W. W. F., and Molina, W. F. (2019). Paracentric inversions differentiate the conservative karyotypes in two *Centropomus* species (Teleostei: Centropomidae). *Cytogenet. Genome Res.* 157, 239–248. doi: 10.1159/000499748

Bowen, B. W., and Karl, S. A. (2008). Population genetics and phylogeography of sea turtles. *Mol. Ecol.* 16, 4886–4907. doi: 10.1111/j.1365-294x.2007.03542.x

Brainerd, E. L., Slutz, S. S., Hall, E. H., and Phillis, R. W. (2001). Patterns of genome size evolution in Tetraodontiform fishes. *Evolution* 55, 2363–2368. doi: 10.1111/j.0014-3820.2001.tb00750.x

Brum, M. J. I., and Galetti, P. M. Jr. (1997). Teleostei ground plan karyotype. *J. Comp. Biol.* 2, 91–102.

Burt, D. B. (2001). Evolutionary stasis, constraint and other terminology describing evolutionary patterns. *Biol. J. Linn. Soc.* 72, 509–517. doi: 10.1111/j.1095-8312.2001.tb01334.x

Calado, L. L., Bertollo, L. A. C., Cioffi, M. B., Costa, G. W. W. F., Jacobina, U. P., and Molina, W. F. (2014). Evolutionary dynamics of rDNA genes on chromosomes of the *Eucinostomus* fishes: cytotaxonomic and karyoevolutive implications. *Genet. Mol. Res.* 13, 9951–9959. doi: 10.4238/2014.November.27.24

Centofante, L., Bertollo, L. A. C., and Moreira-Filho, O. (2002). A ZZ/ZW sex chromosome system in a new species of the genus *Parodon* (Pisces, Parodontidae). *Caryologia* 55, 139–150. doi: 10.1080/00087114.2002.10589270

Charlesworth, B., Lande, R., and Slatkin, M. (1982). A neo-darwinian commentary on macroevolution. *Evolution* 36, 474–498. doi: 10.1111/j.1558-5646.1982.tb05068.x

Costa, G. W. W. F., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2016). The evolutionary dynamics of ribosomal genes, histone H3, and transposable Rex elements in the genome of Atlantic snappers. *J. Hered.* 107, 173–180. doi: 10.1093/jhered/esv136

Ellegren, H. (2010). Evolutionary stasis: the stable chromosomes of birds. *Trends Ecol. Evol.* 25, 283–291. doi: 10.1016/j.tree.2009.12.004

Fricke, R., Eschmeyer, W. N., and Fong, J. D. (2019). *Eschmeyer's Catalog of Fishes: Species by Family/Subfamily*. http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp (accessed on May 16 2019).

Galetti, P. M. Jr., Aguilar, C. T., and Molina, W. F. (2000). An overview on marine fish cytogenetics. *Hydrobiologia* 420, 55–62. doi: 10.1023/A:1003977418900

Gold, J. R., Li, C., Shipley, N. S., and Powers, P. K. (1990). Improved methods for working with fish chromosomes with a review of metaphase chromosome banding. *J. Fish Biol.* 37, 563–575. doi: 10.1111/j.1095-8649.1990.tb05889.x

Gornung, E. (2013). Twenty years of physical mapping of major ribosomal RNA genes across the Teleosts: a review of research. *Cytogenet. Genome Res.* 141, 90–102. doi: 10.1159/000354832

Howell, W. M., and Black, D. A. (1980). Controlled silver staining of nucleolus organizer region with protective colloidal developer: a 1st-step method. *Experientia* 36, 1014–1015. doi: 10.1007/bf01953855

Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., et al. (2018). Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6249–6254. doi: 10.1073/pnas.1719358115

Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N., Mauceli, E., et al. (2004). Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957. doi: 10.1038/nature03025

Kahl, G. (2015). *The Dictionary of Genomics, Transcriptomics and Proteomics*. Verlag: John Wiley & Sons.

King, M. (1981). "Chromosome change and speciation in lizards," in *Evolution and Speciation. Essays in Honor of M. J. D. White*, eds W. R. Atchley, and D. Woodruff, (London: Cambridge University Press), 262–285.

Kirkpatrick, M. (1982). Sexual selection and the evolution of the female choice. *Evolution* 36, 1–12. doi: 10.1111/j.1558-5646.1982.tb05003.x

Lerner, I. M. (1954). *Genetic Homeostasis*. Edinburgh: Oliver & Boyd.

Levan, A., Fredga, K., and Sandberg, A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Lima, L. C. B., and Molina, W. F. (2004). Homeostase cariotípica em três espécies do gênero *Haemulon* (Haemulidae, Perciformes) do litoral Potiguar. Anais X Simpósio de Citogenética e Genética de Peixes. *Universidade Federal do Rio Grande do Norte Natal Brasil* 1:113.

Mandáková, T., Heenan, P. B., and Lysak, M. A. (2010). Island species radiation and karyotypic stasis in Pachycladon allopolyploids. *BMC Evol. Biol.* 10:367. doi: 10.1186/1471-2148-10-367

Mayr, E. (1982). *The Growth of Biological Thought*. Cambridge, MA: Harvard Univ. Press.

Molina, W. F. (2007). "Chromosomal changes and stasis in marine fish groups," in *Fish Cytogenetics*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor, (Enfield: Science Publishers), 69–110.

Molina, W. F., Alves, D. E. O., Araujo, W. C., Martinez, P. A., Silva, M. F. M., and Costa, G. W. W. F. (2010). Performance of human immunostimulating agents in the improvement of fish cytogenetics. *Genet. Mol. Res.* 9, 1807–1810. doi: 10.4238/vol9-3gmr840

Molina, W. F., Costa, G. W. W. F., Soares, R. X., and Affonso, P. R. A. M. (2013). Extensive chromosome conservatism in Atlantic butterflyfishes, genus *Chaetodon* Linnaeus, 1758: implications for the high hybridization success. *Zool. Anz.* 253, 137–142. doi: 10.1016/j.jcz.2013.10.001

Molina, W. F., and Galetti, P. M. Jr. (2004). Karyotypic changes associated to the dispersive potential on Pomacentridae (Pisces, Perciformes). *J. Exp. Mar. Biol. Ecol.* 309, 109–119. doi: 10.1016/j.jembe.2004.03.011

Molina, W. F., Martinez, P. A., Bertollo, L. A. C., and Bidau, C. J. (2014). Preferential accumulation of sex and Bs chromosomes in biarmed karyotypes by meiotic drive and rates of chromosomal changes in fishes. *An. Acad. Bras. Ciênc.* 86, 1801–1812. doi: 10.1590/0001-3765201420130489

Motta-Neto, C. C., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2011a). Extensive chromosomal homologies and evidence of karyotypic stasis in Atlantic grunts of the genus *Haemulon* (Perciformes). *J. Exp. Mar. Biol. Ecol.* 401, 75–79. doi: 10.1016/j.jembe.2011.02.044

Motta-Neto, C. C., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2011b). Molecular cytogenetic analysis of Haemulidae fish (Perciformes): evidence of evolutionary conservation. *J. Exp. Mar. Biol. Ecol.* 407, 97–100. doi: 10.1016/j.jembe.2011.07.014

Motta-Neto, C. C., Lima-Filho, P. A., Araújo, W. C., Bertollo, L. A. C., and Molina, W. F. (2012). Differentiated evolutionary pathways in *Haemulidae* (Perciformes): karyotype stasis versus morphological differentiation. *Rev. Fish Biol. Fish.* 22, 457–465. doi: 10.1007/s11160-011-9236-4

Nakayama, C. M., Feldberg, E., and Bertollo, L. A. C. (2012). Karyotype differentiation and cytotaxonomic considerations in species of Serrasalmidae (Characiformes) from the Amazon basin. *Neotrop. Ichthyol.* 10, 53–58. doi: 10.1590/S1679-62252012000100005

Near, T. J., Eytan, R. I., Dornburg, A., Kuhn, K. L., Moore, J. A., Davis, M. P., et al. (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13698–13703. doi: 10.1073/pnas.1206625109

Nelson, J. S., Grande, T. C., and Wilson, M. V. H. (2016). *Fishes of the World*. New York: John Wiley & Sons Inc.

Nirchio, M., Gaviria, J. I., Oliveira, C., Ferreira, I. A., and Martins, C. (2007). Cytogenetic analysis of three species of the genus *Haemulon* (Teleostei: Haemulinae) from Margarita Island, Venezuela. *Genetica* 131, 135–140. doi: 10.1007/s10709-006-9123-4

Nirchio, M., and Oliveira, C. (2014). Cytogenetics as a taxonomic tool in fish. *Artículo de Revisión Agrobiología* 26:11.

Pendás, A. M., Morán, P., and García-Vázquez, E. (1994). Organization and chromosomal location of the major histone cluster in brown trout, Atlantic salmon and rainbow trout. *Chromosoma* 103, 147–152. doi: 10.1007/BF00352324

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2934–2938. doi: 10.1073/pnas.83.9.2934

Quintero, P. E. P., Tavera, J. J., Millán-Márquez, A. M., and Acero, P. A. (2018). Amplification of microsatellites in two haemulid species (*Haemulon aurolineatum* and *Haemulon steindachneri*). *Bol. Invest. Mar. Cost* 47, 151–157. doi: 10.25268/bimc.invemar.2018.47.1.743

Rocha, L., Lindeman, K., Rocha, C., and Lessios, H. A. (2008). Historical biogeography and speciation in the reef fish genus *Haemulon* (Teleostei: Haemulidae). *Mol. Phylogenet. Evol.* 48, 918–928. doi: 10.1016/j.ympev.2008.05.024

Rocha, L. A. (2003). Patterns of distribution and processes of speciation in Brazilian reef fishes. *J. Biogeogr.* 30, 1161–1171. doi: 10.1046/j.1365-2699.2003.00900.x

Rocha, L. A., and Bowen, B. W. (2008). Speciation in coral reef fishes. *J. Fish Biol.* 72, 1101–1121. doi: 10.1111/j.1095-8649.2007.01770.x

Sá-Gabriel, L. G., and Molina, W. F. (2005). Karyotype diversification in fishes of theBalistidae, Diodontidae and Tetraodontidae (Tetraodontiformes). *Caryologia* 58, 229–237. doi: 10.1080/00087114.2005.10589455

Samad, N. A., Dagher-Kharrat, M. B., Hidalgo, O., El Zein, R., Douaihy, B., and Siljak-Yakovlev, S. (2016). Unlocking the karyological and cytogenetic diversity of iris from lebanon: *Oncocyclus* section shows a distinctive profile and relative stasis during its continental radiation. *PLoS One* 11:e0160816. doi: 10.1371/journal.pone.0160816

Santini, F., and Tyler, J. C. (2003). A phylogeny of the families of fossil and extant Tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), upper cretaceous to recent. *Zool. J. Linnean Soc.* 139, 565–617. doi: 10.1111/j.1096-3642.2003.00088.x

Schweizer, D. (1980). Simultaneous fluorescent staining of R bands and specific heterochromatic regions (DA-DAPI bands) in human chromosomes. *Cytogenet. Cell Genet.* 27, 190–193. doi: 10.1159/000131482

Sena, D. C. S., and Molina, W. F. (2007). Chromosomal rearrangements associated with pelagic larval duration in Labridae (Perciformes). *J. Exp. Mar. Biol. Ecol.* 353, 203–210. doi: 10.1016/j.jembe.2007.08.020

Sessions, S. K., and Kezer, J. (1991). "Evolutionary cytogenetics of Bolitoglossine salamanders (family Plethodontidae)," in *Amphibian Cytogenetics and Evolution*, eds D. M. Green, and S. K. Sessions, (San Diego: Academic Press), 89–130. doi: 10.1016/B978-0-12-297880-7.50009-3

Sola, L., Cataudella, S., and Capanna, E. (1981). New developments in vertebrate cytotaxonomy III. Karyology of bony fishes: a review. *Genetica* 54, 285–328. doi: 10.1007/BF00135048

Souza, C. S., and Mafalda-Junior, P. O. (2019). Large-scale spatial and temporal variability of larval fish assemblages in the tropical Atlantic

Ocean. *Acad. Bras. Ciênc.* 91:e20170567. doi: 10.1590/0001-37652018201 70567

Sumner, A. T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 75, 304–306. doi: 10.1016/0014-4827(72) 90558-7

Tavera, J., Acero, P. A., and Wainwright, P. C. (2018). Multilocus phylogeny, divergence times, and a major role for the benthic-topelagic axis in the diversification of grunts (Haemulidae). *Mol. Phylogenet. Evol.* 121, 212–223. doi: 10.1016/j.ympev.2017.12.032

Tavera, J. J., and Wainwright, P. C. (2019). Geography of speciation affects rate of trait divergence in haemulid fishes. *Proc. R. Soc. Lond.* 286:20182852. doi: 10.1098/rspb.2018.2852

Tavera, J. T., Acero, A. P., Balart, E. F., and Bernardi, G. (2012). Molecular phylogeny of grunts (Teleostei, Haemulidae), with an emphasis on the ecology, evolution, and speciation history of new world species. *BMC Evol. Biol.* 12:57. doi: 10.1186/1471-2148-12-57

Templeton, A. R. (1982). "Adaption and the integration of evolutionary forces," in *Perspectives on Evolution*, ed. R. Milkman, (Sunderland,MA: Sinauer Associates), 15–31.

Vicente, V. E., Jesus, C. M., and Moreira-Filho, O. (2001). Chromosomal localization of 5S and 18S rRNA genes in three *Parodon* species (Pisces, Parodontidae). *Caryologia* 54, 365–369. doi: 10.1080/00087114.2001.105 89247

Wake, D. B., Roth, G., and Wake, M. H. (1983). On the problem of stasis in organismal evolution. *J. Theor. Biol.* 101, 211–224. doi: 10.1016/0022-5193(83) 90335-1

Westneat, M. W., and Alfaro, M. E. (2005). Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Mol. Phylogenet. Evol.* 36, 370–390. doi: 10.1016/j.ympev.2005.02.001

White, M. J. D. (1973). *Animal Cytology and Evolution*. Cambridge: University Press.

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics," in *PCR Protocols: A Guide to Methods and Applications*, eds M. A. Innis, D. H. Gelfand, J. J. Shinsky, and T. J. White, (San Diego: Academic Press Inc), 315–322. doi: 10. 1016/b978-0-12-372180-8.50042-1

Yang, H., Xia, J., Zhang, J. E., Yang, J., Zhao, H., Wang, Q., et al. (2018). Characterization of the complete mitochondrial genome sequences of three croakers (Perciformes, Sciaenidae) and novel insights into the phylogenetics. *Int. J. Mol. Sci.* 19:1741. doi: 10.3390/ijms19061741

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 16 October 2019
doi: 10.3389/fmars.2019.00637

# Artificial Induction and Genetic Structure Analysis of Tetraploid Turbot *Scophthalmus maximus*

*Zhihao Wu[1,2], Lijuan Wang[1,2], Yunliang Lu[3], Xiangping Zhu[3], Xinlu Yue[4] and Feng You[1,2]**

[1] Key Laboratory of Experimental Marine Biology, Center for Ocean Mega-Science, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2] Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China, [3] Marine Science and Engineer College, Qingdao Agricultural University, Qingdao, China, [4] Shenghang Aquatic Science and Technology Co., Ltd., Weihai, China

Artificial tetraploid induction is one of the important techniques of fish chromosome manipulation, and it is the first step for triploid breeding. There are a few reports to artificial induction of tetraploid in marine fish. The induction and survival rates were usually low. We firstly optimized the tetraploid induction conditions in turbot *Scophthalmus maximus*, one of the most important maricultural fish in China and Europe. For the initiate time of treatment, which is the most important factor in tetraploid induction, the first cleavage index (FCI) was used to reduce the influences of genetic origin and environment factors. Overall, the optimal initiation time for pressure shock was 15 min before the first cleavage at 14.8–15.5°C. The optimal treatment pressure and treatment duration were 67.5 MPa and 6 min. The regression equation prediction model was: The optimal initiation time = 0.982 FCI − 12.182 or the optimal initiation time = 0.85 FCI. Then two tetraploid induction (4n1 and 4n2) populations were obtained under the optimal conditions with diploid controls (2n1 and 2n2). The induction rates in tetraploid induction (4n) populations at hatched larvae stage could reach 100%. The genetic structure of these two 4n populations was also studied. Two to four alleles in each locus were detected in diploid (2n) and 4n populations, respectively. Private alleles were only appeared at locus Sma-USC21, with two alleles lost in 4n populations. Eleven and fourteen loci in 2n and 4n populations respectively showed a negative genetic deviation index. 3D-FCA analysis showed that the two 2n and two 4n populations have obvious differences. The numbers of locus deviating from Hardy–Weinberg equilibrium in 2n1, 4n1, 2n2, and 4n2 populations were 6, 9, 12, and 7, respectively. Overall, 12 loci in either 2n or 4n population deviated from Hardy–Weinberg equilibrium. Tetraploid induction population showed lower heterozygosity and higher heterozygote deletion.

**Keywords: turbot *Scophthalmus maximus*, tetraploid, artificial induction, first cleavage index, genetic diversity**

OPEN ACCESS

***Edited by:***
*Dongdong Xu,*
*Marine Fisheries Research Institute of Zhejiang, China*

***Reviewed by:***
*Jie Mei,*
*Huazhong Agricultural University, China*
*Tian Yongsheng,*
*National Astronomical Observatories (CAS), China*
*Haiyang Yu,*
*Ocean University of China, China*

***Correspondence:**
*Feng You*
*youfeng@qdio.ac.cn*

***Specialty section:***
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

***Received:*** *25 May 2019*
***Accepted:*** *30 September 2019*
***Published:*** *16 October 2019*

***Citation:***
*Wu Z, Wang L, Lu Y, Zhu X, Yue X and You F (2019) Artificial Induction and Genetic Structure Analysis of Tetraploid Turbot Scophthalmus maximus. Front. Mar. Sci. 6:637. doi: 10.3389/fmars.2019.00637*

## INTRODUCTION

Polyploidization results in a wide variety of beneficial effects for genetic breeding. The most common chromosome manipulation in fish is to use a physical shock, usually a pressure or heat shock, to induce retention of the second polar body during meiosis, resulting in triploid fish. An alternative method for producing triploids is to first make tetraploids by suppression of early cell

division in the zygote (Chourrout, 1984; Zhang and Onozato, 2004; Zhu et al., 2017) and then mate the fertile tetraploids with normal diploids to generate triploid offspring. The tetraploid-derived triploids are not exposed to the trauma of the induction shock and show higher survival rate than induced triploids. Generation of tetraploid-derived triploids has been achieved in rainbow trout (*Oncorhynchus mykiss*) (Chourrout et al., 1986; Blanc et al., 1993). However, tetraploid-derived triploid was unavailable in marine fish until now, due to the low induction rate of tetraploid.

Research on chromosome set doubling has been going on since the first part of the last century. Tetraploid has been induced in many finfish species, including rainbow trout (Thorgaard et al., 1981; Chourrout et al., 1986), channel catfish (*Ictalurus punctatus*) (Bidwell et al., 1985), tilapia (*Oreochromis aureus*) (Don and Avtalion, 1988) and silver crucian carp (*Carassius auratus*) (Gui et al., 1993). In marine fishes, induction of tetraploid has been reported in European sea bass (*Dicentrarchus labrax*) (Peruzzi and Chatain, 2003), yellow perch (*Perca flavescens*) (Malison et al., 1993), olive flounder (*Paralichthys olivaceus*) (Yi et al., 2012), and half-smooth tongue sole (*Cynoglossus semilaevis*) (Li et al., 2012). However, the optimal values of treatment parameters, such as initiate treatment time, treatment intensity (temperature or pressure) and treatment duration, differ greatly among species. We also studied the artificial induction of tetraploid turbot (*Scophthalmus maximus*) (Wu et al., 2014), but the result was not stable and the induction rate was low. The viability of tetraploids was low in most instances and even unviable in several cases (Rothbard et al., 1997). The low yield of tetraploids, high frequency of abnormality and mosaicism, and especially the relatively low fertility of tetraploid females have prevented the establishment of tetraploid broodstock (Myers and Hershberger, 1991). Some possible mechanisms for the occurrence of mosaicism have been proposed by several investigators. Mosaics may be produced from tetraploid by conversion of some of tetraploid to diploid cells resulting from tetrapolar division (Yamaki et al., 1999) or somatic meiosis (Zhang and Onozato, 2004), or may be generated from the start by the simultaneous appearance of both types of ploidy cells after treatment. Maybe it is a reason to lead to the ploidy instability of artificial induced tetraploid. So, most reports on the production of tetraploid finfish have based their claims on analyses of early embryos or fry, and successful examples have been reported in a limited number of freshwater fish as mentioned above.

Induction of tetraploid requires the precise timing of the application of the shock treatment, usually a pressure shock. In salmonid, the embryo development lasts a long period. The initiation time of tetraploid induction is based on a percentage of the first cleavage interval (FCI), which is the time span between insemination and first cleavage of the zygote (Chourrout, 1984; Myers et al., 1986). The initiation time was also expressed as the relative dimensionless unit τ0, which is equivalent to the duration of one mitotic cycle during synchronous cell divisions in the initial stages of embryogenesis (Cherfas et al., 1993). By comparison, in turbot and most marine fish, the embryo development was relatively short. It seems that the optimal initiation time is easier to ascertain. The initiation time of tetraploid was mostly calculated by minutes after fertilization or minutes before the first cleavage. However, the precise timing of the shock treatment was hard to get due to the different embryo development speed caused by different genetic origin and breeding temperature.

Tetraploid usually shows higher heterozygosity because it has four sets of chromosome. It could help to produce relatively large fish (Chourrout et al., 1986). The tetraploid also plays an important role in evolution due to its high heterozygosity (Amores et al., 2004; Pasquier et al., 2017). In artificial induced tetraploid, all the chromosome sets come from the diploid parents. The genetic diversity in tetraploid should be equal to the diploid full sib family in theory. Or the genetic diversity should be even lower in tetraploid because of the loss of some genetic information in induction process. However, how the genetic diversity varied in artificial induced tetraploid is still unknown.

In this study, one of our goals was to determine the most effective and repeatable timing of hydrostatic pressure shock for suppression of the mitotic cleavage of the embryo in turbot, an important maricultured fish in China and Europe. Subsequently, tetraploid turbot with high tetraploid rates were produced. Finally, the genetic structure of tetraploid populations was also studied.

# MATERIALS AND METHODS

## Brood Stock Management and Gamete Collection

The turbot broodstocks used in the present study were the cultured stocks introduced from Europe. Turbot broodstocks were cultured in the fish farm of Shenghang Aquatic Science and Technology Co., Ltd., Weihai, China under controlled conditions (photoperiod 16 h light: 8 h dark; temperature, 14 ± 1°C). The eggs were stripped from each female turbot and stored at wet box under dark condition. The eggs were divided into small subsamples before fertilization so that each parameter could be tested with one batch (∼5,000 eggs per batch). Semen was drawn from mature male by the application of gentle pressure to their abdomens and was transferred into 5 mL EP tubes to be stored on ice until required. Semen contaminated with water or urine was discarded.

## Artificial Fertilization

Before artificial fertilization, the quality of sperms and eggs were checked under microscope. The sperm motility was Grades IV–V (Valdebenito et al., 2013). The unfertilized eggs were buoyant, transparent, spherical in shape, and 1 mm in average diameter with a single oil globule. The gametes below standards were discarded. The artificial fertilization procedure was as follows: 0.5 mL sperm was activated using 15–20 mL filtered seawater at 15°C and then was added to plastic beakers each containing 5,000 eggs. After a further 30 s of gentle agitation, 50–100 mL filtered seawater was added and the eggs were left undisturbed.

## Tetraploid Induction

The appropriate conditions for the production of tetraploid turbot were examined by varying the moment of hydrostatic pressure shock after fertilization and by altering the pressure and durations of hydrostatic pressure shock. Hydrostatic pressure induction was carried out in manual hydrostatic pressure chamber. There were three experiments, including five groups in each experiment: (1) to determine the optimal moment of shock induction with a single treatment of pressure shock at 65 MPa for 6 min at 5, 10, 15, 20, or 25 min before the appearance of the cleavage furrow; (2) to determine the appropriate pressure intensity of shock with a single treatment of pressure shock at 15 min before the appearance of the cleavage furrow for 6 min at 55, 60, 65, 70, or 75 MPa; (3) to determine the appropriate pressure duration of shock with a single treatment of pressure shock at 15 min before the appearance of the cleavage furrow at 65 MPa for 4, 5, 6, 7, or 8 min. The preset pressure was achieved in 15 s, and recovered to normal pressure in 30 s after shock. Temperature was constantly monitored throughout the experimentation. After treatments, shocked eggs were acclimated to 15 ± 0.2℃ water for incubation. All experiments were replicated up to five times using egg batches derived from different five males and females, respectively. Eggs without hydrostatic pressure shock were used as diploid control.

Fertilization rate in each group was determined at blastocyst stage by examining ∼200 floating eggs. Hatching rate was assessed at 6 h after hatching. Induction rate was the percentage of tetraploid larvae in hatched larvae.

The FCI and mitotic interval (τ0) were calculated in diploid controls. The FCI was defined as the time at which more than 80% of the zygotes had reached first cleavage, and was calculated according to Hershberger and Hostuttler (2005). The τ0 was also calculated according to Shelton et al. (1997). The FCI and τ0 were calculated and recorded in each group of each experiment.

## Assessment of Ploidy Level

The ploidy of control and treatment groups was determined by cellular DNA content and chromosome preparation. Flow cytometric analysis was performed to detect average cellular DNA contents of larvae in hatched control and treated groups with a PARTEC cell counter analyzer CCA-II (PARTEC, Germany), and 30 hatched larvae were sampled and prepared each group (You et al., 2001; Luckenbach et al., 2004). Diploid larvae were used as a diploid standard for the calibration of the cytometer. Chromosomal metaphases were prepared from embryos at gastrula stage by regular air-dried method after colchicine and hypotonic pretreatments. Slides were stained in 15% Giemsa for 15 min and observed under the light microscope after drying (You et al., 1991).

## Induction of Tetraploid Populations and Embryo Development Observation

According to the results of induction conditions of tetraploid turbot, we carried out a massive tetraploid induction of turbot. About 200 mL eggs acquired from three to four female turbot were fertilized with 5 mL sperms from three to four male ones, and then shocked under the optimal parameters. Then eggs were incubated in net cages under 15 ± 0.2℃. The hatched larvae were reared in indoor tanks with flow-through sea water at a temperature of 18–21℃ and fed with rotifer, artemia, and commercial dry feed (salinity 28–30, pH 7.8–8.2, dissolved oxygen > 6 mg/L, water exchange rate was 50% for fry, 100% for larvae, and >800% for juveniles). Two tetraploid induction (4n1 and 4n2) populations with different parental origin were induced, respectively. The control diploid (2n1 and 2n2) populations were also conducted.

The embryo development was observed under stereoscope. The morphological character and embryonic development time were calculated in each tetraploid induction (4n) and diploid control (2n) population.

## Genetic Structure Analysis

The genetic diversity of 4n induction populations was measured using microsatellite markers. More than 50 new hatched larvae in each 2n or 4n population were sampled and stored in 100% ethanol under −20℃ until DNA extraction. Total genomic DNA was extracted using the rapid salt-extraction of genomic DNA according to Aljanabi and Martinez (1997). Extracted DNA was checked using 0.8% agarose gel electrophoresis and then stored at −20℃ for PCR amplification.

Twenty out of fifty-five previously published microsatellite loci were screened to be analyzed (**Table 1**, Iyengar et al., 2000; Castro et al., 2003; Chen et al., 2007; Pardo et al., 2010). Primers were synthesized by Sangon Biological Engineering Technology & Services Co., Ltd. (Shanghai, China). PCR was conducted in a volume of 25 μL containing 100 ng of template DNA, 0.4 μmol/L of each primer, 1.5 mmol/L of $MgCl_2$, 100 μmol/L of dNTPs, 1 U of Taq polymerase, and 1 × PCR buffer (Promaga). Reactions were processed as follows: an initial denaturation step of 5 min at 94℃, followed by 1 min at 94℃, 30 s at annealing temperature and 45 s at 72℃ for 35 cycles with a final 5 min extension at 72℃. Amplification products (5 μL load) were separated by electrophoresis through a 12% non-denaturing polyacrylamide gel in 1 × TBE buffer. Detection of microsatellite alleles was achieved by silver staining method. DNA fragments were visualized with a modified silver stain method described by Xu et al. (2002). Alleles were designated according to the PCR product size relative to a molecular size marker (M1041 50–500 bp, Dongsheng Biotech, China) in combination with Quantity One software (Bio-Rad). More than 2% PCR products with representative alleles in each population were screened out, and were blindly scored after electrophoresis on the same plate to reduce genotyping errors.

The genetic diversities of the two 4n induction populations were calculated separately and integrally, as well as the two 2n populations. Measurements of genetic diversity were conducted within population: standard genetic diversity parameters, including the number of alleles (*A*), effective alleles (*Ae*), observed heterozygosity (*Ho*), and expected heterozygosity (*He*), were calculated using GENALEX v. 6. Deviations from Linkage and Hardy–Weinberg Equilibrium (HWE) for each microsatellite locus were tested using the package GENEPOP v. 4.0, and the significance was adjusted

**TABLE 1 |** The microsatellite primer and annealing temperature.

| Primer | Sequence 5′-3′ | Annealing temperature/℃ |
|---|---|---|
| F1-0CA19 | F:AGTTACACCAGTGCACAGAG | 56 |
| | R:CCAGGCCATCCACATTTAAC | |
| Smac-01 | F: TGTTGCTTTGCTCCTTTTCC | 56 |
| | R: TAGTGGAACGGCGTCTAGGT | |
| Smac-02 | F: TGTCTCTCGTCAGTGGCAGT | 56 |
| | R: AAACTGCAGCCTCCAAGATG | |
| Smac-05 | F: TTCGAATTCACCAAGTGTCC | 56 |
| | R: GTCAGCCATCATCTCCACCT | |
| Smac-06 | F: GACCCAACGAGCACTGTT | 57 |
| | R: GGGCCAACATCATTATGG | |
| Smac-08 | F: GTACACTTATTGGTGCAAGGC | 58 |
| | R: TTCTGACAGATTTGCTGGCACT | |
| Smac-10 | F: GGTGGCTGGGTAAATCTGTT | 60 |
| | R: CTTCCCTCCGTCTACGCTCC | |
| Smac-11 | F: AACCACTAGCTGGAAATCAGACC | 64 |
| | R: CTTCCCTTCAACTCGGCAAAA | |
| Sma3-129INRA | F:GCACTGCCTTTTCATTGG | 58 |
| | R:CAGCTCTAGATTGTTTATCCC | |
| Sma-USC9 | F:CAAGATGGAGAAGCTGGACTG | 58 |
| | R:GCAGGAAAGAGGGAAGATCG | |
| Sma-USC13 | F:CATTTGTGGCACTTTTAG | 50 |
| | R:CTTTCTGTCAGTCTCATCC | |
| Sma-USC17 | F:TCGCCTGCTATCTGCTTACA | 55 |
| | R:TCGTTCCCACACTTGACTTG | |
| Sma-USC18 | F:TGGGACTGCTTGTGTGTGTT | 55 |
| | R:TCACACTCCTAAATTCCCTCTT | |
| Sma-USC21 | F:TGGGAGAGTGGGACTTTCAG | 55 |
| | R:CGCTCGTCTTTCTTTCCATC | |
| Sma-USC23 | F:CCTGGCACTGTCTGGGCT | 56 |
| | R:CACTGGAGCGGGAATGATG | |
| Sma-USC25 | F:AGCCCACTGCCATGAATAGA | 56 |
| | R:CACAGTTGAAGCACACAGCA | |
| Sma-USC26 | F:CAAACCAACGGACTAACAAACA | 56 |
| | R:TCTTCATTACCAGCCCATCA | |
| Sma-USC27 | F:GCATTACCGCCATCTACTGG | 56 |
| | R:GGTGCAGTTTGAATCTCCTTG | |
| Sma-USC28 | F:CCCGCAGAGACAGAGGTAAA | 58 |
| | R:CGTGTGCAGGATTGTTTGG | |
| Sma-USC30 | F:GTGCTTCTAACACATCTACTGT | 54 |
| | R:GTTCAGACTCGGATTATGTA | |

by applying the sequential Bonferroni correction. Linkage disequilibrium and HWE exact tests were tested using the Markov chain method (10,000 dememorization steps, 100 batches, 5,000 iterations). The Discriminant Analysis of Principal Components (DAPC) was used to cluster genotypes independently of *a priori* haplotype designation using the R package adegenet v. 1.4.2.

## Statistical Analysis

Data from control diploid and tetraploid induction groups were analyzed with the SPSS package for Windows (Version 15.0, SPSS, Chicago, IL, United States). Data are shown as mean ± SD. And the regression analyses of optimal initiation time, τ0 and FCI in nine groups were conducted by linear regression analysis in the SPSS package. Distributions were examined for departures from normality by the Kolmogorov–Smirnov test and the homogeneity of variances was verified by the Levene's test. Significant differences were determined using One-way ANOVA tests followed by Duncan's multiple comparison tests at the probability level of 0.05.

# RESULTS

## The Optimal Conditions of Pressure Shock Induction

The optimal conditions of pressure shock induction were evaluated according to the survival rate of new hatched larvae (hatching rate × tetraploid rate). Overall, the optimal initiation time for pressure shock was 15 min before the first cleavage at 14.8–15.5℃. The optimal treatment pressure and treatment duration were 67.5 MPa and 6 min (**Figure 1**). The initiation time for pressure shock was the most important factor for the survival of tetraploid induction. The peak of survival rate was obtained when the shock started around 15 min (12.5–17.5 min) before the appearance of the cleavage furrow, which was higher than those obtained from shocks applied at earlier or later time. The induction rate was determined by cellular DNA content of hatched larvae and was verified by chromosome preparation at gastrula stage (**Figure 2**). The induction rate ranged from 70 to 100% under the optimal condition.

## The Regression Analysis of Optimal Initiation Time, FCI, and τ0

The variation in FCI and τ0 among the populations tested in this study was fairly large (**Table 2**). On the other hand, the coefficient of variation within the groups was rather small, suggesting homogeneity among the eggs within each group.

The FCI had extremely significant correlation with the optimal initiation time ($P = 0.002$). Meanwhile, there was no significant correlation between τ0 and the optimal initiation time (**Table 3**). Therefore, the factor τ0 could be excluded, and the regression analyses of optimal initiation time and FCI were conducted. The FCI had extremely significant correlation with the optimal initiation time ($P = 0.001$, **Table 3**). A regression equation prediction model was as below:

$$\text{The optimal initiation time} = 0.982\ \text{FCI} - 12.182.$$

Or

$$\text{The optimal initiation time} = 0.85\ \text{FCI}.$$

## Tetraploid Induction Population Establishment and Embryo Development Observation

According to the optimal induction conditions of tetraploid turbot and the prediction model of optimal initiation time

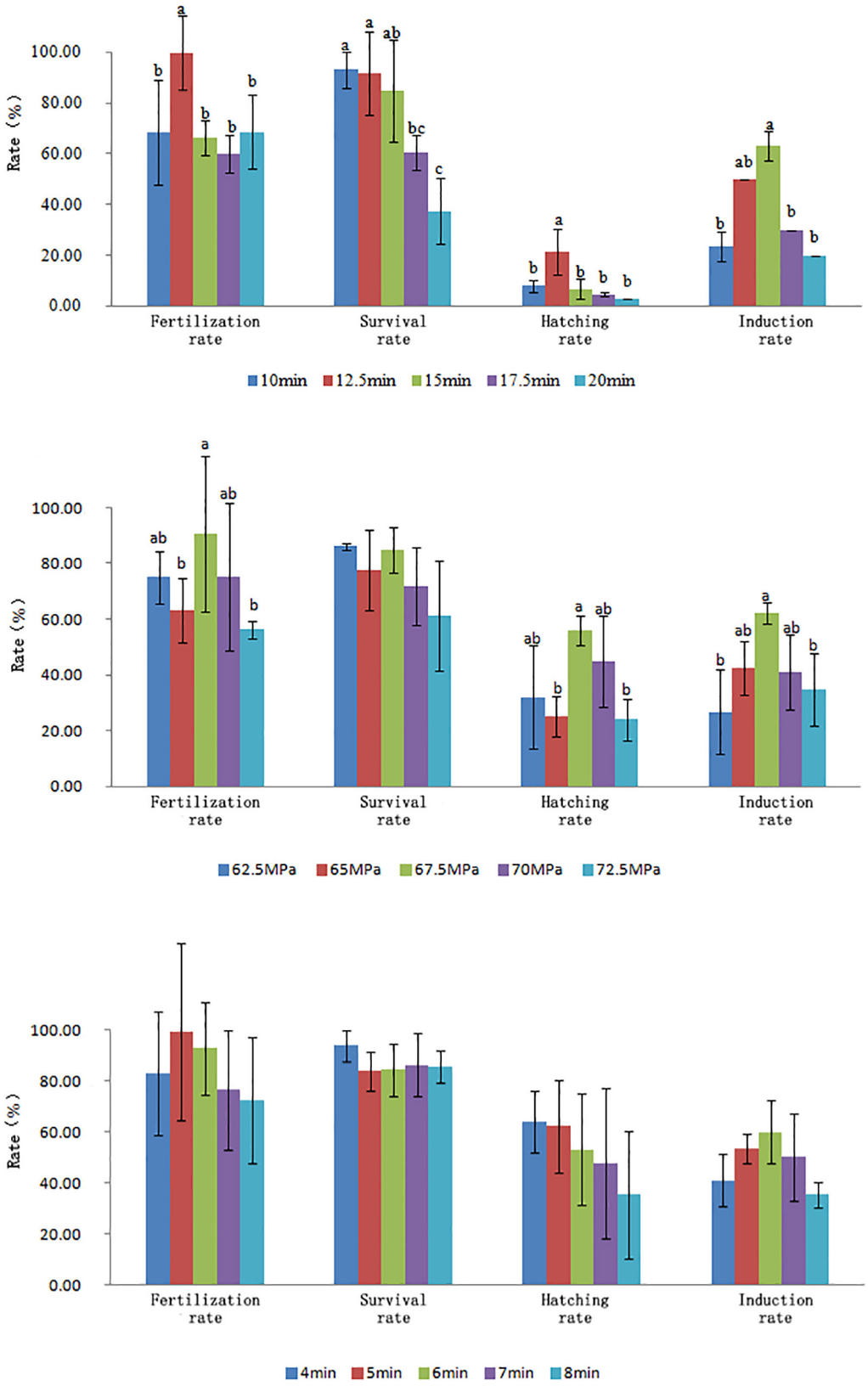


**FIGURE 1 |** Effects of initiation treatment time, treatment pressure, and treatment duration on fertilization rate, survival rate at gastrula stage, hatching rate, and induction rate.

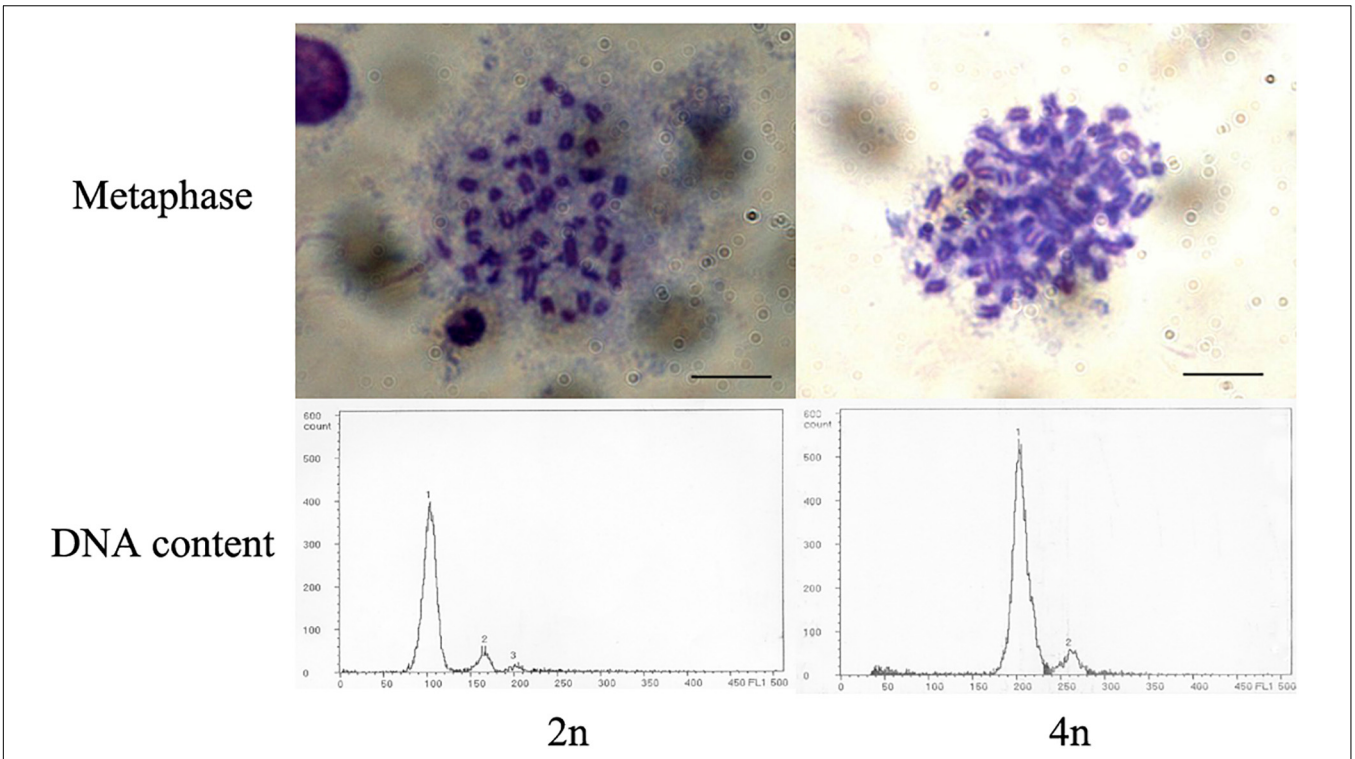


**FIGURE 2 |** The polyploid determination by chromosome preparation (above) and FCM (below).

**TABLE 2 |** The optimal initiation time, FCI, and τ0 in nine different tetraploid induction (4n) groups.

| No. | Water temp/℃ | Optimal initiation time/mpf | FCI/min | FCI/% | τ 0/min |
|---|---|---|---|---|---|
| 1 | 15.0 | 86.0 | 96.0 | 89.6 | 56.0 |
| 2 | 15.0 | 82.5 | 95.0 | 86.8 | 58.0 |
| 3 | 15.4 | 81.5 | 94.0 | 86.7 | 60.0 |
| 4 | 15.1 | 76.0 | 91.0 | 83.5 | 58.3 |
| 5 | 15.1 | 81.5 | 96.5 | 84.5 | 60.0 |
| 6 | 15.0 | 72.5 | 86.5 | 83.8 | 50.0 |
| 7 | 14.9 | 83.5 | 99.5 | 83.9 | 62.0 |
| 8 | 15.0 | 77.0 | 92.0 | 83.7 | 62.5 |
| 9 | 14.8 | 82.0 | 97.0 | 84.5 | 64.0 |

described above, we carried out a massive tetraploid induction of turbot, and obtained two 4n induction populations. The ploidy levels in these two populations were tested by FCM at 1 dph. The 1 dph diploid larvae were also tested as control. The tetraploid rates in both 4n induction populations were 100%. The observation of embryo development showed that the 2n embryos took about 100 h to hatch under 15℃. Hydrostatic pressure shock induced higher deformity rate of embryos which could not survive to hatching stage. The rest embryos showed no difference in morphology with diploid embryo and larvae, and had a prolonged hatching period (110 h). About 8 g (7,200–8,800) embryos in each population were moved into a 1.5 $m^3$ indoor tank before hatching. The hatching rates in 4n1 and 4n2 induction populations were 70 and 75%, respectively. The hatching rates in 2n1 and 2n2 populations were slightly higher (81 and 88%, respectively). In embryo development, two mortality peaks were observed at stages gastrula and hatching in 4n induction populations. Some embryos in 2n populations also died in these two stages, but the mortality rates were much lower (**Figure 3**).

## Genetic Structure of Tetraploid Induction Populations

The population genetic analysis with 20 microsatellite markers showed that two to four alleles in each locus were respectively detected in 2n and 4n populations (**Figure 4**). Private alleles were only appeared at locus Sma-USC21, with low allele frequencies. Two alleles in locus Sma-USC21 lost in 4n induction populations. The genetic diversity indexes were shown in **Table 4**. Eleven and fourteen loci in 2n and 4n populations showed a negative genetic deviation index (*D*). The numbers of locus deviating from Hardy–Weinberg equilibrium in 2n1, 4n1, 2n2, and 4n2 populations were 6, 9, 12, and 7, respectively. Overall, 12 loci in 2n and 4n populations, respectively deviated from Hardy–Weinberg equilibrium (**Table 5**).

The scatter plots from the DAPC clearly showed four major clusters. In the 3D-FCA, the plots of 4n1 scatted more widely than those of 2n1 population, meanwhile plots of 4n2 showed narrower distribution range than those of 2n2 population (**Figure 5**).

**TABLE 3 |** The coefficient of regression analysis of optimal initiation time, FCI, and τ 0.

| Model | | Unstandardized coefficients | | Standardized coefficients | *t* | Sig. |
|---|---|---|---|---|---|---|
| | | *B* | SE | Beta | | |
| 1 | (Constant) | −16.224 | 15.786 | | −1.028 | 0.344 |
| | FCI | 1.242 | 0.230 | 1.133 | 5.408 | 0.002 |
| | τ0 | −0.347 | 0.212 | −0.342 | −1.633 | 0.154 |
| 2 | (Constant) | −12.182 | 17.349 | | −0.702 | 0.505 |
| | FCI | 0.982 | 0.184 | 0.896 | 5.334 | 0.001 |

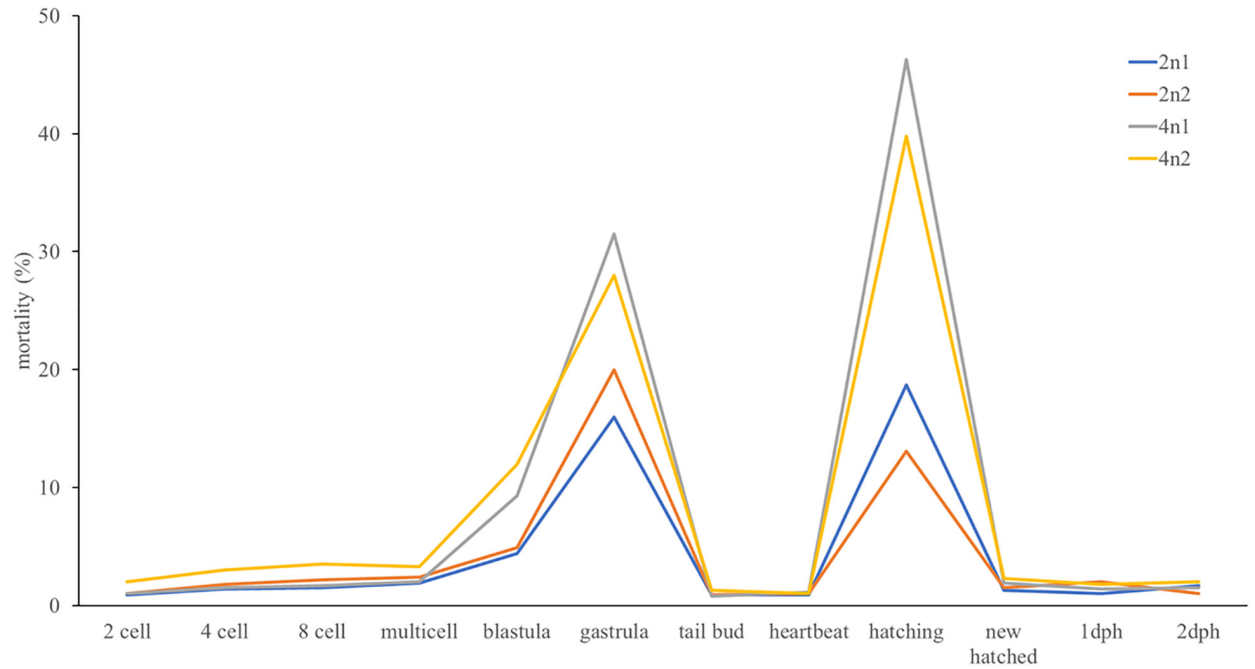


**FIGURE 3 |** The mortality of diploid (2n) and tetraploid induction (4n) populations.

## DISCUSSION

Tetraploid could benefit in triploid breeding, because it can avoid the inductive effect in artificial triploid induction. However, tetraploid-derived triploids were only acquired in a few fish, such as rainbow trout (Chourrout et al., 1986; Weber et al., 2014) and Atlantic salmon (*Salmo salar*) (Benfey, 2016). In some ancient tetraploid fresh water fish, triploids could also be acquired by distant hybridization, such as female Japanese crucian carp (*Carassius cuvieri*) and male blunt snout bream (*Megalobrama amblycephala*) (Hu et al., 2018), and female grass carp (*Ctenopharyngodon idellus*) and male topmouth culter (*Erythroculter ilishaeformis*) (Wu et al., 2019). In turbot and other marine fish, no report was found about the tetraploid-derived triploids. The main reason is the difficulty in acquirement of tetraploid parent fish because of the low survival and unstable ploidy of artificial induced tetraploid.

Hydrostatic treatment is the most frequently used method of tetraploid induction in marine fish. Usually, the effect of hydrostatic treatment is mainly affected by three factors: initiate treatment time, treatment pressure intensity and treatment duration (Peruzzi and Chatain, 2003). The optimal values of these parameters differ greatly among species. Especially the initiate treatment time could range from 6 to 70 min after fertilization (Peruzzi and Chatain, 2003; Li et al., 2012; Yi et al., 2012). Our results showed a longer optimal initiate treatment time (about 85 min after fertilization) in turbot, which may concern with the longer FCI. There were significant differences in fertilization rate, hatching rate and induction rate among induction groups at different initiate time and pressure ($P < 0.05$); but there were no significant differences at different treatment time, which indicated that the effect of treatment time on the experiment results was small, and the determination of appropriate treatment time and pressure were more important. Hydrostatic pressure doubles the chromosomes to induce tetraploid fish by inhibiting cleavage of fertilized eggs (Zhou and Gui, 2017).

Generally, before the first cleavage, the low synchronization of fertilized eggs, the complexity of cleavage regulation and

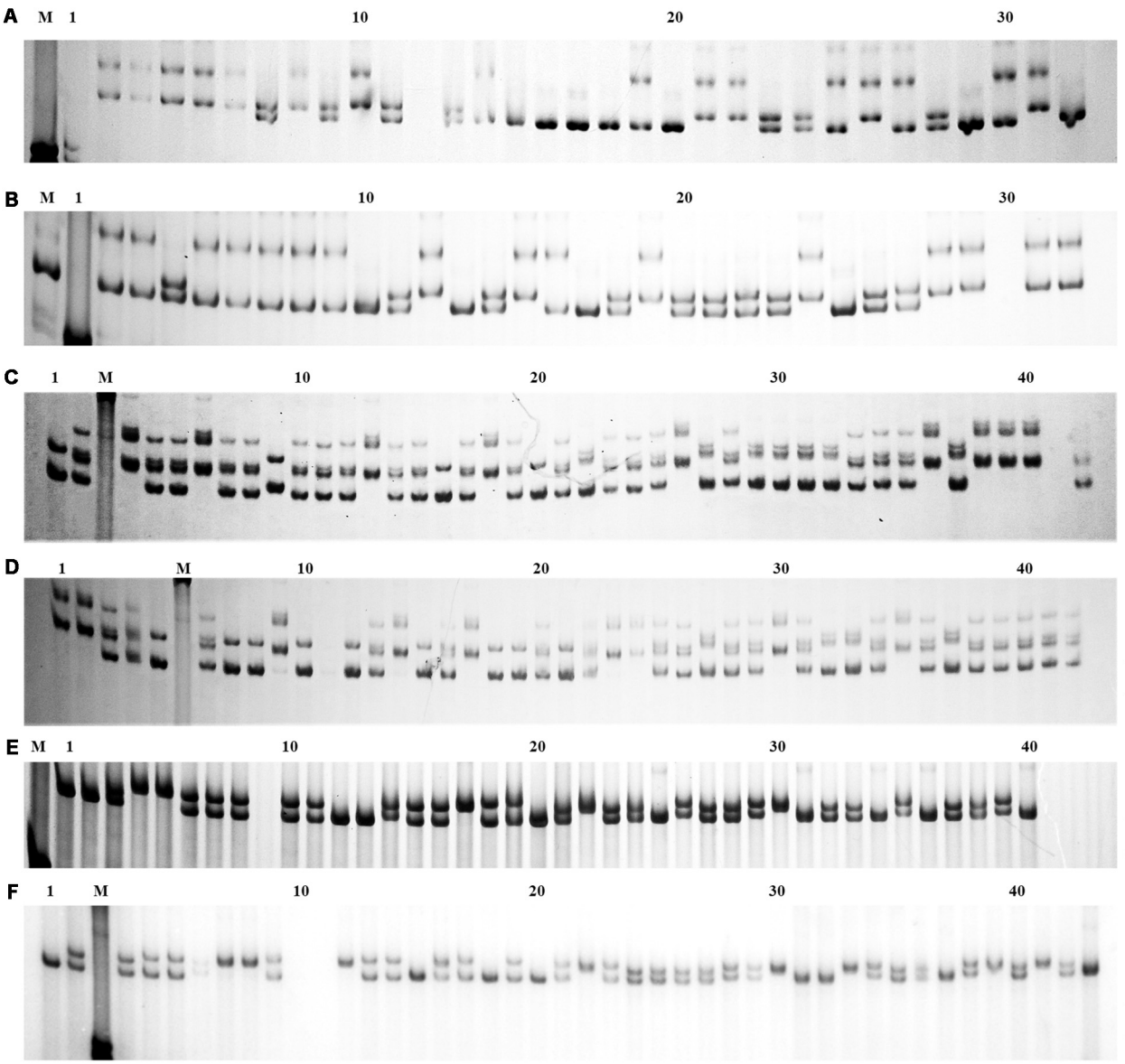


**FIGURE 4 |** The electrophoresis patterns of some primers in 2n and 4n populations. M, DNA 2000 plus maker. **(A)** Locus Smac-05 in 2n populations; **(B)** Locus Smac-05 in 4n populations; **(C)** Locus Sma3-129INRA in 2n populations; **(D)** Locus Sma3-129INRA in 4n populations; **(E)** Locus Sma-USC23 in 2n populations; **(F)** Locus Sma-USC23 in 4n populations.

environmental changes will affect the development of fertilized eggs, making it difficult to determine the optimal treatment time (Jeuthe et al., 2016). If treated in advance, the fertilized eggs will be shocked before they reach the expected level, which greatly reduces the induction efficiency. If the treatment is postponed, the fertilized eggs will begin to cleavage, which will affect their inhibition effect on cleavage, leading to the generation of aneuploidy and an increase in mortality (Sakao et al., 2006). In salmonids, early studies suggested that small changes in the moment of shock induction could have large effects on success of tetraploid induction (Chourrout, 1984; Chourrout et al., 1986; Zhang et al., 2005). Hershberger and Hostuttler (2007) found that applying pressure 15 min late could decrease the tetraploid induction rate from 100 to 0%. It was also found that the variation in embryo development significantly affected the tetraploid induction in terms of both tetraploid induction rate and viability of tetraploid progeny (Weber and Hostuttler, 2012). Thereafter, the FCI index was used to alleviate the effect of asynchronism in embryo development. And the induction success rate was increased (Weber and Hostuttler, 2012). The temperature-dependent measure of $\tau 0$ can also help to standardize chromosome manipulation in fish eggs (Shelton et al., 1997). It is the duration of one mitotic cycle during early embryonic cleavage stage when the cleavage is synchronous.

**TABLE 4 |** The genetic diversity index of diploid (2n) and 4n populations.

| Locus | 2n | | | | | 4n | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *A* | *Ae* | *Ho* | *He* | *D* | *A* | *Ae* | *Ho* | *He* | *D* |
| F1-0CA19 | 2 | 1.980 | 0.450 | 0.501 | 0.091 | 2 | 1.982 | 0.524 | 0.501 | −0.057 |
| Smac-01 | 2 | 1.904 | 0.725 | 0.481 | −0.527 | 2 | 1.999 | 0.738 | 0.506 | −0.477 |
| Smac-02 | 2 | 1.923 | 0.450 | 0.486 | 0.062 | 2 | 1.747 | 0.619 | 0.433 | −0.448 |
| Smac-05 | 2 | 1.987 | 0.270 | 0.504 | 0.456 | 2 | 1.968 | 0.308 | 0.498 | 0.374 |
| Smac-06 | 2 | 1.835 | 0.700 | 0.461 | −0.538 | 2 | 1.426 | 0.366 | 0.303 | −0.224 |
| Smac-08 | 2 | 1.632 | 0.475 | 0.392 | −0.227 | 2 | 1.536 | 0.450 | 0.353 | −0.290 |
| Smac-10 | 2 | 1.663 | 0.550 | 0.404 | −0.379 | 2 | 1.660 | 0.548 | 0.402 | −0.377 |
| Smac-11 | 2 | 1.988 | 0.564 | 0.503 | −0.135 | 2 | 1.980 | 0.600 | 0.501 | −0.212 |
| Sma3-129INRA | 2 | 1.999 | 0.525 | 0.506 | −0.051 | 2 | 1.981 | 0.463 | 0.501 | 0.064 |
| Sma-usc09 | 3 | 2.888 | 0.579 | 0.662 | 0.114 | 3 | 2.795 | 0.463 | 0.650 | 0.278 |
| Sma-usc13 | 3 | 2.796 | 0.658 | 0.651 | −0.024 | 3 | 2.832 | 0.707 | 0.655 | −0.093 |
| Sma-usc17 | 2 | 1.904 | 0.225 | 0.481 | 0.526 | 2 | 1.908 | 0.244 | 0.482 | 0.488 |
| Sma-usc18 | 3 | 2.456 | 0.675 | 0.600 | −0.139 | 3 | 2.292 | 0.610 | 0.571 | −0.082 |
| Sma-usc21 | 4 | 2.102 | 1.000 | 0.531 | −0.907 | 2 | 2.000 | 1.000 | 0.506 | −1.000 |
| Sma-usc23 | 4 | 3.233 | 0.641 | 0.700 | 0.072 | 4 | 3.666 | 0.625 | 0.736 | 0.141 |
| Sma-usc25 | 2 | 1.838 | 0.378 | 0.462 | 0.170 | 2 | 1.521 | 0.439 | 0.347 | −0.281 |
| Sma-usc26 | 3 | 2.791 | 0.641 | 0.650 | 0.001 | 3 | 2.779 | 0.756 | 0.648 | −0.181 |
| Sma-usc27 | 4 | 3.399 | 0.872 | 0.715 | −0.235 | 4 | 3.816 | 0.902 | 0.747 | −0.223 |
| Sma-usc28 | 3 | 2.817 | 0.450 | 0.653 | 0.302 | 3 | 2.894 | 0.524 | 0.662 | 0.200 |
| Sma-usc30 | 2 | 1.257 | 0.231 | 0.207 | −0.130 | 2 | 1.584 | 0.488 | 0.373 | −0.323 |
| Mean | 2.550 | 2.219 | 0.553 | 0.527 | −0.075 | 2.450 | 2.218 | 0.569 | 0.519 | −0.136 |

**TABLE 5 |** Hardy–Weinberg equilibrium in 2n and 4n populations.

| Locus | Population | | | | | |
|---|---|---|---|---|---|---|
| | 2n1 | 4n1 | 2n2 | 4n2 | 2n-total | 4n-total |
| F1-0CA19 | ns | ns | ns | ns | ns | ns |
| Smac-01 | ns | *** | *** | ns | *** | ** |
| Smac-02 | ns | ** | ns | ns | ns | ** |
| Smac-05 | ns | ns | * | ns | ** | * |
| Smac-06 | ns | ns | ** | ns | *** | ns |
| Smac-08 | ns | ns | ns | ns | ns | ns |
| Smac-10 | ns | ns | * | * | * | * |
| Smac-11 | ns | ns | * | ns | ns | ns |
| Sma3-129INRA | ns | ns | ns | ns | ns | ns |
| Sma-usc09 | ns | ** | * | * | ** | *** |
| Sma-usc13 | * | ns | * | ns | ** | ns |
| Sma-usc17 | ns | *** | *** | ns | *** | ** |
| Sma-usc18 | ns | ns | ns | ns | ns | ns |
| Sma-usc21 | *** | *** | *** | *** | *** | *** |
| Sma-usc23 | ** | ** | *** | *** | ** | ** |
| Sma-usc25 | ns | ns | ns | ns | ns | ns |
| Sma-usc26 | * | * | ns | * | * | ** |
| Sma-usc27 | * | *** | ** | ** | ** | *** |
| Sma-usc28 | ** | ** | * | *** | *** | *** |
| Sma-usc30 | ns | ns | ns | ns | ns | * |

*ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001.*

τ0 is affected by temperatures, and is also a reliable indicator of developmental rates of egg incubation. So it can be used to estimate the optimal times for chromosome manipulation. Both the FCI and τ0 were widely used in Salmonids, paddlefish (*Polyodon spathula*) and shovelnose sturgeon (*Scaphirhynchus platorynchus*) (Shelton et al., 1997). It is mainly because the optimal moment of treatment in fish is long, and the embryo development is hard to find out. In flatfish, the embryo development was far shorter than that in Salmonids, so it was believed that the deviation from the actual initial time to the optimal initial time was small. The optimal initiate time of treatment was conducted as minutes after the fertilized time or before the appearance of the first cleavage furrow (Lin et al., 2016). In the present study, we found that FCI and τ0 varied in a fairly large range for the optimal initiation time for pressure shock (**Table 2**). These indexes could be affected by the water temperature and genetic background. Due to the buoyant trait of turbot eggs, they may also be affected by the air temperature. Besides, our previous study showed that about 5 min derived from the optimal moment of shock induction would sharply decrease the induction rate (Zhu et al., 2017). Therefore, the relatively short embryo development also shortens the time window of initiation time for success tetraploid induction. It needs to determine the initiation time more precisely. In order to determine the treatment time accurately and reduce the genetic and environment influences, the relation among optimal initiate time of treatment, FCI and τ0 was calculated. We found that the FCI is the best index to determine the optimal initiate time of treatment in turbot. In this way, the influences of genetic origin and environmental factors such as water temperature and air temperature can be reduced. Under the optimal conditions, even fully tetraploid population could be induced, which is far better than the original method.

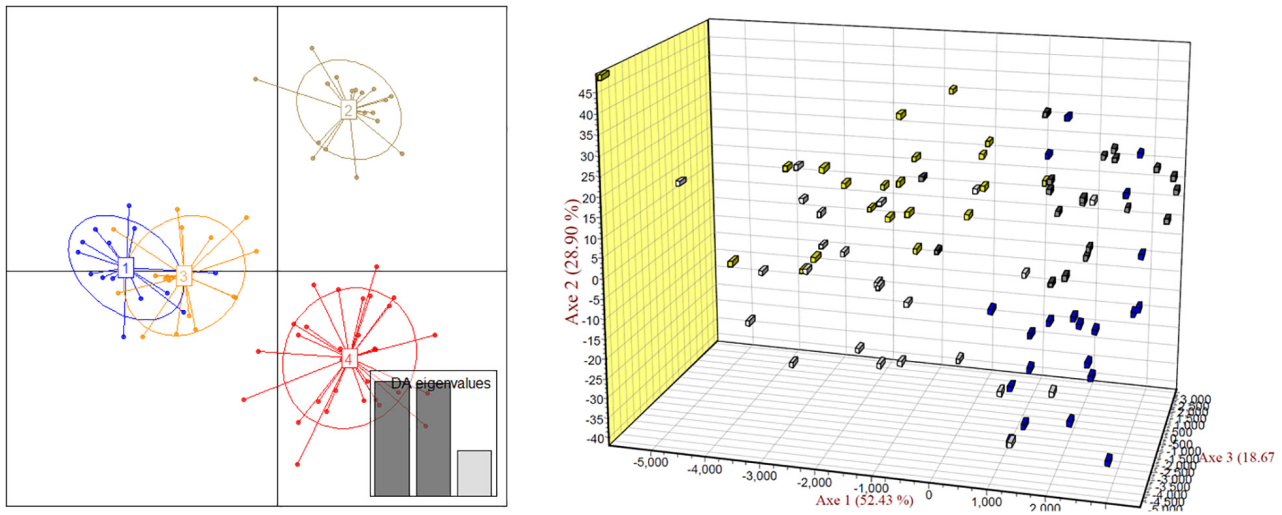


**FIGURE 5 |** Discriminant analysis of principal components (DAPC, left) and three-dimensional factorial correspondence analysis (3D-FCA, right) showing the relationships among turbot individuals of two 2n and two 4n populations. Yellow block, 2n1; white block, 4n1; blue block, 2n2; gray block, 4n2.

In present study, we also tried to find out how the artificial tetraploid induction affected the genetic diversity in turbot. The genetic structure of natural or spontaneous tetraploid fish, such as Amazon molly (*Poecilia formosa*) (Lampert et al., 2008) and minnow (*Squalius alburnoides*) (Crespo-López et al., 2007), has been reported before. The number of alleles and heterozygosity were higher in tetraploid populations than those in diploid populations in these fishes due to their hybrid origin. However, the genetic structure of artificial induced tetraploids was different. Overall, the number of effective alleles in the 2n population of turbot was basically the same as that in the 4n induction population. The observed heterozygosity was slightly lower and the expected heterozygosity was slightly higher in 2n population than those in the 4n induction population. The average observed heterozygosity *Ho* of 2n and 4n populations were 0.553 and 0.569, respectively. The average expected heterozygosity *He* of 2n and 4n populations were 0.527 and 0.519, respectively. Yu et al. (2009) analyzed seven artificial breeding families of turbot, and the average heterozygosity was 0.7206. Gu et al. (2009) used self-developed microsatellite markers to analyze the expected heterozygosity of 31 turbot individuals ranging from 0.5061 to 0.8995. It was found that the heterozygosity of turbot populations in the present study were lower than those has been reported. The screened out microsatellite markers were reported to have more alleles in the literature, but the number of alleles in this study is only 2–4. Heterozygote deletion was also found at many loci. It indicates that inbreeding may occur in parent turbot breeding, and it is necessary to introduce parents from different sources to improve their genetic diversity. By comparing the allele frequencies of turbot, two alleles of Sma-USC21 were deleted in the 4n population, suggesting the existence of recessive lethal genes. Tetraploid induction population showed lower heterozygosity and higher heterozygote deletion. We also, respectively, analyzed the genetic diversity of two 4n populations, and the scatter plots from the DAPC clearly showed four major clusters. In the 3D-FCA, the plots of 4n1 population scatted more widely than those of 2n1 population, meanwhile plots of 4n2 population showed narrower distribution range than those of 2n2 population. Since the 4n population was obtained by inhibiting the cleavage doubling of the 2n population, its genetic diversity level should theoretically be the same as that in the 2n population (Ferris and Whitt, 1980). Besides, this induction process can be regarded as artificial selection of population. Some individuals with specific alleles or genotypes are sensitive to high-pressure induction and cannot survive, resulting in the deletion or frequency reduction of the allele. Such phenomena could also be found in artificial induced triploid (Liu et al., 2018). Therefore, it is easy to explain the decrease of genetic diversity in 4n2 population compared to that in 2n2 population. However, it is hard to explain the wider expand of scatter of 4n1 population than that of 2n1 population. The 4n1 and 4n2 induction populations came from different parents. Weather the genetic origin affected the genetic diversity in tetraploid induction is still unknown. Further analysis is needed in combination with genotypes.

In conclusion, the tetraploid induction rate of turbot can be increased to be more than 90% or even 100% under the optimized conditions by using FCI. The optimal initiation time is 0.982 FCI − 12.182 or the optimal initiation time is 0.85 FCI. And the optimal treatment pressure and treatment duration were 67.5 MPa and 6 min. Then two tetraploid induction populations were obtained under the optimal conditions. They showed lower heterozygosity and higher heterozygote deletion. Besides, it seems that the hydrostatic pressure has different effects on turbot with different genetic origin. The exact mechanism still needs further study. The results would help in tetraploid induction and relatively studies in marine fish.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/supplementary files.

## ETHICS STATEMENT

The animal study was reviewed and approved by Institutional Animal Care and Use Committee of the Institute of Oceanology, Chinese Academy of Sciences.

## AUTHOR CONTRIBUTIONS

ZW and FY designed the study and wrote the manuscript. ZW, LW, YL, XZ, and XY conducted the field work. ZW generated the data. LW performed the data analyses. All authors read, commented, and agreed on the manuscript.

## FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31502156 and 41276171), the Key Research and Development Plan of Shandong Province (No. 2017CXGC0101), and the National Key R&D Program of China (No. 2018YFD0901202).

## ACKNOWLEDGMENTS

We are grateful to Mr. Yan Li for his help in the laboratory.

## REFERENCES

Aljanabi, S., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality gnomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Amores, A., Suzuki, T., Yan, Y. L., Pomeroy, J., Singer, A., Amemiya, C., et al. (2004). Developmental roles of pufferfish hox clusters and genome evolution in ray-fin fish. *Genome Res.* 14, 1–10. doi: 10.1101/gr.1717804

Benfey, T. J. (2016). Effectiveness of triploidy as a management tool for reproductive containment of farmed fish: atlantic salmon (*Salmo salar*) as a case study. *Aquaculture* 8, 264–282. doi: 10.1111/raq.12092

Bidwell, C. A., Chrisman, C. L., and Libey, G. S. (1985). Polyploidy induced by heat-shock in channel catfish. *Aquaculture* 51, 25–32. doi: 10.1016/0044-8486(85)90237-6

Blanc, J. M., Poisson, H., Escaffre, A. M., Aguirre, P., and Vallee, F. (1993). Inheritance of fertilizing ability in male tetraploid rainbow-trout (*Oncorhynchus mykiss*). *Aquaculture* 110, 61–70. doi: 10.1016/0044-8486(93)90434-Z

Castro, J., Bouza, C., Sánchez, L., Cal, R. M., Piferrer, F., and Martínez, P. (2003). Gynogenesis assessment using microsatellite genetic markers in turbot (*Scophthalmus maximus*). *Mar. Biotechnol.* 5, 584–592. doi: 10.1007/s10126-003-0004-x

Chen, S. L., Ma, H. Y., Jiang, Y., Liao, X. L., and Meng, L. (2007). Isolation and characterization of polymorphic microsatellite loci from an EST library of turbot (*Scophthalmus maximus*) and cross-species amplification. *Mol. Ecol. Resour.* 5, 215–217. doi: 10.1111/j.1471-8286.2007.01725.x

Cherfas, N. B., Gomelsky, B., Peretz, Y., Bendom, N., Hulata, G., and Moav, B. (1993). Induced gynogenesis and polyploidy in the Israeli common carp line Dor-70. *Isr. J. Aquac.* 45, 59–72.

Chourrout, D. (1984). Pressure-induced retention of second polar body and suppression of first cleavage in rainbow trout: production of all-triploids, all-tetraploids, and heterozygous and homozygous diploid gynogenetics. *Aquaculture* 36, 111–126. doi: 10.1016/0044-8486(84)90058-9

Chourrout, D., Chevassus, B., Krieg, F., Happe, A., Burger, G., and Renard, P. (1986). Production of second generation triploid and tetraploid rainbow trout by mating tetraploid males and diploid females - Potential of tetraploid fish. *Theor. Appl. Genet.* 72, 193–206. doi: 10.1007/Bf00266992

Crespo-López, M. E., Pala, I., Duarte, T. L., Dowling, T. E., and Coelho, M. M. (2007). Genetic structure of the diploid-polyploid fish *Squalius alburnoides* in southern Iberian basins Tejo and Guadiana, based on microsatellites. *J. Fish Biol.* 71, 423–436. doi: 10.1111/j.1095-8649.2007.01688.x

Don, J., and Avtalion, R. R. (1988). Production of viable tetraploid tilapias using the cold shock technique. *Isr. J. Aquac.* 40, 17–21.

Ferris, S. D., and Whitt, G. S. (1980). Genetic variability in species with extensive gene duplication: the tetraploid catostomid fishes. *Am. Nat.* 115, 650–666. doi: 10.1086/283590

Gu, Y., Guo, Z. B., Gu, J. J., Mao, R. X., Lu, C. Y., and Sun, X. W. (2009). Isolation of novel microsatellite markers from genome library and SSR polymorphic loci identification in turbot *Scophthalmus maximus*. *J. Dalian Fish. University* 24, 366–370.

Gui, J. F., Liang, S. C., Zhu, L. F., and Jiang, Y. G. (1993). Discovery of 2 different reproductive development modes of the eggs of artificial multiple tetraploid allogynogenetic silver crucian carp. *Chin. Sci. Bull.* 38, 332–341. doi: 10.1016/0167-8809(93)90098-A

Hershberger, W. K., and Hostuttler, M. A. (2005). Variation in time to first cleavage in rainbow trout *Oncorhynchus mykiss* embryos: a major factor in induction of tetraploids. *J. World Aquac. Soc.* 36, 96–102. doi: 10.1111/j.1749-7345.2005.tb00135.x

Hershberger, W. K., and Hostuttler, M. A. (2007). Protocols for more effective induction of tetraploid rainbow trout. *North Am. J. Aquac.* 69, 367–372. doi: 10.1577/a06-022.1

Hu, F., Wu, C., Zhou, Y., Cao, L., Xiao, J., Wang, S., et al. (2018). Production of androgenetic, triploid and tetraploid hybrids from the interspecific hybridization of female Japanese crucian carp and male blunt snout bream. *Aquaculture* 491, 50–58. doi: 10.1016/j.aquaculture.2018.03.014

Iyengar, A., Piyapattanakorn, S., Heipel, D. A., Stone, D. M., Howell, B. R., Child, A. R., et al. (2000). A suite of highly polymorphic microsatellite markers in turbot (*Scophthalmus maximus* L.) with potential for use across several flatfish species. *Mol. Ecol.* 9, 368–371. doi: 10.1046/j.1365-294x.2000.00874-3.x

Jeuthe, H., Brännäs, E., and Nilsson, J. (2016). Effects of variable egg incubation temperatures on the embryonic development in Arctic charr *Salvelinus alpinus*. *Aquacu. Res.* 47, 3753–3764. doi: 10.1111/are.12825

Lampert, K., Lamatsch, D. K., Fischer, P., and Schartl, M. (2008). A tetraploid amazon molly. Poecilia formosa. *J. Heredity* 99, 223–226. doi: 10.1093/jhered/esm102

Li, W. L., Chen, S. L., Ji, X. S., Xie, M. S., Xu, Y., and Deng, H. (2012). Induction and identification of tetraploid fry in *Cynoglossus semilaevis*. *J. Fish. Sci. China* 19, 196–201. doi: 10.3724/SP.J.1231.2011.17344

Lin, Z., Zhu, X., Zhang, T., You, F., Wu, Z., and Cao, Y. (2016). Effects of hydrostatic pressure on microtubule organization and nucleus changes in gynogenetically activated eggs of olive flounder (*Paralichthys olivaceus*). *Theriogenology* 85, 1610–1624. doi: 10.1016/j.theriogenology.2016.01.020

Liu, H., Wu, Z., Zhu, X., Song, Z., Hu, J., Wang, L., et al. (2018). Comparative performance of growth, vertebral structure and muscle composition in diploid and triploid *Paralichthys olivaceus*. *J. Fish Dis.* 41, 1495–1504. doi: 10.1111/jfd.12846

Luckenbach, J. A., Godwin, J., Daniels, H. V., Beasley, J. M., Sullivan, C. V., and Borski, R. J. (2004). Induction of diploid gynogenesis in southern flounder (*Paralichthys lethostigma*) with homologous and heterologous sperm. *Aquaculture* 237, 499–516. doi: 10.1016/j.aquaculture.2004.05.005

Malison, J. A., Kayes, T. B., Held, J. A., Barry, T. P., and Amundson, C. H. (1993). Manipulation of ploidy in yellow perch (*Perca flavescens*) by heat-shock,

hydrostatic-pressure shock, and spermatozoa inactivation. *Aquaculture* 110, 229–242. doi: 10.1016/0044-8486(93)90371-5

Myers, J. M., and Hershberger, W. K. (1991). Early growth and survival of heat-shocked and tetraploid-derived triploid rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 96, 97–107. doi: 10.1016/0044-8486(91)90142-T

Myers, J. M., Hershberger, W. K., and Iwamoto, R. N. (1986). The induction of tetraploidy in Salmonids. *J. World Aquac. Soc.* 17, 1–7. doi: 10.1111/j.1749-7345.1986.tb00546.x

Pardo, B. G., Hermida, M., Fernández, C., Bouza, C., Pérez, M., Llavona, A., et al. (2010). A set of highly polymorphic microsatellites useful for kinship and population analysis in turbot (*Scophthalmus maximus* L.). *Aquac. Res.* 37, 1578–1582. doi: 10.1111/j.1365-2109.2006.01600.x

Pasquier, J., Braasch, I., Batzel, P., Cabau, C., Montfort, J., Nguyen, T., et al. (2017). Evolution of gene expression after whole-genome duplication: new insights from the spotted gar genome. *J. Exp. Zool. Part B Mol. Dev. Evol.* 328, 709–721. doi: 10.1002/jez.b.22770

Peruzzi, S., and Chatain, B. (2003). Induction of tetraploid gynogenesis in the European sea bass (*Dicentrarchus labrax* L.). *Genetica* 119, 225–228. doi: 10.1023/a:1026077405294

Rothbard, S., Shelton, W. L., Kulikovsky, Z., Rubinshtein, I., Hagani, Y., and Moav, B. (1997). Chromosome set manipulations in the black carp. *Aquac. Int.* 5, 51–64. doi: 10.1007/Bf02764787

Sakao, S., Fujimoto, T., Kimura, S., Yamaha, E., and Arai, K. (2006). Drastic mortality in tetraploid induction results from the elevation of ploidy in masu salmon *Oncorhynchus masou*. *Aquaculture* 252, 147–160. doi: 10.1016/j.aquaculture.2005.06.048

Shelton, W., Mims, S., Clark, J. A., Hiott, A. E., and Wang, C. (1997). A temperature-dependent index of mitotic interval (τ0) for chromosome manipulation in paddlefish and shovelnose sturgeon. *Progress. Fish* 59, 229–234. doi: 10.1577/1548-8640(1997)059<0229:atdiom>2.3.co;2

Thorgaard, G. H., Jazwin, M. E., and Stier, A. R. (1981). Polyploidy induced by heat shock in rainbow trout. *Trans. Am. Fish. Soc.* 110, 546–550. doi: 10.1577/1548-8659(1981)110<546:pibhsi>2.0.co;2

Valdebenito, I., Gallegos, P. C., and Roldán, B. (2013). Gamete quality in fish: evaluation parameters and determining factors. *Zygote* 23, 1–21. doi: 10.1017/S0967199413000506

Weber, G. M., and Hostuttler, M. A. (2012). Factors affecting the first cleavage interval and effects of parental generation on tetraploid production in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 34, 231–238. doi: 10.1016/j.aquaculture.2012.03.017

Weber, G. M., Hostuttler, M. A., Cleveland, B. M., and Leeds, T. D. (2014). Growth performance comparison of intercross-triploid, induced triploid, and diploid rainbow trout. *Aquaculture* 433, 85–93. doi: 10.1016/j.aquaculture.2014.06.003

Wu, C., Huang, X., Hu, F., Ouyang, Y., Zhao, L., Wang, S., et al. (2019). Production of diploid gynogenetic grass carp and triploid hybrids derived from the distant hybridization of female grass carp and male topmouth culter. *Aquaculture* 504, 462–470. doi: 10.1016/j.aquaculture.2018.12.056

Wu, Z. H., You, F., Song, Z. C., Hu, J. W., Wang, L. J., Zhu, X. P., et al. (2014). Induction of tetraploid in turbot *Scophthalmus maximus*. *Oceanol. Limnol. Sinica* 45, 657–662.

Xu, S. B., Tao, Y. F., Yang, Z. Q., and Chu, J. Y. (2002). A simple and rapid methods used for silver staining and gel preservation. *Hereditas* 24, 335–336.

Yamaki, M., Satou, H., Taniura, K., and Arai, K. (1999). Progeny of the diploid-tetraploid mosaic amago salmon. *Nippon Suisan Gakkai.* 65, 1084–1089. doi: 10.2331/suisan.65.1084

Yi, Q. L., Yu, H. Y., Wang, X. L., Wang, Z. G., Qi, J., and Zhang, Q. Q. (2012). Production of viable tetraploid olive flounder (*Paralichthys olivaceus*) by hydrostatic pressure shock. *Oceanol. Limnol. Sinica* 43, 382–388.

You, F., Liu, J., Wang, X. C., Xu, Y. L., Huang, R. D., and Zhang, P. J. (2001). Study on embryonic development and early growth of triploid and gynogenetic diploid left-eyed flounder. *Paralichthys olivaceus* (T. et S.). *Chin. J. Oceanol. Limnol.* 19, 147–151. doi: 10.1007/bf02863039

You, F., Sha, X. S., and Ruan, H. C. (1991). Preliminary study on induced triploidy in *Sparus microcephalus* (Basilewsky). *Oceanol. Limnol. Sinica* 22, 489–491.

Yu, F., Wang, W. J., Kong, J., and Ruan, X. H. (2009). The application of microsatellite markers for genealogical identification in selective breeding program of turbot (*Scophthalmus maximus* L.). *Acta Oceanol. Sinica* 31, 127–136.

Zhang, X., Mutsukawa, K., and Onozato, H. (2005). Correlation between delay in the earlier cleavage stage and the tetraploidization rate in rainbow trout *Oncorhynchus mykiss* embryos treated with heat or hydrostatic pressure shock during the first cell cycle. *Fish. Sci.* 71, 239–241. doi: 10.1111/j.1444-2906.2005.00954.x

Zhang, X. L., and Onozato, H. (2004). Hydrostatic pressure treatment during the first mitosis does not suppress the first cleavage but the second one. *Aquaculture* 240, 101–113. doi: 10.1016/j.aquaculture.2004.07.004

Zhou, L., and Gui, J. F. (2017). Natural and artificial polyploids in aquaculture. *Aquacu. Fish.* 2, 103–111. doi: 10.1016/j.aaf.2017.04.003

Zhu, X., Lin, Z., Wu, Z., Li, J., and You, F. (2017). Effect of initiation time of hydrostatic pressure shock on chromosome set doubling of tetraploidization in turbot *Scophthalmus maximus*. *Mar. Biotechnol.* 19, 528–540. doi: 10.1007/s10126-017-9771-7

**Conflict of Interest:** XY was employed by Shenghang Aquatic Science and Technology Co., Ltd., Weihai, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 14 January 2020
doi: 10.3389/fmars.2019.00782

# Genome-Wide Discovery of Single-Nucleotide Polymorphisms and Their Application in Population Genetic Studies in the Endangered Japanese Eel (*Anguilla japonica*)

*Bing-Jian Liu[1,2,3]*, Yu-Long Li[1,3,4], Bai-Dong Zhang[1,3] and Jin-Xian Liu[1,3,4]**

[1] CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2] National Engineering Laboratory of Marine Germplasm Resources Exploration and Utilization, Zhejiang Ocean University, Zhoushan, China, [3] Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, [4] Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China

OPEN ACCESS

**Edited by:**
*Dongdong Xu,*
*Marine Fisheries Research Institute of Zhejiang, China*

**Reviewed by:**
*Xiubao Li,*
*Hainan University, China*
*Junlong Zhang,*
*Institute of Oceanology (CAS), China*

***Correspondence:**
*Bing-Jian Liu*
*bjetbj@163.com*
*Jin-Xian Liu*
*jinxianliu@gmail.com*

**Specialty section:**
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

**Received:** *09 September 2019*
**Accepted:** *04 December 2019*
**Published:** *14 January 2020*

**Citation:**
*Liu B-J, Li Y-L, Zhang B-D and Liu J-X (2020) Genome-Wide Discovery of Single-Nucleotide Polymorphisms and Their Application in Population Genetic Studies in the Endangered Japanese Eel (Anguilla japonica). Front. Mar. Sci. 6:782. doi: 10.3389/fmars.2019.00782*

The Japanese eel (*Anguilla japonica*) is a commercially important aquatic species in East Asia. The number of the Japanese eels has been dramatically declining over the last four decades, and it is now listed as an endangered species (International Union for Conservation of Nature [IUCN] 2014). To manage and conserve this endangered species, it is necessary to assess population genetic diversity, genetic structure, and identify regions of the genomes that are under selection. Here, we generated a catalog of novel single-nucleotide polymorphism (SNP) markers for the Japanese eel using restriction site-associated DNA (RAD) sequencing of 24 individuals from two geographic locations. The 73,557 identified SNPs were widely distributed across the draft genome of the Japanese eel. No genetic differentiation between the two populations was detected based on all loci or neutral loci. However, highly significant genetic differentiation was detected based on loci that appeared to be under selection. BLAST2GO annotations of the outlier SNPs yielded hits for 61 (72%) of 85 significant BLASTX matches. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis identified some of the putative targets of local selection, including genes in several important pathways such as calcium signaling pathway and intestinal immune network for IgA production. This SNP catalog will provide a valuable resource for future population genetic and genomic studies and allows for targeting specific genes and genomic regions under selection in the Japanese eel genome.

**Keywords: population genomics, genetic diversity, restriction site-associated DNA sequencing, single-nucleotide polymorphism discovery, local adaptation**

## INTRODUCTION

Recent advances in next-generation sequencing (NGS) technologies have greatly improved the speed of genome sequencing, facilitating the application of genomic approaches in many research areas, including ecology, conservation, fisheries genetics, and so on (Hudson, 2008; Harismendy et al., 2009; Allendorf et al., 2010; Ekblom and Galindo, 2011; Dudgeon et al., 2012). The high-throughput sequencing methods in NGS can result in extremely large collections of data, which

enables the identification of numerous sensitive markers for a high-resolution genetic analysis, such as single-nucleotide polymorphisms (SNPs) (Ekblom and Galindo, 2011). However, because of the large and complex genomes of research organisms, researchers have typically used complexity-reduction technology for SNP discovery (Slate et al., 2009). One promising approach is restriction site-associated DNA (RAD) sequencing, which reduces genome complexity by sequencing the same loci across the genomes of numerous individuals, enabling comparisons between individuals, and reducing sequencing cost. This approach consists of employing specific restriction enzymes to cleave double-stranded genomic DNA into random fragments and then amplifying and sequencing the regions flanking the restriction enzyme cut sites through specific adapters and NGS platforms, such as Illumina. The paired-end reads from each RAD tag can then be assembled into long contiguous sequences or aligned to a known reference to identify and score 1000s of genetic markers (Miller et al., 2007; Baird et al., 2008; Hohenlohe et al., 2010, 2011). To date, RAD sequencing has been successfully applied to SNP discovery in many species, such as three-spined stickleback (*Gasterosteus aculeatus*) (Baird et al., 2008; Hohenlohe et al., 2010), European eel (*Anguilla anguilla*) (Pujolar et al., 2013), sea mullet (*Mugil cephalus*) (Kruck et al., 2013), and rainbow trout (*Oncorhynchus mykiss*) (Palti et al., 2014). As sequencing cost continues to drop, these methods will have numerous applications in a genetic analysis of complex genomes.

The Japanese eel, *Anguilla japonica*, is an important economical marine fishery species distributed in East Asian countries, which is characterized as a temperate catadromous fish with a complex life cycle and an extensive migratory loop. The spawning area of this species is located in the western Mariana Islands, near 14°–16° N, 142°–143° E (Tsukamoto, 1992, 2006). After being spawned between April and November (Tsukamoto, 1992), leptocephalus larvae disperse via the North Equatorial Current (NEC) and the Kuroshio Current (KC). Four to six months later, they reach the coasts of East Asia and then develop into juvenile Japanese eels (which are typically referred to as "glass eels"). Glass eels then move into continental (freshwater, brackish, or coastal) growth habitats and become yellow eels. After more than four 4 years of feeding, they develop into mature silver eels and then migrate approximately 2,000 to 3,500 km back to the western Mariana Islands, spawn once, and die (Han et al., 2010b). The abundance of Japanese eels has sharply decreased by more than 90% since the 1970s, owing to factors such as commercial overfishing, habitat destruction, pollution, and other environmental changes (Han et al., 2010a). In March 2014, the Japanese eel was listed as an endangered species in the International Union for Conservation of Nature (IUCN) Red List of Threatened Species[1]. Hence, an assessment of population genetic structure, genetic diversity, and regions of the genome that are under selection for Japanese eel is urgently needed in order to better manage and conserve this endangered species.

Despite occupying a broad range of habitats spanning almost the entire coastal area of East Asia, the Japanese eel has been regarded as a panmictic species. The concept of panmictic populations for Japanese eel was originally established by genetic studies focusing on allozyme and mitochondrial DNA (mtDNA) markers (Sang et al., 1994; Ishikawa et al., 2001). In addition, Han et al. (2010a) used eight polymorphic microsatellite DNA loci to investigate the spatial population genetic structures of Japanese eels collected from nine locations in East Asia, resulting in no identified genetic differentiation between different geographical locations. Gong et al. (2014) used mtDNA and microsatellite DNA loci to investigate the population structure of Japanese glass eels collected from 10 locations in China, which provided further evidence for panmixia in Japanese eel. However, significant geographical clines have been detected in this species. For example, Chan et al. (1997) found that *A. japonica* in the western Pacific fringe exhibited clear geographic clines in two allozyme loci. Moreover, Tseng et al. (2006) conducted an analysis of eight microsatellite loci and found that Japanese eel populations could be divided into two major groups.

Owing to the inconclusive and often conflicting results of earlier research, a better understanding of crucial aspects of the population genetic structure in the Japanese eel is still needed. Here, we developed a genome-wide catalog of SNPs for the Japanese eel by employing Illumina HiSeq 2500 paired-end sequencing of RAD tags (Baird et al., 2008). Previous studies in this area either sequenced single-end reads (Hohenlohe et al., 2011; Palti et al., 2014; Guo et al., 2015) or sequenced paired-end reads but only used the single reads (Baird et al., 2008; Wang et al., 2013; Pujolar et al., 2014), which resulted in less accurate alignments. In the present study, we compared the number of SNPs generated from single-end and paired-end data, and we found that paired-end sequencing resulted in more high-quality SNPs. The RAD tags in this study were generated from a total of 24 glass eels from two separate sampling locations, enabling the discovery of novel candidate SNP markers for this widespread, highly fecund marine fish species. The SNPs were then used in a population genetic analysis of two Japanese eel populations, which improved the current understanding of population structure and identified genomic regions that are subject to selection. This improved our understanding of Japanese eel populations, which have significant implications for the sustainable management and utilization of these important resources.

## MATERIALS AND METHODS

### Ethics Statement

All methods were performed in strict accordance with the relevant guidelines and regulations in China.

### Sample Preparation and Restriction Site-Associated DNA Tag Sequencing

Samples of glass eels were collected from two geographically distant locations: one is in Dandong, Liaoning Province (39°58′ N; 124°12′ E), and the other is in Xiapu, Fujian Province (26°55′ N; 120°13′ E). Tissues were kept in 95% ethanol and stored at −80°C. Genomic DNA was extracted from a total of

[1]http://www.iucnredlist.org/details/166184/0

24 individuals (12 from each location) using standard phenol–chloroform extraction. The quality of the DNA samples was checked on 1% agarose gels and measured using a Nanodrop 2000 spectrophotometer and Qubit fluorometric quantitation, in order to ensure that the concentration was more than 25 ng/μl and the total DNA weight was no less than 1 μg. RAD libraries were prepared as described by Zhang et al. (2016).

## Restriction Site-Associated DNA Data Analysis and Single-Nucleotide Polymorphism Identification

Sequencing reads from the Illumina runs were filtered to generate clean reads and then were aligned to the Japanese eel genome draft[2] using BWA version 0.7.12 (Li and Durbin, 2010). The BWA-MEM algorithm was used for mapping with default parameters except for adjusting the seeding length to 32. SNP calling was performed by the SAMTOOLS 0.1.19 software package (Li et al., 2009) with a maximum read depth of 1,000. This package consists of two key programs, SAMTOOLS and BCFTOOLS, which incorporate genotype likelihood and call variants using Bayesian inference. In order to remove false positives, putative SNPs were required to fit the following criteria: (i) only bi-allelic SNPs were retained; (ii) SNPs with an average phred-scaled quality score > 30 (Q30) were retained; (iii) a minimum read depth of 10 for each individual was required; (iv) the minimum individual coverage for each population was set to 80%; v) a global minor allele frequency (MAF) of 0.05 was applied; vi) SNPs with $F_{IS}$ values < -0.3 or > 0.3 and observed heterozygosity ($H_O$) values > 0.5 were discarded to exclude false SNPs, which might have resulted from the paralog clusters (Hohenlohe et al., 2011). All subsequent datasets used were reformatted with PGDSPIDER version 2.0.5.2 (Lischer and Excoffier, 2012).

## Outlier Detection

$F_{ST}$-based outlier tests implemented in ARLEQUIN v3.5.2.2 (Excoffier and Lischer, 2010) were applied to test for selection between the two samples. This approach obtains the distribution of $F_{ST}$ across loci as a function of heterozygosity between populations by performing coalescent simulations under a symmetrical two island model assuming near-random mating in order to reduce the number of false-positive loci (Beaumont and Nichols, 1996). For each outlier locus, $F_{ST}$ values located above the 99.5th quantile of the simulated distribution were selected.

## Population Genetic Analyses

The package VCFTOOLS (Danecek et al., 2011) was used to estimate genome-wide nucleotide $F_{ST}$ values, as well as observed ($H_O$) and expected ($H_E$) heterozygosity at each SNP for all individuals. Population pairwise $F_{ST}$ values were calculated using ARLEQUIN v3.5.2.2 (Excoffier and Lischer, 2010) and significance was determined using 10,000 permutations. Furthermore, ADMIXTURE v1.23 (Alexander et al., 2009) was used for maximum likelihood estimation of individual ancestries from multi-locus SNP genotype datasets. The analysis was performed with values of $K$ ranging from 1 to 3, and the $K$ with the lowest cross-validation error was selected. NETVIEW P 0.7.1 (Steinig et al., 2016), which is an implementation of NETVIEW (Neuditschko et al., 2012) in Python, was used to infer population structure including individual- and family-level relationships of populations on the basis of putative neutral SNPs and outlier datasets. A $K$-nearest neighbor (KNN) value of 10 was selected for visualization in CYTOSCAPE 3.2.1 (Shannon et al., 2003) as suggested by the author. The analyses of population pairwise $F_{ST}$, ADMIXTURE, and NETVIEW were conducted based on all SNPs, neutral SNPs, and outlier datasets.

## BLAST Analyses and Gene Ontology Annotation

Contigs containing outlier SNPs were used as queries in nucleotide searches using local BLASTX against the Genome Reference Consortium Zebrafish Build 10 (GRCz10) proteins with an $E$-value threshold of 1E-5. Next, functional annotations of these genes were obtained using the BLAST2GO suite (Götz et al., 2008), with parameters set according to previous recommendations (Pujolar et al., 2013). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis implemented in the Database for Annotation, Visualization and Integrated Discovery (DAVID) web-server v6.7[3] was used for a more systematic functional interpretation of the set of candidate genes. A gene functional analysis in DAVID was conducted with default settings.

## Comparison of the Number of Single-Nucleotide Polymorphisms Generated From Paired and Unpaired Reads

In order to compare the number of SNPs generated from paired-end reads with that of SNPs generated from single-end reads, SNP calling was performed on each single-end read group separately in addition to SNP calling on reads mapped as pairs. Single-end BAM files were extracted from paired-end BAM files by using SAMTOOLS. SNPs were called by using SAMTOOLS and BCFTOOLS, followed by additional filtering based on the following criteria: (i) retained only bi-allelic SNPs; (ii) retained SNPs with a minimum quality score of 30; and (iii) retained SNPs with a minimum read depth of 10. In order to increase read depth, individuals were pooled together before filtering.

# RESULTS

## Restriction Site-Associated DNA Tag Sequencing and Data Filtering

Sequencing of the RAD libraries generated a total of 42.10 Gb of raw data with an average of 1.75 Gb per individual, ranging from 1.02 to 7.84 Gb (**Table 1**). A total of 40.00 Gb of clean data was retained after quality filtering, with an average effective rate

[2]www.eelgenome.com

[3]http://david.abcc.ncifcrf.gov

**TABLE 1 |** Statistics describing the distribution of different properties of RAD sequences after each step of filtering.

| Sample | Raw data (G) | Clean data (G) | Error rate (%) | Q20 (%) | Q30 (%) | GC content (%) | Clean reads | Removed duplication reads | Clean duplication rate (%) | Digestion reads | Digestion ratio (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MD03 | 1.50 | 1.44 | 0.06 | 90.72 | 83.15 | 40.14 | 5,768,077 | 5,526,159 | 4.19 | 5,182,323 | 93.78 |
| MD05 | 1.50 | 1.43 | 0.06 | 90.54 | 82.83 | 40.53 | 5,732,607 | 5,475,147 | 4.49 | 5,090,677 | 92.98 |
| MD09 | 1.50 | 1.43 | 0.06 | 90.56 | 83.13 | 40.91 | 5,737,303 | 5,508,928 | 3.98 | 5,122,880 | 92.99 |
| MD22 | 1.50 | 1.44 | 0.06 | 90.87 | 83.36 | 40.02 | 5,774,108 | 5,555,422 | 3.79 | 5,196,842 | 93.55 |
| MD23 | 1.50 | 1.43 | 0.04 | 94.65 | 90.08 | 41.49 | 5,723,214 | 4,853,016 | 15.2 | 4,672,254 | 96.28 |
| MD24 | 1.23 | 1.15 | 0.04 | 94.62 | 89.97 | 40.75 | 4,587,874 | 3,788,128 | 17.43 | 3,587,807 | 94.71 |
| MD25 | 1.67 | 1.59 | 0.04 | 94.82 | 90.36 | 40.55 | 6,342,873 | 5,152,866 | 18.76 | 4,923,362 | 95.55 |
| MD26 | 1.50 | 1.44 | 0.04 | 94.66 | 89.93 | 41.28 | 5,755,968 | 5,088,562 | 11.6 | 4,940,675 | 97.09 |
| MD27 | 1.83 | 1.73 | 0.04 | 94.81 | 90.13 | 40.5 | 6,922,638 | 5,532,694 | 20.08 | 5,324,355 | 96.23 |
| MD28 | 1.28 | 1.17 | 0.04 | 93 | 87.2 | 40.8 | 4,698,576 | 4,069,704 | 13.38 | 3,726,116 | 91.56 |
| MD29 | 1.02 | 0.92 | 0.04 | 94.27 | 89.51 | 40.33 | 3,669,955 | 2,945,528 | 19.74 | 2,700,099 | 91.67 |
| MD30 | 1.86 | 1.79 | 0.04 | 94.6 | 90.34 | 41.78 | 7,160,053 | 6,314,835 | 11.8 | 6,058,807 | 95.95 |
| MS13 | 2.00 | 1.90 | 0.04 | 94.76 | 90.24 | 41.72 | 7,580,224 | 6,083,254 | 19.75 | 5,796,104 | 95.28 |
| MS14 | 1.50 | 1.44 | 0.03 | 95.27 | 91.09 | 41.41 | 5,773,666 | 5,027,364 | 12.93 | 4,856,877 | 96.61 |
| MS20 | 1.50 | 1.44 | 0.04 | 94.25 | 89.47 | 40.69 | 5,777,831 | 5,138,407 | 11.07 | 4,968,681 | 96.7 |
| MS23 | 1.17 | 1.04 | 0.04 | 94.47 | 89.84 | 40.43 | 4,151,896 | 3,263,872 | 21.39 | 2,849,829 | 87.31 |
| MS24 | 1.50 | 1.45 | 0.06 | 90.18 | 82.31 | 40.27 | 5,800,946 | 5,605,437 | 3.37 | 5,262,247 | 93.88 |
| MS25 | 7.84 | 7.62 | 0.04 | 92.58 | 86.72 | 41.7 | 30,498,830 | 26,873,349 | 11.89 | 26,175,981 | 97.4 |
| MS26 | 1.50 | 1.44 | 0.06 | 90.28 | 82.53 | 40.5 | 5,769,821 | 5,554,405 | 3.73 | 5,179,480 | 93.25 |
| MS27 | 1.68 | 1.58 | 0.04 | 94.84 | 90.16 | 40.56 | 6,334,869 | 5,040,194 | 20.44 | 4,850,227 | 96.23 |
| MS31 | 1.31 | 1.20 | 0.04 | 95 | 90.6 | 40.39 | 4,784,800 | 3,366,384 | 29.64 | 3,163,174 | 93.96 |
| MS32 | 1.25 | 1.19 | 0.07 | 89.97 | 81.96 | 40.7 | 4,764,364 | 4,531,800 | 4.88 | 4,179,097 | 92.22 |
| MS33 | 1.40 | 1.30 | 0.04 | 93.41 | 87.89 | 40.32 | 5,186,488 | 4,599,961 | 11.31 | 4,252,827 | 92.45 |
| MS34 | 1.57 | 1.42 | 0.04 | 94.94 | 90.57 | 40.64 | 5,699,749 | 4,061,840 | 28.74 | 3,752,240 | 92.38 |
| AVG | 1.75 | 1.67 | 0.05 | 93.25 | 87.64 | 40.77 | 6,666,530 | 5,789,886 | 13.48 | 5,492,207 | 94.17 |

*GC, guanine–cytosine; RAD, restriction site-associated DNA.*

(clean reads compared with raw reads) of 95.01%. The clean data had a guanine–cytosine (GC) content that ranged from 40.02 to 41.78%, and the proportion of Q20 (the bases of one individual with Phred Q > 20) ranged from 89.97 to 95.27%, whereas that of Q30 ranged from 81.96 to 91.09%.

## Single-Nucleotide Polymorphism Discovery and Population Genetic Structure

Of the retained sequences, an average of 10.51 million (97.35%) reads per individual were aligned to the Japanese eel draft genome contigs, and then 9,905,533 SNPs were called from 24 individuals using SAMTOOLS. After quality filtering, a total of 73,557 SNPs were retained. The final retained SNPs were widely distributed across the genome and were found in a total of 28,407 contigs.

The expected and observed heterozygosity ($H_E$ and $H_O$) per individual ranged from 0.194 to 0.198 and from 0.157 to 0.312, respectively (**Table 2**). The $F_{ST}$ values between the two populations ranged from −0.0667 to 0.4354 for each SNP. The final dataset had a transition:transversion ratio of 1.83:1 (**Figure 1**).

ARLEQUIN 3.5 obtained an $F_{ST}$ value of 0.00271 ($P > 0.05$) between the two populations using all the retained SNPs. The population pairwise $F_{ST}$ was 0.00112 ($P > 0.05$) based on the neutral SNPs and 0.28473 ($P < 0.001$) based on the outlier SNPs. Results of ADMIXTURE and NETVIEW showed that all individuals from Dandong and Xiapu were grouped into a single cluster when using the neutral or all SNPs. However, when the analysis was based on the outlier SNPs, the two populations were clearly separated (**Figure 2**).

**TABLE 2 |** Statistics of the expected and observed heterozygosity ($H_E$ and $H_O$) per individual.

| INDV | $H_O$ | $H_E$ | Mean depth |
|---|---|---|---|
| MD03 | 0.166 | 0.194 | 17.670 |
| MD05 | 0.167 | 0.194 | 18.525 |
| MD09 | 0.163 | 0.194 | 20.422 |
| MD22 | 0.165 | 0.194 | 18.217 |
| MD23 | 0.281 | 0.194 | 23.857 |
| MD24 | 0.265 | 0.195 | 16.100 |
| MD25 | 0.224 | 0.194 | 20.995 |
| MD26 | 0.248 | 0.194 | 24.821 |
| MD27 | 0.228 | 0.194 | 23.254 |
| MD28 | 0.263 | 0.195 | 17.686 |
| MD29 | 0.312 | 0.198 | 11.490 |
| MD30 | 0.168 | 0.194 | 30.835 |
| MS13 | 0.200 | 0.194 | 28.360 |
| MS14 | 0.172 | 0.194 | 22.347 |
| MS20 | 0.213 | 0.194 | 23.576 |
| MS23 | 0.269 | 0.196 | 11.495 |
| MS24 | 0.164 | 0.194 | 18.503 |
| MS25 | 0.157 | 0.194 | 98.453 |
| MS26 | 0.169 | 0.194 | 19.806 |
| MS27 | 0.191 | 0.194 | 22.110 |
| MS31 | 0.238 | 0.196 | 12.258 |
| MS32 | 0.172 | 0.194 | 16.861 |
| MS33 | 0.286 | 0.194 | 18.053 |
| MS34 | 0.232 | 0.194 | 16.001 |

## BLAST Analyses of the Divergent Loci

A total of 250 outlier SNPs representing 238 unique contigs in the draft genome were detected, with all showing high $F_{ST}$ values (0.209–0.435). BLASTX similarity searches demonstrated that 85 contigs (36%) significantly matched to zebrafish proteins, with 61 of these contigs corresponding to zebrafish proteins Gene Ontology (GO) functional annotations (**Figure 3**). These annotations are represented by GO terms including signaling, response to stimulus, metabolic process, biological regulation, catalytic activity, development process, and others (**Figure 4**). Classification of biological and molecular functionalities for these hits is listed in **Supplementary Table S1**. A total of 228 zebrafish genes, which were homologous to the 85 contigs with significant BLASTX hits, were loaded into the DAVID database and were found to correspond to 224 DAVID IDs. Enriched KEGG pathways including 52 genes classified into nine pathways were summarized in **Table 3**. The pathway with the highest number of genes was endocytosis (10 genes), including several genes such as those encoding chemokine (C-X-C motif) receptor 4, fibroblast growth factor receptor, and par-6 partitioning defective 6 homolog. Other pathways of particular interest were neuroactive ligand–receptor interaction (nine genes), calcium signaling pathway (eight genes), intestinal immune network for IgA production (three genes), and mitogen-activated protein kinase (MAPK) signaling pathways (eight genes).

## The Number of Single-Nucleotide Polymorphisms Generated Based on Single Reads or Paired Reads

A total of 4,227,450 SNPs were called using the first reads before filtering, whereas 7,741,237 SNPs were called when using the second reads. A total of 2,449,359 for the first reads, 3,724,441 for the second reads, and 5,515,225 for both of the paired reads were retained after applying filtration. However, much fewer SNPs were retained for the second reads while applying the filter on each individual, suggesting low depth of the second reads. This is a common phenomenon in RAD, as the randomly sheared fragments are expected to be sequenced at a much lower depth than the other ends. However, the longer fragment coverage of the second reads increases the chance for targeting loci under local adaptation and is also suitable for annotations.

# DISCUSSION

Compared with microsatellites, SNPs are the most widespread type of sequence variation in genomes. Despite microsatellites presenting higher diversity per locus (with a mutation rate of $10^{-4}$ per generation), the large number of unlinked SNPs may overcome their relatively low mutation rates ($10^{-8}$–$10^{-9}$ per generation), and panels of several 1000 SNPs are likely to be more informative than the 10–20 microsatellite loci used in

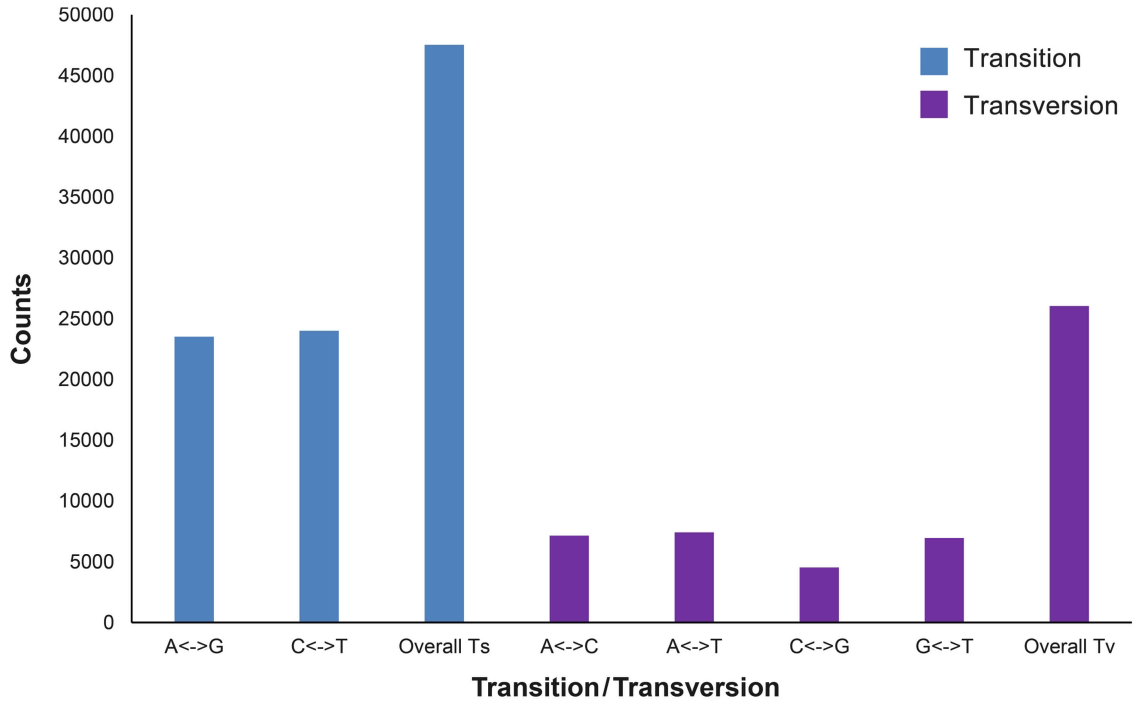


**FIGURE 1 |** Transitions and transversions occurring within a set of 73,557 single-nucleotide polymorphisms (SNPs).

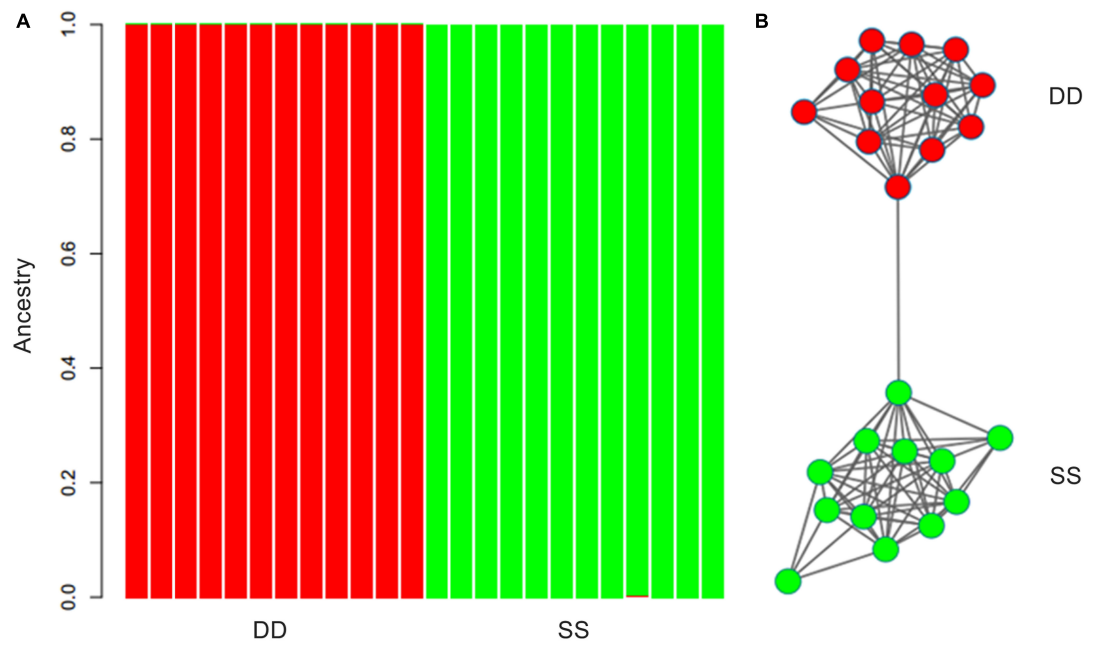


**FIGURE 2 |** Results of ADMIXTURE **(A)** and NETVIEW **(B)** based on outlier dataset.

standard population genetic studies (Helyar et al., 2011; Seeb et al., 2011). Here, we reported the discovery of a large number of genome-wide SNPs in the Japanese eel using RAD sequencing. After quality filtering, 73,557 high-quality SNPs were retained, which further improved the amount of genomic resources available for this endangered species. Most previous population genetic studies of *A. japonica* were conducted using less than 20 microsatellite loci, mitochondrial sequence, or isozyme markers, resulting in inconsistent findings (Sang et al., 1994; Chan et al., 1997; Tseng et al., 2003, 2006, 2009; Han et al., 2010a). Compared

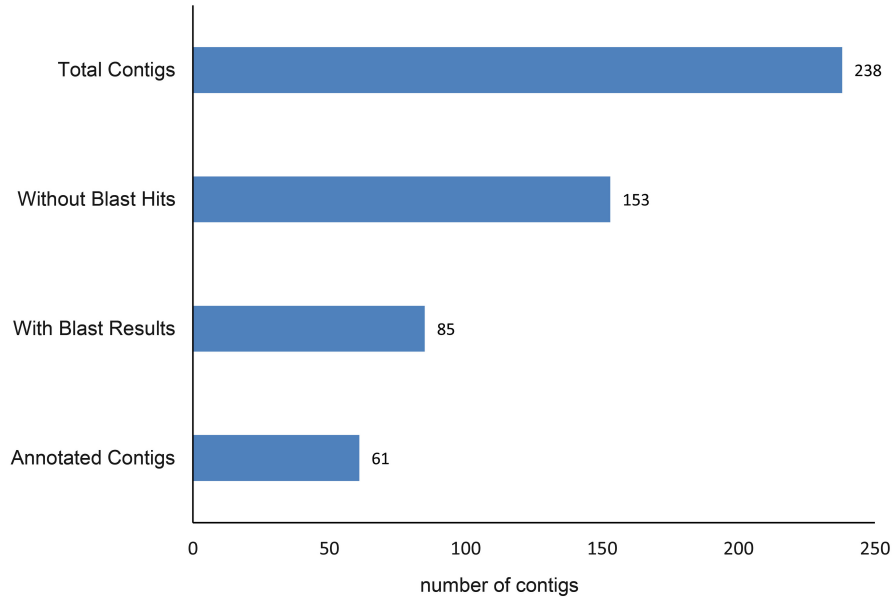


**FIGURE 3 |** The numbers of the annotated contigs.

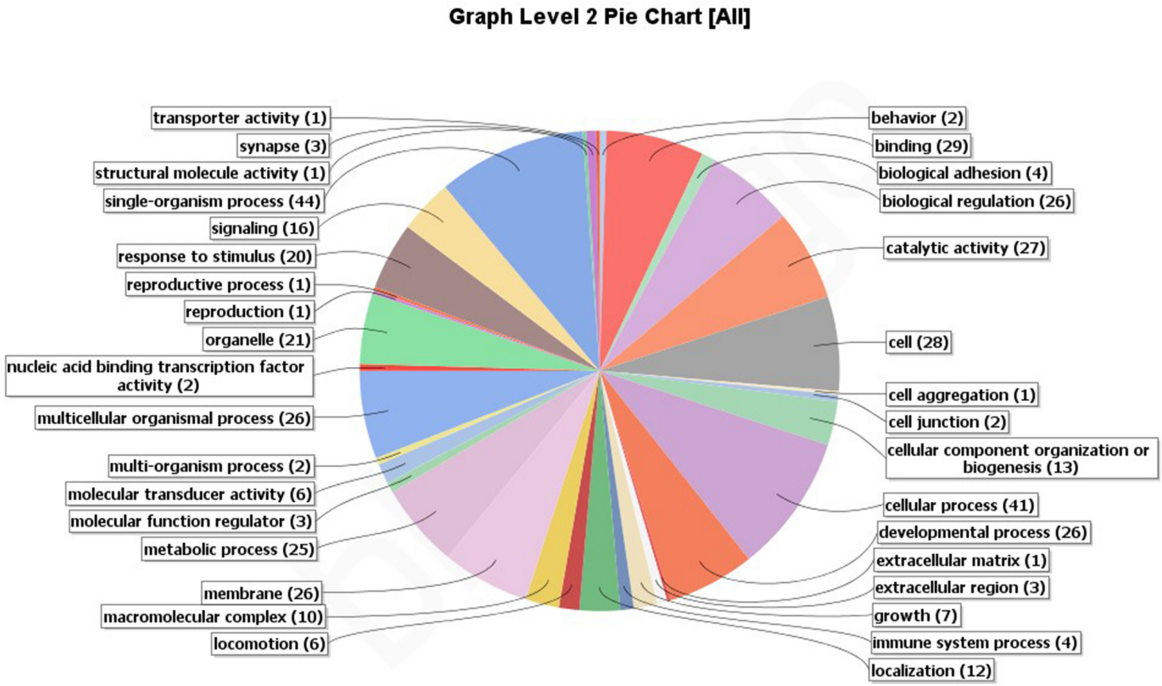


**FIGURE 4 |** Putative functional categorization and distribution of 61 significant BLAST hits (*E*-value ≤ 1E-5) of *Anguilla japonica* according to BLAST2GO annotation.

with other studies of eels on the basis of RAD (Pujolar et al., 2013, 2014), our study made full use of the paired-end reads, which resulted in significantly more markers and enabled the identification of genomic regions under selection.

The transition:transversion ratio of the SNPs in our dataset was 1.83:1, which is in keeping with similar ratios that have been reported in the study of European eel (*Anguilla anguilla*) (1.6:1) (Pujolar et al., 2013) and great tit (*Parus major*) (1.7:1) (Bers et al., 2010). The overall non-significant pairwise $F_{ST}$ (0.00271, $P > 0.05$) indicated that most of the genome is homogenized by gene flow, supporting the theory of panmixia in the Japanese eel. However, the $F_{ST}$ based on the outlier SNPs was 0.285 and significant ($P < 0.001$). In addition, both ADMIXTURE and NETVIEW analyses based on the outlier

**TABLE 3 |** KEGG pathways and genes identified using the $F_{ST}$ outlier approach implemented in ARLEQUIN.

| KEGG pathway | Count | Genes |
|---|---|---|
| Endocytosis | 10 | Chemokine (C-X-C motif) receptor 4a<br>Chemokine (C-X-C motif), receptor 4b<br>Fibroblast growth factor receptor 2<br>Fibroblast growth factor receptor 3<br>Fibroblast growth factor receptor 4<br>Par-6 partitioning defective 6 homolog beta (*Caenorhabditis elegans*)<br>Par-6 partitioning defective 6 homolog gamma A (*C. elegans*); similar to par-6 partitioning defective 6-like protein gamma<br>Par-6 partitioning defective 6 homolog gamma B (*C. elegans*)<br>Similar to CXC chemokine receptor-2<br>zgc:162756; par-6 partitioning defective 6 homolog alpha (*C. elegans*) |
| Heparan sulfate biosynthesis | 4 | Heparan sulfate 6-*O*-sulfotransferase 1b<br>Heparan sulfate 6-*O*-sulfotransferase 2<br>Heparan sulfate 6-*O*-sulfotransferase 3<br>Similar to heparan sulfate 6-*O*-sulfotransferase 3 |
| Neuroactive ligand–receptor interaction | 9 | Apelin receptor a<br>Apelin receptor b<br>Similar to G protein-coupled receptor 83<br>Similar to neuromedin-K receptor (NKR) (neurokinin B receptor) (NK-3 receptor) (NK-3R) (tachykinin receptor 3)<br>Similar to neuromedin-K receptor (NKR) (neurokinin B receptor) (NK-3 receptor) (NK-3R) (tachykinin receptor 3)<br>Similar to substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (tachykinin receptor 1)<br>Similar to type 1B angiotensin II receptor (AT1B) (AT3)<br>Similar to tachykinin receptor 3<br>Similar to tachykinin receptor 3 |
| Calcium signaling pathway | 8 | Phospholipase C, delta 4a<br>Similar to neuromedin-K receptor (NKR) (neurokinin B receptor) (NK-3 receptor) (NK-3R) (tachykinin receptor 3)<br>Similar to neuromedin-K receptor (NKR) (neurokinin B receptor) (NK-3 receptor) (NK-3R) (tachykinin receptor 3)<br>Similar to substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (tachykinin receptor 1)<br>Similar to type 1B angiotensin II receptor (AT1B) (AT3)<br>Similar to phospholipase C, gamma 2; si:ch211-260p9.3<br>Similar to tachykinin receptor 3<br>Similar to tachykinin receptor 3 |
| Intestinal immune network for IgA production | 3 | Chemokine (C-X-C motif) receptor 4a<br>Chemokine (C-X-C motif), receptor 4b<br>Hypothetical protein LOC553366; sb:eu630 |
| ECM–receptor interaction | 4 | Thrombospondin 1<br>Thrombospondin 2<br>Thrombospondin 3a<br>Thrombospondin 4b |
| Fatty acid biosynthesis | 2 | im:7138837<br>similar to Acetyl-CoA carboxylase 2 (ACC-beta) |
| Cytokine–cytokine receptor interaction | 5 | Chemokine (C-X-C motif) receptor 4a<br>Chemokine (C-X-C motif), receptor 4b<br>Hypothetical protein LOC553366; sb:eu630<br>Similar to CXC chemokine receptor-2<br>zgc:136557 |
| MAPK signaling pathway | 8 | Brain-derived neurotrophic factor<br>Fibroblast growth factor receptor 1a<br>Fibroblast growth factor receptor 2<br>Fibroblast growth factor receptor 3<br>Fibroblast growth factor receptor 4<br>Mitogen-activated protein kinase kinase kinase 7<br>Nerve growth factor (beta polypeptide)<br>Neurotrophin 3 |

*ECM, extracellular matrix; KEGG, Kyoto Encyclopedia of Genes and Genomes.*

SNPs demonstrated strong genetic differentiation between the two populations. Taken together, these results indicated that some genomic regions of the Japanese eel have undergone divergent natural selection.

Furthermore, RAD sequencing-based population genomic studies can also detect signatures of selection and local adaptation. Candidate genes and genomic regions were identified using an $F_{ST}$ outlier approach by detecting loci showing increased or decreased differentiation across populations compared with neutral expectations, which are suggestive of directional or purifying natural selection. For a panmictic species like Japanese eel, SNP-based genome scans can be used for test selection within a single generation, which results from geographically varying environmental conditions encountered by glass eels across different regions. Such outlier detection methods have been illustrated by several recent population genomic studies in marine fishes displaying extensive gene flow (Russello et al., 2012; Milano et al., 2014). For instance, Hess et al. (2013) detected fine-scale population structure in Pacific lamprey (*Entosphenus tridentatus*) by using 162 putatively adaptive SNPs identified by RAD sequencing. Larson et al. (2014) found three potential regions of adaptive divergence in Chinook salmon (*Oncorhynchus tshawytscha*) using RAD sequencing, showing that RAD sequencing is a promising approach for marine species with large population sizes and shallow structures.

The two populations of *A. japonica* analyzed were collected from two environmentally heterogeneous locations, and different environmental factors may lead to local adaptation. BLASTX and BLAST2GO revealed that 61 out of 238 highly divergent contigs were located in functional genes or genomic regions, which might indicate divergent natural selection (Foll and Gaggiotti, 2008). The low frequency of BLASTX-annotated outliers suggested that substantial outlier polymorphisms detected might be in the intergenic or intronic regions of the genome. The KEGG pathway analysis showed that some of the putative targets under local selection were associated with genes involved in several important pathways, such as calcium signaling pathway, intestinal immune network for IgA production, cytokine–cytokine receptor interaction, and MAPK signaling pathways. Differences in these pathways might indicate local adaptation of *A. japonica* to different geographical environments, which vary in salinity and temperature. For instance, the calcium signaling pathway plays an important role in fertilization, development, learning and memory, ion exchange, and osmoregulation (Berridge et al., 2000; Daverat et al., 2006). The intestine can generate large amounts of non-inflammatory immunoglobulin A (IgA) antibodies that serve as the first line of defense against microorganisms (Reinking et al., 1999) and is also crucial for adaptive inflammatory host defenses, cell growth, differentiation, cell death, and angiogenesis (Lackie, 2010). Similar results were also reported in other studies. For instance, Guo et al. (2015) investigated genome-wide patterns of genetic variability of Atlantic herring using the RAD sequencing approach, and outlier analyses uncovered 100s of directionally selected loci that were associated with salinity and temperature. Pujolar et al. (2014) found a small set of SNPs showing high genetic differentiation, which is consistent with single-generation signatures of spatially varying selection acting on European eel (glass eels) using RAD sequencing. After 50,354 SNPs were screened, a total of 754 potentially locally selected SNPs were identified. Candidate genes for local selection include those involved in calcium signaling, neuroactive ligand–receptor interaction, circadian rhythm, and others.

It remains an open question whether the adaptive loci inferred in this study were directly related to the local adaptation of *A. japonica* populations to specific environments or were linked to genes underlying specific traits. Further analyses are necessary to confirm divergent selection on outliers through a correlation analysis between allele frequencies and environmental factors among populations (Seeb et al., 2011; Limborg et al., 2012; DeFaveri et al., 2013; Teacher et al., 2013).

## DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article and its **Supplementary Information Files**.

## ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Ethics Committee of Zhejiang Ocean University.

## AUTHOR CONTRIBUTIONS

J-XL and B-JL conceived and designed the research. B-JL, Y-LL, and B-DZ conducted the experiments, analyzed the data, and wrote the manuscript. All authors critically reviewed and approved the manuscript.

## FUNDING

The study was supported by the National Natural Science Foundation of China (NSFC) (Nos. 41676137 and 41806156); Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology (2017-2B05); Natural Science Foundation of Zhejiang Province (LY17C200010 and LY15D060004); Fundamental Research Funds for Zhejiang Provincial Universities and Research Institutes (No. 2019J00048); Starting Research Fund from the Zhejiang Ocean University; The Fundamental Research Funds for Zhejiang Provincial Universities and Research Institutes; and Open Foundation from Marine Sciences in the First-Class Subjects of Zhejiang Province.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00782/full#supplementary-material

## REFERENCES


Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11, 697–709. doi: 10.1038/nrg2844

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD Markers. *PLoS One* 3:e3376. doi: 10.1371/journal.pone.0003376

Beaumont, M. A., and Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. B Biol. Sci.* 263, 1619–1626. doi: 10.1098/rspb.1996.0237

Berridge, M. J., Lipp, P., and Bootman, M. D. (2000). The versatility and universality of calcium signalling. *Nat. Rev. Mol. Cell Biol.* 1, 11–21. doi: 10.1038/35036035

Bers, N. E., Oers, K. V., Kerstens, H. H., Dibbits, B. W., Crooijmans, R. P., Visser, M. E., et al. (2010). Genome-wide SNP detection in the great tit Parus major using high throughput sequencing. *Mol. Ecol.* 19, 89–99. doi: 10.1111/j.1365-294X.2009.04486.x

Chan, I. K. K., Chan, D. K. O., Lee, S. C., and Tsukamoto, K. (1997). Genetic variability of the Japanese eel Anguilla japonica (*Temminck* & *Schlegel*) related to latitude. *Ecol Freshw Fish.* 6, 45–49. doi: 10.1111/j.1600-0633.1997.tb00141.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

Daverat, F., Limburg, K. E., Thibault, I., Shiao, J. C., Dodson, J. J., Caron, F. O., et al. (2006). Phenotypic plasticity of habitat use by three temperate eel species, *Anguilla anguilla*, A-japonica and A-rostrata. *Mar. Ecol. Prog. Ser.* 308, 231–241. doi: 10.3354/meps308231

DeFaveri, J., Shikano, T., Shimada, Y., and Merila, J. (2013). High degree of genetic differentiation in marine three-spined sticklebacks (*Gasterosteus aculeatus*). *Mol. Ecol.* 22, 4811–4828. doi: 10.1111/mec.12430

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., et al. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *J. Fish Biol.* 80, 1789–1843. doi: 10.1111/j.1095-8649.2012.03265.x

Ekblom, R., and Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107, 1–15. doi: 10.1038/hdy.2010.152

Excoffier, L., and Lischer, H. E. L. (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. doi: 10.1111/j.1755-0998.2010.02847.x

Foll, M., and Gaggiotti, O. E. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221

Gong, X. L., Ren, S. J., Cui, Z. K., and Yue, L. J. (2014). Genetic evidence for panmixia of Japanese eel (*Anguilla japonica*) populations in China. *Genet. Mol. Res.* 13, 768–781. doi: 10.4238/2014.January.31.3

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176

Guo, B., DeFaveri, J., Sotelo, G., Nair, A., and Merilä, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol.* 13:19. doi: 10.1186/s12915-015-0130-8

Han, Y. S., Hung, C. L., Liao, Y. F., and Tzeng, W. N. (2010a). Population genetic structure of the Japanese eel *Anguilla japonica*: panmixia at spatial and temporal scales. *Mar. Ecol. Prog. Ser.* 401, 221–232. doi: 10.3354/meps08422

Han, Y.-S., Iizuka, Y., and Tzeng, W.-N. (2010b). Does Variable Habitat Usage by the Japanese Eel Lead to Population Genetic Differentiation? *Zool. Stud.* 49, 392–397.

Harismendy, O., Ng, P. C., Strausberg, R. L., Wang, X., Stockwell, T. B., Beeson, K. Y., et al. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol.* 10:R32. doi: 10.1186/gb-2009-10-3-r32

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., et al. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Mol. Ecol. Resour.* 11:36. doi: 10.1111/j.1755-0998.2010.02943.x

Hess, J. E., Campbell, N. R., Close, D. A., Docker, M. F., and Narum, S. R. (2013). Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Mol. Ecol.* 22, 2898–2916. doi: 10.1111/mec.12150

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., and Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Mol. Ecol. Resour.* 11, 117–122. doi: 10.1111/j.1755-0998.2010.02967.x

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., and Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Gene.*6:e1000862. doi: 10.1371/journal.pgen.1000862

Hudson, M. E. (2008). Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.* 8, 3–17. doi: 10.1111/j.1471-8286.2007.02019.x

Ishikawa, S., Aoyama, J., Tsukamoto, K., and Nishida, M. (2001). Population structure of the Japanese eel Anguilla japonica as examined by mitochondrial DNA sequencing. *Fish Sci.* 67, 246–253. doi: 10.1046/j.1444-2906.2001.00227.x

Kruck, N. C., Innes, D. I., and Ovenden, J. R. (2013). New SNPs for population genetic analysis reveal possible cryptic speciation of eastern Australian sea mullet (*Mugil cephalus*). *Mol. Ecol. Resour.* 13, 715–725. doi: 10.1111/1755-0998.12112

Lackie, J. (2010). "Cytokine," in *John Lackie. A Dictionary of Biomedicine*, ed. C. A. O'Callaghan, (Oxford: Oxford University Press).

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., and Seeb, J. E. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evol. App.* 7:69. doi: 10.1111/eva.12128

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Limborg, M. T., Helyar, S. J., Bruyn, M. D., Taylor, M. I., Nielsen, E. E., Ogden, R., et al. (2012). Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Mol. Ecol.* 21, 2355–2356. doi: 10.1111/j.1365-294X.2012.05639.x

Lischer, H., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299. doi: 10.1093/bioinformatics/btr642

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., et al. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Mol. Ecol.* 23, 118–135. doi: 10.1111/mec.12568

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., and Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17, 240–248. doi: 10.1101/gr.5681207

Neuditschko, M., Khatkar, M. S., and Raadsma, H. W. (2012). NETVIEW: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. *PLoS One* 7:e48375. doi: 10.1371/journal.pone.0048375

Palti, Y., Gao, G. T., Miller, M. R., Vallejo, R. L., Wheeler, P. A., Quillet, E., et al. (2014). A resource of single-nucleotide polymorphisms for rainbow trout generated by restriction-site associated DNA sequencing of doubled haploids. *Mol. Ecol. Resour.* 14, 588–596. doi: 10.1111/1755-0998.12204

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jonsson, B., et al. (2014). Genome-wide single-generation signatures of local selection in the panmictic European eel. *Mol. Ecol.* 23, 2514–2528. doi: 10.1111/mec.12753

Pujolar, J. M., Jacobsen, M. W., Frydenberg, J., Als, T. D., Larsen, P. F., Maes, G. E., et al. (2013). A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. *Mol. Ecol. Resour.* 13, 706–714. doi: 10.1111/1755-0998.12117

Reinking, R. R., Raedler, A., and Möckel, G. (1999). The intestinal immune system. *Medizinische Monatsschrift Für Pharmazeuten.* 22, 198–204.

Russello, M. A., Kirk, S. L., Frazer, K. K., and Askey, P. J. (2012). Detection of outlier loci and their utility for fisheries management. *Evol. App.* 5, 39–52. doi: 10.1111/j.1752-4571.2011.00206.x

Sang, T. K., Chang, H. Y., Chen, C. T., and Hui, C. F. (1994). Population structure of the japanese eel. *Anguilla japonica*. *Mol. Biol. Evol.* 11, 250–260.

Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., and Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Mol. Ecol. Resour.* 11(Suppl. 1), 1–8. doi: 10.1111/j.1755-0998.2010.02979.x

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Slate, J., Gratten, J., Beraldi, D., Stapley, J., Hale, M., and Pemberton, J. M. (2009). Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* 136, 97–107. doi: 10.1007/s10709-008-9317-z

Steinig, E. J., Neuditschko, M., Khatkar, M. S., Raadsma, H. W., and Zenger, K. R. (2016). netview p: a network visualization tool to unravel complex population structure using genome-wide SNPs. *Mol. Ecol. Resour.* 16, 216–227. doi: 10.1111/1755-0998.12442

Teacher, A. G. F., Andre, C., Jonsson, P. R., and Merila, J. (2013). Oceanographic connectivity and environmental correlates of genetic structuring in Atlantic herring in the Baltic Sea. *Evol. App.* 6, 549–567. doi: 10.1111/eva.12042

Tseng, M. C., Tzeng, W. N., and Lee, S. C. (2003). Historical decline in the Japanese eel *Anguilla japonica* in northern Taiwan inferred from temporal genetic variations. *Zool. Stud.* 42, 556–563.

Tseng, M. C., Tzeng, W. N., and Lee, S. C. (2006). Population genetic structure of the Japanese eel *Anguilla japonica* in the northwest Pacific Ocean: evidence of non-panmictic populations. *Mar. Ecol. Prog. Ser.* 308, 221–230. doi: 10.3354/meps308221

Tseng, M. C., Tzeng, W. N., and Lee, S. C. (2009). Genetic differentiation of the Japanese eel. *Am. Fish. Soc. Sym.*. 58, 59–69.

Tsukamoto, K. (1992). Discovery of the spawning area for japanese eel. *Nature* 356, 789–791. doi: 10.1038/356789a0

Tsukamoto, K. (2006). Oceanic biology: spawning of eels near a seamount. *Nature* 439:929. doi: 10.1038/439929a

Wang, X. Q., Zhao, L., Eaton, D. A. R., Li, D. Z., and Guo, Z. H. (2013). Identification of SNP markers for inferring phylogeny in temperate bamboos (*Poaceae*: *Bambusoideae*) using RAD sequencing. *Mol. Ecol. Resour.* 13, 938–945. doi: 10.1111/1755-0998.12136

Zhang, B., Xue, D., Wang, J., Li, Y. L., Liu, B. J., Liu, J. X., et al. (2016). Development and preliminary evaluation of a genomewide single nucleotide polymorphisms resource generated by RAD -seq for the small yellow croaker (*Larimichthys polyactis*). *Mol. Ecol. Resour.* 16, 755–768. doi: 10.1111/1755-0998.12476

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer JZ declared a shared affiliation, with no collaboration, with one of the authors, B-DZ, to the handling Editor at the time of review.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 27 February 2020
doi: 10.3389/fmars.2020.00094

# Genetic Assessment of a Black Rockfish, *Sebastes schlegelii*, Stock Enhancement Program in Lidao Bay, China Based on Mitochondrial and Nuclear DNA Analysis

*Lijuan Wang[1,2], Zhihao Wu[1,2], Yingjun Wang[3], Mengxia Liu[3], Aihuan Song[3], Hongjun Liu[3] and Feng You[1,2]**

[1] Key Laboratory of Experimental Marine Biology, Center for Ocean Mega-Science, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, [2] Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China, [3] Marine Biology Institute of Shandong Province, Qingdao, China

OPEN ACCESS

**Edited by:**
*Dongdong Xu,*
*Marine Fisheries Research Institute of Zhejiang, China*

**Reviewed by:**
*Xiaoshou Liu,*
*Ocean University of China, China*
*Qi Liu,*
*Dalian Ocean University, China*

***Correspondence:**
*Feng You*
*youfeng@qdio.ac.cn*

**Specialty section:**
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

**Received:** *07 October 2019*
**Accepted:** *04 February 2020*
**Published:** *27 February 2020*

**Citation:**
*Wang L, Wu Z, Wang Y, Liu M, Song A, Liu H and You F (2020) Genetic Assessment of a Black Rockfish, Sebastes schlegelii, Stock Enhancement Program in Lidao Bay, China Based on Mitochondrial and Nuclear DNA Analysis. Front. Mar. Sci. 7:94. doi: 10.3389/fmars.2020.00094*

The black rockfish *Sebastes schlegelii* is a commercially important fish species for marine fishery stock enhancement in Asia. This work aimed to evaluate the potential genetic impacts of releasing hatchery-reared juvenile black rockfish on wild stock in Lidao Bay, China. A partial sequence of the mitochondrial DNA (mtDNA) control region and 38 microsatellite DNA loci were used to assess the genetic impact. The haplotype diversity parameter ($h$) and nucleotide diversity parameter ($\pi$) in hatchery-released stock were 0.902 and 0.00483, respectively. The $h$ values in wild stock before stock enhancement and the mixed stock after enhancement were 0.970 and 0.939 ($p$ = 0.025), respectively, and the $\pi$ values were 0.00581 and 0.00526 ($p$ = 0.150), respectively. The mean effective number of alleles (*Ae*) in hatchery-released stock was 4.76, the mean polymorphism information content (*PIC*) was 0.674, the observed heterozygosity (*Ho*) was 0.668, and the expected heterozygosity (*He*) was 0.697. In wild stock before stock enhancement and the mixed stock after enhancement, the *Ae*, *PIC*, *Ho*, and *He* values were 6.01 and 5.82, 0.698 and 0.716, 0.709 and 0.741, and 0.735 and 0.754, respectively. These results indicated no marked decrease in the hatchery-released stock, although it displayed slightly lower levels of genetic diversity and heterozygosity than the wild stock. And the mixed stock after release exhibited similar genetic diversity to that of the wild stock before release. Accordingly, we propose that stock enhancement may not cause genetic diversity reduction on wild *S. schlegelii* stock in Lidao Bay over the short term. Whereas, discriminant analysis of principal components (DAPC) analysis identified three major clusters. The fixation index ($F_{ST}$) and analysis of molecular variance (AMOVA) also clearly showed low level but significant differentiations between the hatchery-released and wild stocks, and the wild stock before enhancement and the mixed stock after enhancement. Consequently, long-term genetic evaluation might be required.

**Keywords: *Sebastes schlegelii*, mitochondrial DNA, microsatellite DNA, genetic variation, stock enhancement, Lidao Bay**

## INTRODUCTION

Over the past few decades, many natural systems have been negatively affected by anthropogenic stresses, such as overfishing and habitat degradation, and as a result, capture fisheries can no longer meet the demand for seafood (Garibaldi, 2012). Therefore, hatchery-reared organisms are often released into the wild to increase abundance and fishery yields (Leber, 2013). These restoration practices are applied for three main purposes: restocking, stock enhancement, and sea ranching (Bell et al., 2008). The mass release of juvenile hatchery-produced individuals into the wild has increased sharply in recent years worldwide, and more than 180 species were released in 20 countries from 2011 to 2016 (Kitada, 2018). In China, fisheries enhancements are practiced by both the country and the provinces along the country's vast acreage of coastal regions, and 94.543 billion marine fish, shrimp, and scallops were released during 2004–2013 (Luo and Zhang, 2014).

Although stock enhancement programs can be demographically beneficial, they also raise numerous questions related to their eventual ecological and genetic impacts on the recipient stock (Kitada, 2018). Compared with the ecological impacts, the genetic impacts of enhanced production on wild populations are more difficult to be assessed, because such effects may not be as apparent as short-term shifts in abundance (Grant et al., 2017). Domesticated strains of a species are usually produced from a small broodstock that harbors only a small portion of the total genetic diversity within a species. Therefore, they often differ genetically from their wild counterparts by showing lower genetic diversity, particularly when they have been systematically selected for multiple generations (Champagnon et al., 2012; Lorenzen et al., 2012). As stock enhancement aims to enhance both intra- and inter-generational stocking (Kitada, 2018), there may be both short-term and long-term genetic effects. The long-term effects are mainly caused by gene flow. Hybridization between hatchery and wild individuals has the potential to reduce fitness and genetic variation and change the genetic composition and population structure of the wild population (Laikre et al., 2010). The negative genetic impacts of stock enhancement have been documented in a wide range of fish species (Laikre et al., 2010), such as red drum (*Sciaenops ocellatus*), red sea bream (*Pagrus major*), and steelhead trout (*Oncorhynchus mykiss*) (Araki et al., 2007; Gold et al., 2008; Hamasaki et al., 2010). Moreover, even stock enhancement that does not result in gene flow can still destroy the gene pool by decreasing the size of the wild population through competition or disease transmission in the short term. In addition, the effective population size (*Ne*) of the wild population may be reduced due to increasing mortality. The *Ne* of cultured stocks is typically much lower than that of wild stocks, potentially leading to a depletion of genetic diversity through increased random drift. Consequently, enhancing wild populations with hatchery individuals can reduce the *Ne* of the entire mixed population (Ryman and Laikre, 1991; Ryman et al., 1995). Thus, evaluation of the genetic variability of released stock and recipient stock in such enchainment activities can yield valuable insights and ensure a responsible stock enhancement program (Romo et al., 2006; Ward, 2006).

Black rockfish (*Sebastes schlegelii*) is an important commercial fish inhabiting the coasts of China, Japan, and Korea. It has strong site fidelity and limited capacity for movement (Kang and Shin, 2006; Zhang et al., 2015). Therefore, black rockfish is considered to be an ideal candidate fish species for stock enhancement. The catch per unit of effort (CPUE) and individual weight of *S. schlegelii* have declined since the 1980s (Xu and Jin, 2005), and stock enhancement programs were initiated in these three countries to recover overexploited stocks. In China, enhancement programs for *S. schlegelii* began in 1995, and this species has gradually become one of the most popular fish species for release. Millions of juveniles are released in the northern coastal region of China every year (Lü et al., 2014). However, scant information on the genetic effects of release is available, barring a few reports on enhancement technology (Nihira and Takashima, 1999), growth and movement (Nakagawa, 2008), rates of return and economic return (Nakagawa et al., 2007; Noda et al., 2012), and release habitat (Chin et al., 2013). Although a few studies have dealt with the genetic diversity and differentiation of different geographical populations (Yoshida et al., 2005; Ding et al., 2014; Zhang et al., 2016), no study has focused on the genetic variability of the released rockfish and recipient stocks. A *S. schlegelii* stock enhancement activity was implemented in Lidao Bay in 2013. The present study was aimed to examine the genetic variability of the hatchery-released stock, wild stock before stock enhancement and the mixed stock after enhancement of black rockfish in Lidao Bay as well as the divergence among them. The results will provide basic data to ensure a healthy and sustainable restoration grant for *S. schlegelii*.

## MATERIALS AND METHODS

### Sample Collection and DNA Extraction

The black rockfish specimens evaluated in this study comprised a total of 288 fish from three different stocks. The first stock included 96 wild individuals [average total length (TL) = 12.74 ± 5.27 cm] in Lidao Bay sampled using a cage net in May, August, and November, 2013 (before release). The second sample group included 96 individual specimens (average TL = 9.48 ± 1.12 cm) sampled from the approximately 13,000 hatchery-reared *S. schlegelii* individuals released into Lidao Bay in November 2013. Finally, 96 samples from the mixed stock after release (average TL = 13.52 ± 4.68 cm) were collected in December 2013, and May, August and November 2014. As *S. schlegelii* has a short-distance seasonal movement, in order to obtain more representative samples, fishes before and after release were collected in different seasons. The sampling locations (122°34′30″–122°36′00″E, 37°14′00″–37°16′00″N) are shown in **Figure 1**. This study followed the guidelines for the experimental use of animals of the Institute of Oceanology, Chinese Academy of Sciences. The study was approved by the ethics committee of Institute of Oceanology, Chinese Academy of Sciences. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed by the authors.

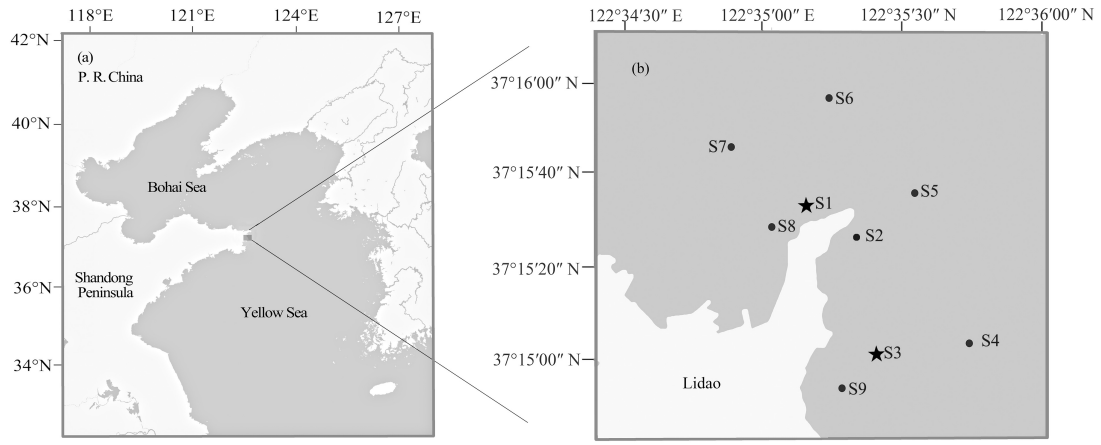


**FIGURE 1 |** Map of Lidao Bay showing sampling sites. Locations of Lidao Bay **(a)** and sampling sites **(b)**. S1–S9 indicating the sampling sites before and after release; S1 and S3 (asterisks) indicating the releasing sites.

Muscle sample from the dorsal region was excised from each fish specimen and immediately stored at −20°C. Total genomic DNA was extracted from 30 mg of muscle sample using the TIANamp Marine Animals DNA Kit (Tiangen Biotech Co., Ltd., Beijing, China). The concentration and purity of the extracted DNA were measured by using 1.0% agarose gel electrophoresis and a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, United States), and then stored at −20°C.

## mtDNA Sequencing

A pair of primers (SS-CR-F1: 5′-TAGTAGCTCAGCGTCAG AGCC-3′ and SS-CR-R1: 5′-GGGCCCATCTTAACATCTTCA-3′) designed according to the previously published *S. schlegelii* mitogenome (GenBank accession no. AY491978.1) was used for PCR amplification of the whole mtDNA control region. As there was an approximately 289 base pair (bp) tandem repeat situated within the 3′ end of the control region, one more pair of primers were designed (SS-CR-F2: 5′-CGTCAGAGCCCTGGTCTTGTA-3′ and SS-CR-R2: 5′-AAGCCAAAGGGGGTAATATAGA-3′) and used to obtain an approximately 800 bp fragment from the 5′ region of the mitochondrial control region.

The PCR amplification was performed in a 50 μL reaction containing 50–100 ng of genomic DNA as a template, 1 μL of KOD FX neo polymerase (Toyobo, Osaka, Japan), 25 μL of 2× PCR Buffer, 10 μL of dNTP mix (2 mM each), 2 μL (10 μM) of each primer, and ultrapure water to 50 μL. The cycling conditions consisted of an initial denaturation step at 94°C for 2 min, followed by 35 cycles of 98°C for 10 s, 55°C for 30 s, and 68°C for 1 min, with a final extension at 68°C for 7 min. The PCR products were purified with the EZNA$^{TM}$ gel extraction kit (Omega Bio-Tek, Norcross, GA, United States), and the purified PCR products were bidirectionally sequenced by Shanghai Sunny Biotechnology Co., Ltd. (Shanghai, China) using an ABI 3730 capillary sequencer.

## Microsatellite Genotyping

A total of 38 previously published *S. schlegelii* microsatellite loci were chosen for the analysis (**Table 1**). Primers were synthesized by Life Technology Co., Ltd. (Shanghai, China), and the forward primers were labeled with the fluorescent dyes FAM, HEX, and TAMRA. The PCR amplification was performed as described in the references by Yoshida et al. (2005), An et al. (2009), Bai et al. (2011), Yasuike et al. (2013), and Jia (2014). Amplified fragments were electrophoresed using an ABI 3130xl capillary DNA sequencer (Applied Biosystems, Inc., CA, United States). The ROX-500 size standard was used, and GENEMARKER software was used for scoring.

## Genetic Diversity Estimation

Sequences in the mtDNA control region were assembled and edited using Bioedit software (Hall, 1999), and aligned using ClustalX2.1 (Larkin et al., 2007). The haplotype diversity ($h$) and nucleotide diversity ($\pi$) indices and other genetic indices, such as polymorphic site and the number of haplotypes, were calculated with DnaSP 5.10 software (Rozas et al., 2003). Haplotype networks were constructed based on statistical parsimony using TCS 1.21 (Clement et al., 2000).

The standard genetic diversity parameters of microsatellite markers, including the number of alleles (*A*), effective number of alleles (*Ae*), observed heterozygosity (*Ho*), and expected heterozygosity (*He*) were calculated by using GENALEX v.6 (Peakall and Smouse, 2012). Allelic richness (*Ar*) was calculated by using FSTAT v.2.9.3 (Goudet, 2001). Micro-Checker software (Van Oosterhout et al., 2004) was used to search for null alleles at the loci. Deviations from linkage disequilibrium and Hardy–Weinberg equilibrium (HWE) for each microsatellite locus were tested by using the GENEPOP v.4.0 package (Saitoh and Rousset, 2008), and the significance was adjusted by applying sequential Bonferroni correction. The mean polymorphism information content (*PIC*) was calculated according to Botstein et al. (1980) by using PIC_Calc 0.6. ADZE v.1.0 (Szpiech et al., 2008), which was used to determine the number of private alleles with standardized sample sizes.

**TABLE 1 |** Information for microsatellite loci and primers used in this study.

| Locus | Repeat motif | Primer sequence (5′–3′) | Tm (°C) | Fluorescent label | References |
|---|---|---|---|---|---|
| RFs304 | $(TAA)_{16}$ | GAGTCGTACCTGTTGTGAACCA<br>GGTCATGTGACCTCAGATAGCA | 57 | FAM | Yasuike et al., 2013 |
| RFs307 | $(ACT)_{16}$ | GCTGACAGAGAGCTGTAACAGTG<br>TAGTTAAATACACACCGCACGC | 57 | HEX | |
| RFs309 | $(TTA)_{15}$ | CCGTGACAACATGATCAGAAAT<br>CGCAAATTTCAAAACGAATATG | 52 | TAMRA | |
| RFs312 | $(TTA)_{18}$ | CCACACAGGAAGTGGTACAAAA<br>ATTGCAGGTAGTTTGTCACAGC | 52 | FAM | |
| RFs315 | $(TAT)_{15}$ | TGTAACTTTGTTTTGAAAAGTGCTG<br>AAGGATGAGGGTAAAGGAGAGG | 57 | FAM | |
| RFs317 | $(ATT)_{15}$ | CCCATCACTGAATAAAGAAGCA<br>AGGTCTCTGTGGACTGAAGGAG | 57 | HEX | |
| RFs418 | $(TCTA)_{18}$ | TTTTTCAAAAGAAATTGGGCAT<br>GGAAAACAGATCCTTTCTGAACA | 52 | FAM | |
| RFs419 | $(ACGC)_{14}$ | ATGAGTTTGCTGTCATCACTGG<br>CATGTTAGCTGAATGGAAAGCA | 57 | HEX | |
| RFs421 | $(CAGA)_{18}$ | GGCGAGCCATCTAATAGTTGTT<br>TTTTTAGCAATAGCGACGAGAC | 57 | TAMRA | |
| RFs423 | $(TATC)_{18}$ | CACACTACTTGTTGAAGGGACG<br>GAACTATTGAAGGTCATTGTGAGC | 57 | TAMRA | |
| RFs527 | $(ATATA)_{10}$ | TAAAAATGCCCATAGTTGCAGA<br>ACCATGGAGTAACCCTTCTCAC | 57 | HEX | |
| RFs530 | $(ACAAT)_{10}$ | TGCAGATATTGGGTATGACTGG<br>ACAGCCAAAACTGGAACTCAAT | 57 | TAMRA | |
| Py2-12 | (AC)17 | TGACCAACAGGAAAATAC<br>TTGAAAGATGACCCATTA | 55 | FAM | Bai et al., 2011 |
| Py3-6 | $(TG)_{12}$ | ACGTATGTTGGCTGAAAC<br>TTTGGATAATGTGGCTTT | 55 | HEX | |
| Py3-29 | $(TGG)_5$ | TGCGGTGACTTATCCAGC<br>TCAACAAGGGAGCAAAGG | 55 | TAMRA | |
| Py3-41 | $(TG)_8AGT(GA)_5$ | CTGTTGGAGGGAGGTTAT<br>GTGTCTGGTTGAGCGAGT | 55 | FAM | |
| Py4-5 | $(AC)_{21}$ | ATGCACAGACAGAAATAC<br>TCGGATGATAATCAATAC | 54 | HEX | |
| Pyzj16 | $(AC)_5\ldots(CT)_9\ldots(CA)_5$ | ATTTCACAGCCTCGTTTAG<br>AGTAGGACAGGTGACTTCG | 57.8 | TAMRA | |
| KSs2A | $(TG)_{22}$ | CCCATAGCCTTGTTTACCT<br>TTTTGGTTATTGTCTTGGTTT | 55 | FAM | An et al., 2009 |
| KSs3 | $(TG)_{15}AG(TG)_2$ | TTGCCCACATCCTTTCT<br>ACTTCATGAATCCACTGACAT | 56 | HEX | |
| KSs5 | $(TG)_6TC(TG)_6$ | TGAAGCAATAAGGTAAAGGTG<br>GCAGAAGCCTCAGGAAAG | 61 | TAMRA | |
| KSs12B | $(CA)_{12}$ | TCATAAATGTGCTTAGTGAGG<br>AAGGGAATGCTAATGCTG | 55 | FAM | |
| KSs16 | $(TG)_{31}$ | TGTATTATGCCAATGAGGAGA<br>CAGTCGCACTTATTTTCCAG | 59 | HEX | |
| KSs17 | $(CA)_{16}$ | CCAAATAATAGCGAACACAC<br>ATGCAGAAGTGGTTAAAGTCT | 60 | TAMRA | |
| KSs18A | $(CA)_{11}$ | GCAGGGATCAATATCAACAA<br>ACATCAGCACATCACCTAATG | 56 | FAM | |
| KSs20 | $(CA)_5GA(CA)_9$ | ACAGACGCTTACTCACAAAAA<br>GAAGCGATTCCACAGGATA | 58 | HEX | |
| KSs26 | $(CA)_{12}$ | GGGGCTACATGTATGCTCA<br>AGACTGCGATACCTAGAAGGA | 55 | TAMRA | |
| HJ1-6 | $(GT)_{23}\ldots(TG)_6$ | TCAGAAAGGAGGCAAACG<br>CGGCTAATGTCCCACAAC | 60 | FAM | Jia, 2014 |
| HJ1-14 | $(AC)_{15}$ | AGGGAGATGTTCGACAAG<br>AGATTGGATGTAAGCGTG | 60 | HEX | |

*(Continued)*

**TABLE 1** | Continued

| Locus | Repeat motif | Primer sequence (5′–3′) | Tm (°C) | Fluorescent label | References |
|---|---|---|---|---|---|
| HJ3-23 | $(TG)_{13}\ldots(GCA)_{14}$ | GCAGCCCCTGACTTTGTT<br>TCGGTGCTCAGTGAAGGA | 64 | TAMRA | |
| HJ4-40 | $(CT)_6\ldots(CT)_1$ | CTTTACGGAGTGTACCTTG<br>TGACCCTTCTGGTCTGATT | 63 | FAM | |
| HJ4-94 | $(AC)_{13}$ | ACCCCACTTCAGAACACT<br>GAGCTCCGTACTGCATAT | 60 | TAMRA | |
| HJ5-1 | $(AG)_7C(GA)_{19}$ | ATACGCTCTGTATTCAACG<br>ACTTCCACATCAAATGTCC | 63 | FAM | |
| HJ5-46 | $(TG)_{15}\ldots(TG)_{10}$ | CCACTGGCAGATAAACGA<br>TTTTAACGGGCAGTTGTG | 63 | HEX | |
| Ssc12 | $(AC)_{20}$ | AACACGCTGAACAGAGAACAAA<br>GCTCCGACTATAGCTGGTCCTA | 59 | TAMRA | Yoshida et al., 2005 |
| Ssc23 | $(TG)_{21}$ | AGTGTCATGCCCTCTTCCAG<br>CACTCGGCATTCTCACCTCA | 57 | FAM | |
| Ssc51 | $(GT)_{20}T(TG)_5$ | GTGCTGATGGAAAACACTACCAG<br>CCTTTCCCTGAACACACTTGA | 57 | HEX | |
| Ssc69 | $(GT)_{15}$ | GGCACCGAGCTCAACCTTACTG<br>TGCTGTGACTATTTCCCTCTGGC | 57 | TAMRA | |

## Genetic Differentiation

To estimate the genetic subdivision, analysis of molecular variance (AMOVA) and the fixation index ($F_{ST}$) for both the mtDNA and microsatellite markers between each pair of sampling stocks were assessed using Arlequin 3.5.2 (Excoffier and Lischer, 2010). Neutrality tests (Tajima's *D* and Fu's *Fs*) were also performed using Arlequin 3.5. Pairwise genetic distances were calculated using the Kimura 2-parameter (K2P) distance model. Neighbor-joining (NJ) trees of K2P distances with 1000 bootstrap replications (Saitou and Nei, 1987) were generated to provide a graphic representation of the divergence among individuals. The K2P distance and NJ tree were calculated and generated using MEGA version 5 (Tamura et al., 2011).

Estimation of Nei's unbiased genetic distance and genetic identity of *S. schlegelii* samples was performed by using GENALEX v.6. The discriminant analysis of principal components (DAPC) was used to cluster genotypes independently of *a priori* haplotype designation using the R package adegenet v. 1.4.2 (Jombart, 2008). An optimum number of cluster was defined by the K-averaging algorithm that made use of the Bayesian information criterion.

Bottleneck 1.2 software (Cornuet and Luikart, 1996) was used to detect the existence of bottlenecks with the infinite allele model (IAM), stepwise-mutation model (SMM), and two-phased model of mutation (TPM) (Luikart and Cornuet, 1998). Significant heterozygosity excess was assessed with the Wilcoxon signed-rank test.

## Effective Population Size and Relatedness

Estimations of *Ne* in *S. schlegelii* stocks were performed with NeEstimator v.1.3 by using the linkage disequilibrium and heterozygote excess methods. *Ne* was also calculated by the sibship method (Wang, 2009) using COLONY (Jones and Wang, 2010). The relatedness of each stock was calculated with COANCESTRY v.1.0 (Wang, 2011) using the triadic likelihood method described by Wang (2007).

# RESULTS

## Genetic Diversity

A 678 bp fragment of the black rockfish mtDNA control region was amplified from all 288 individuals and sequenced. There were 59 polymorphic sites, and a total of 78 haplotypes were identified in the dataset. The genetic diversity indices are presented in **Table 2**. The values of $h$ and $\pi$ in the hatchery-released stock were 0.902 and 0.00483, respectively; whereas the $h$ values in the wild stock before stock enhancement and the mixed stock after enhancement were 0.970 and 0.939 ($p = 0.025$), and the $\pi$ values were 0.00581 and 0.00526 ($p = 0.150$), respectively. These results showed that the hatchery stock displayed slightly lower genetic diversity than the wild stock; however, there was almost no difference between the wild stock before release and the mixed stock after release. Among the 78 haplotypes, 8 were shared among all 3 stocks, and 7 haplotypes were only detected in the hatchery-released stock (**Supplementary Table S1**). Hap2 was shared among all stocks, accounted for 14.58% of all specimens. Tajima's *D* and Fu's *Fs* were significantly different from zero in the wild stock before enhancement ($D = -1.55$, $p = 0.77$ and $Fs = -25.82$, $p = 0.25$) and in the hatchery-released stock ($D = -0.19$, $p = 0.19$ and $Fs = -6.25$, $p = 0.14$).

For the 38 microsatellite loci, no evidence for scoring error caused by stuttering or large allele dropout was observed. A total of 654 alleles were found among the 288 individuals (**Supplementary Table S2**). The measures of genetic diversity for each stock, as calculated from the observed allele distribution, are presented in **Table 2**. The lowest polymorphism was observed at locus Kss18, which had four alleles, while the highest polymorphism was observed at locus HJ4-49, which had 47 alleles. For most loci in all stocks, the hypothesis of linkage

equilibrium was not rejected. Departures from HWE by exact test (Bonferroni correction applies) were observed at nine loci in the wild stock (Py2-12, Py4-5, Kss2, Kss3, Kss18, HJ1-14, HJ4-94, RFs315, and RFs423), and at eight loci in the hatchery stock (Py2-12, Kss16, Kss17, Kss18, HJ4-94, RFs304, RFs315, and RFs418) (**Supplementary Table S3**). The average *A* and *Ae* per stock per locus varied from 11.66 to 14.68 and 4.76 to 6.01, respectively. For most loci (30 out of 38), the *Ho* values were lower than the *He* values, which indicated heterozygote deficiency. The mean *Ae* in the hatchery-released stock was 4.76, the mean *PIC* was 0.674, and the *Ho* and *He* were 0.668 and 0.697, respectively. The *Ae*, *PIC*, *Ho*, and *He* values were 6.01, 0.698, 0.709, and 0.735 in the wild stock before stock enhancement and 5.82, 0.716, 0.741, and 0.754 in the mixed stock after enhancement, respectively. Consistent with the results from mtDNA, the analysis of microsatellite markers also showed that the level of genetic diversity in the wild stock was higher than that in the hatchery-released stock, and the level of genetic diversity in the mixed stock after enhancement was not significantly different from that in the wild stock before stock enhancement. The number of private alleles per stock, with a standardized sample size, is shown in **Figure 2**, and the value in the hatchery-released stock was the highest. There were 141 private alleles in the studied stocks. Among them, 38 were in the wild stock before release, 27 were in the hatchery-released stock, and 76 were in the mixed stock after release.

## Population Genetic Differentiation

The scatter plots from the DAPC and three-dimensional factorial correspondence analysis (3D-FCA) (**Figure 3**) clearly showed three major clusters (i.e., hatchery-released stock, before impact wild stock, and after impact stock). The genetic differences between the clusters identified by DAPC were evaluated by $F_{ST}$ values (**Table 3**). Pairwise $F_{ST}$ value between the wild stock before stock enhancement in Lidao Bay and the hatchery-released stock calculated from mtDNA was 0.01488 ($p = 0.072$) (**Table 3**). While for the microsatellite markers, the $F_{ST}$ was 0.040 ($p < 0.001$), which evidenced weak genetic differentiation between these two stocks. The pairwise $F_{ST}$ value obtained from mtDNA between the before- and after-impact stocks was small (−0.004), and the $p$ value of the exact test not significant ($p = 0.694$; **Table 3**). However, the estimated $F_{ST}$ from microsatellite DNA was 0.034 ($p < 0.001$; **Table 3**), which was larger than that from the mtDNA, and showed low but significant difference between the wild stock before stock enhancement and the mixed stock after release. The hierarchical analysis of AMOVA based on the mtDNA and microsatellite markers (**Table 4**) also showed that only 1.00% and 4.69% ($F_{ST} = 0.019$, $p < 0.001$) of the genetic variation, respectively,

**TABLE 2 |** Genetic diversity parameters of *S. schlegelii* stocks based on mtDNA control region sequences and 38 microsatellite loci.

| Stock | *N* | mtDNA CR | | | | | | | Microsatellite loci | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Haplotype | *h* | $\pi$ | *K* | Genetic distance | Tajima's *D* | Fu's *Fs* | *A* | *Ae* | *Ho* | *He* | *PIC* |
| HRS | 96 | 21 | 0.902 ± 0.016 | 0.00483 ± 0.00023 | 3.27171 | 0.00486 | −0.19031 | −6.25721 | 11.66 | 4.76 | 0.668 | 0.697 | 0.674 |
| WSB | 96 | 51 | 0.970 ± 0.007 | 0.00581 ± 0.00031 | 3.93640 | 0.00586 | −1.55225* | −25.81502 | 13.87 | 6.01 | 0.709 | 0.735 | 0.698 |
| MSA | 96 | 34 | 0.939 ± 0.010 | 0.00526 ± 0.00026 | 3.56288 | 0.00530 | −1.31746 | −22.10682 | 14.68 | 5.82 | 0.741 | 0.754 | 0.716 |
| Overall | 288 | 78 | 0.947 ± 0.006 | 0.00510 ± 0.00182 | 3.61445 | 0.00538 | −1.78355* | −25.66525 | 17.21 | 6.30 | 0.702 | 0.755 | 0.723 |

*HRS, hatchery released stock; WSB, wild stock before release; MSA, mixed stock after release; *p < 0.05.*

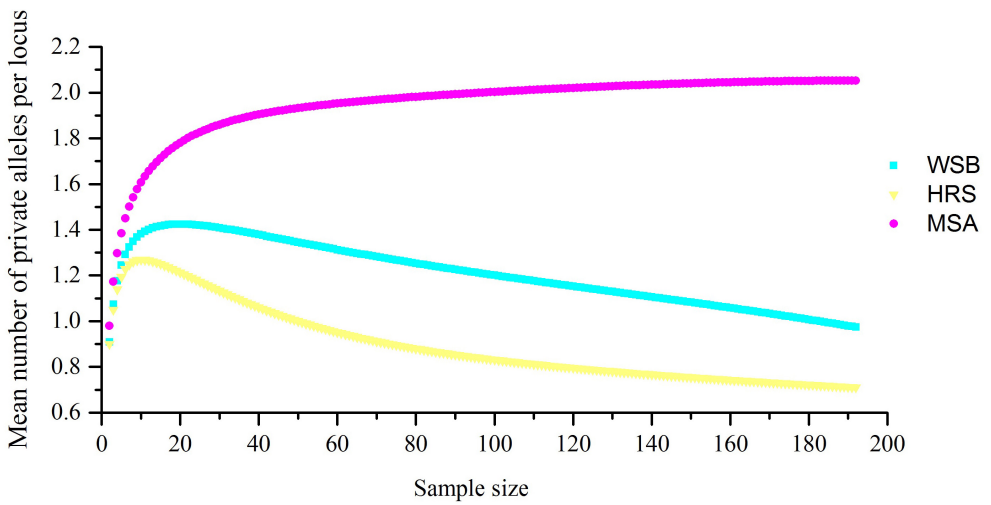


**FIGURE 2 |** The number of *S. schlegelii* private alleles present per locality when sample size was standardized.

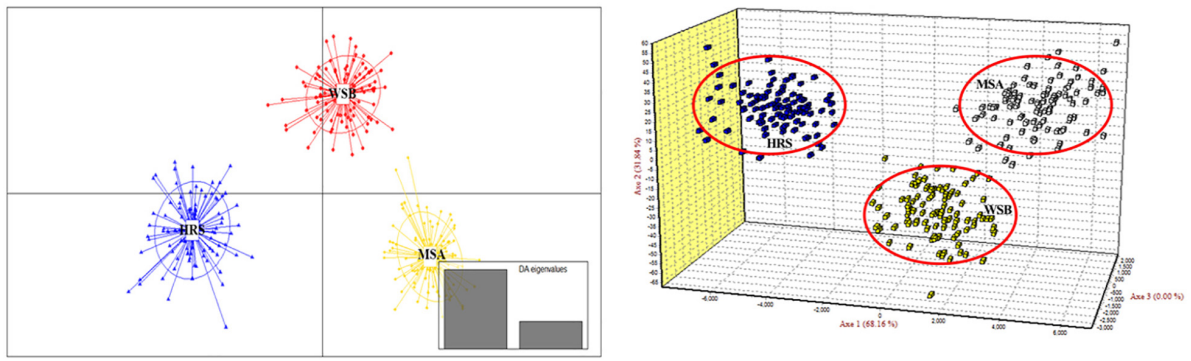


**FIGURE 3 |** Discriminant analysis of principal components (DAPC) and three-dimensional factorial correspondence analysis (3D-FCA) showing the relationships among *S. schlegelii* individuals of three populations based on 38 microsatellite markers.

were attributed to differences among the stocks, although the variance was significant.

The topology of the NJ tree showed no significant branch or cluster corresponding to the stocks (**Supplementary Figure S1**). The shallow phylogeny is consistent with demographic expansion after a bottleneck (Slatkin and Hudson, 1991). The network diagram of the haplotypes was simple, and no clustering that corresponded to sampling site was detected (**Figure 4**).

The IAM, TPM, and SMM models were used to test for population bottlenecks, and the results were shown in **Supplementary Table S4**. All stocks showed significant heterozygosity excess in the IAM model ($p < 0.05$), while in the TPM and SMM models, no significant population bottleneck was detected in any stock. Furthermore, a mode shift was detected in the frequency distribution of alleles, which resulted in a normal L-shaped curve.

## Effective Population Sizes and Relatedness

The *Ne* of each stock was reported in **Table 5**. The *Ne* estimates generated by different methods lacked consistency. When calculated by the linkage disequilibrium method, the *Ne* values of the hatchery-released, before-release wild, and after-release mixed stocks were 662.5, 1043.7, and 602.0, respectively. With the heterozygote excess method, the estimations for all stocks were infinite. *Ne* estimated based on the sibship method for the hatchery-released stock and mixed stock after release were 1013 and 2280, respectively, whereas the estimate for the wild stock before release was infinite. Although the *Ne* values largely varied among the methods, the results in the wild population were systematically higher than those in the hatchery-released population. The *Ne* of the mixed stock was lower than that of the wild stock before enhancement. Relatedness within the hatchery stock was globally higher than that within the wild stock, and value of the mixed stock after release was a little higher than that of the wild stock before stock enhancement (**Figure 5**).

**TABLE 3 |** Estimation of $F_{ST}$ of *S. schlegeli* stocks indicated from the mtDNA (below the diagonal) and microsatellite markers (above the diagonal).

| | HRS | WSB | MSA |
|---|---|---|---|
| HRS | – | 0.03980** | 0.06747** |
| WSB | 0.01488 | – | 0.03359** |
| MSA | 0.01999** | −0.00404 | – |

*HRS, hatchery released stock; WSB, wild stock before release; MSA, mixed stock after release. **p < 0.01.*

# DISCUSSION

The release of hatchery-reared juveniles could augment the marine fishery biomass. It could help fish stocks recover, and has been conducted in many countries. However, numerous studies have shown that the large-scale release of hatchery individuals into the wild can cause negative effects on the genetic diversity of wild populations (Laikre et al., 2010). Stock enhancement programs for black rockfish have been conducted in Asia for more than 30 years; however, scant information is available on its genetics. Therefore, more detailed genetic analyses elucidating the genetics of stock enhancement in this species were required. We conducted *S. schlegelii* stock enhancement analysis in 2013 at an embayment, namely Lidao Bay. Lidao bay is located in the east cost of Shandong Peninsula, China and open to the Yellow Sea. The surface seawater temperature ranged from −1 to 26°C, salinity ranged from 30.2 to 32.6‰, dissolved oxygen ranged from 7.10 to 11.36 mL/L, and pH ranged from 7.20 to 8.38 during our sampling period. This bay has some artificial reefs that may limit the dispersal of released *S. schlegelii* and may render small local native populations more vulnerable than those in the open sea. Thus, the possibility of genetic impact may be higher, which would offer a good opportunity for studying the genetic impacts of the black rockfish restoration. Moreover, in the present study, both mitochondrial and microsatellite makers were used to investigate the genetic effects.

Significant deviations from HWE were observed in all analyzed *S. schlegelii* stocks in the present study, even after sequential Bonferroni's correction. This may be caused by heterozygote deficits, and there was heterozygote deficiency in 30 of the studied loci. Although null allele is also an important

**TABLE 4 |** The AMOVA analysis of *S. schlegelii* based on mtDNA and microsatellite markers.

| Marker | Source of variation | df | Sum of square | Variance composition | Percentage of variation |
|---|---|---|---|---|---|
| mtDNA | Among populations | 2 | 7.087 | 0.01818 Va | 1.00 |
| | Within populations | 285 | 512.583 | 1.79854 Vb | 99.00 |
| | Total | 287 | 519.670 | 1.81671 | |
| | $F_{ST}$ | 0.01000 | | | |
| Microsatellite | Among populations | 2 | 293.497 | 0.68855 Va | 4.69 |
| | Among individuals within populations | 285 | 4145.979 | 0.55403 Vb | 3.77 |
| | Within individuals | 288 | 3870.500 | 13.43924 Vc | 91.54 |
| | Total | 575 | 8309.976 | 14.68181 | |
| | $F_{ST}$ | 0.04690 | | | |

contributor for the heterozygote deviation in marine fish (Dick et al., 2014), it is unlikely that null allele is the main cause of the deviations observed in this study, because only a few loci had a high frequency of null alleles (e.g., RFs315, Kss18).

The overall genetic diversity in *S. schlegelii* in this study was comparable to that previously reported using mtDNA and microsatellite markers (Zhang et al., 2016; Gao et al., 2018). Zhang et al. (2016) used the mitochondrial control region marker to investigate 221 *S. schlegelii* individuals collected from 13 localities across its entire range in China, Japan, and Korea, and the results revealed similar genetic variability at the haplotype and nucleotide levels (overall average $h = 0.91 \pm 0.01$ and $\pi = 0.0063 \pm 0.0092$). In the present study, the genetic diversity indices of the control region in all three stocks were similar to those measured in previous studies. The 38 microsatellite loci used in this study showed slightly lower levels of polymorphism compared to other black rockfish stocks (An et al., 2012; Gao et al., 2018). Many hatchery fishes show lower numbers of alleles and less genetic variability, resulting from the limited numbers of broodstock used as founders in the hatchery stock for juvenile production (Araki and Schmid, 2010; Champagnon et al., 2012; Lorenzen et al., 2012). In this study, although some of the genetic diversity indices for both mtDNA and msDNA were lower in the hatchery-reared fish than in the wild-born fish, most of the differences were slight and not supported statistically. Therefore, no large decrease of genetic diversity was observed between the hatchery stock and the recipient wild stock, consistent with what was observed in other studies of this fish (Han et al., 2016). However, even though the hatchery stock had high gene diversity, these fish had considerably lower allelic richness relative to the wild stock. According to Araki and Schmid (2010), losses of allelic richness are more common than losses of heterozygosity in the early generations of hatchery-reared populations. As the negative genetic impacts caused by stock enhancement have been documented in several fish species (Laikre et al., 2010), the change in the genetic variability of a mixed stock of *S. schlegelii*

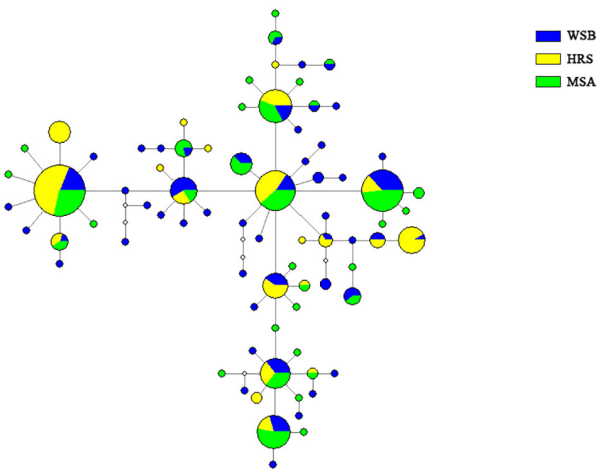


**FIGURE 4 |** Median-Network showing phylogenetic relationship among mtDNA control region haplotypes in *S. schlegelii* stocks of Lidao Bay.

**TABLE 5 |** Results of effective population sizes (*Ne*) for each stock.

| Stock | Linkage disequilibrium | Heterozygote excess | Sibship |
|---|---|---|---|
| HRS | 662.5 (553.6–822.1) | ∞ | 1013 (637–2165) |
| WSB | 1043.7 (831.1–1397.0) | ∞ | ∞ |
| MSA | 602.0 (527.1–700.5) | ∞ | 2280 (1278–8610) |

*HRS, hatchery released stock; WSB, wild stock before release; MSA, mixed stock after release.*

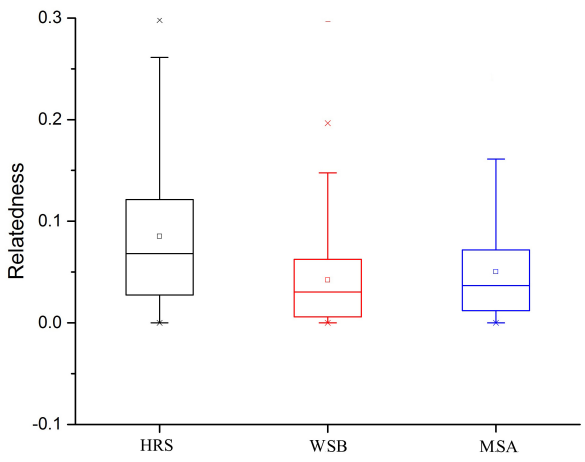


**FIGURE 5 |** Relatedness in the hatchery stock and wild stocks before and after release.

after release was also examined in this study. The results of both mtDNA and msDNA analyses showed that the mixed stock exhibited nearly similar average genetic diversity, albeit slightly lower, than that of the natural stock before release. It is important to note that even if the reductions in allelic diversity and heterozygosity are limited, such reductions could gradually erode the genetic variability of the wild stock by accumulating over generations.

Three clusters were identified by DAPC in this study. Low genetic differentiation level ($F_{ST} < 0.05$) was detected between the hatchery-released stock and the wild stock. The $F_{ST}$ and AMOVA results demonstrated that the wild stock before release and the mixed stock after release showed low but statistically significant levels of genetic differentiation. One possible reason for the low differentiation is that *S. schlegelii* in China has a relatively short domestication history, and it has not been selected in breeding programs over multiple generations. Indeed, most genitors used in artificial selective breeding were captured from natural waters and are only used once. Moreover, the black rockfish is a viviparous species, and the broodstock used to produce hatchery fish is renewed each year by sampling new genitors from the wild, such that a large number of genitors are used in a single year. These breeding practices may result in a high level of genetic diversity in the hatchery stock and low differentiation when compared with the wild stock. Another reason for the observed limited differentiation may be the absence of significant differentiation among different wild populations of this species. Black rock fish is a species that aggregates around drifting seaweed during their early development, and the both mtDNA and microsatellite marker analyses indicated no genetic divergence among different geographical populations and the existence of high gene flow (Wang et al., 2012; Zhang et al., 2016; Gao et al., 2018). The *S. schlegelii* adults spawn during May to July, and the larvae have a flexible feeding stage for 2–3 months (Nagasawa and Domon, 1997). Larvae of *S. schlegelii* might be transported by the China Coast Current. Therefore, although the swimming ability of adult *S. schlegelii* is limited, larval dispersal via drifting seaweed and oceanic circulation could promote genetic homogeneity (Gao et al., 2018). The lack of genetic differentiation in this species may reduce the genetic effects of stock enhancement using hatchery fish from different geographical stocks.

Estimates of *Ne* lack consistency and vary widely among estimation methods (**Table 5**). Although *Ne* is a key parameter for fecundity and supplementation of a population, exact estimation of the *Ne* based on genetic markers remains a challenge. Most methods for estimating the *Ne* were developed based on a hypothetical "ideal" population of constant size, with non-overlapping generations, an equal sex ratio, random mating and random variation in reproductive success (Waples and Do, 2010). These conditions may be not satisfied by black rockfish, which display complex life-history traits. Regardless, we observed that the *Ne* values obtained by using different estimators were systematically higher in the wild stock than in the hatchery-released stock and mixed stock after release.

There was no obvious reduction in genetic diversity in the hatchery-released stock compared with that of the wild stock, and no significant change in genetic diversity indices was detected in the local stock after enhancement in short term. Thus, our study demonstrated that black rockfish is an ideal candidate fish for releasing into the wild with an aim to increase fishery yields. However, although the observed reduction in genetic parameters, such as allelic diversity, heterozygosity and *Ne*, did not reach significance, a slight reduction could gradually affect the genetic diversity of the wild stock through accumulation over generations. Despite subtle genetic divergence was detected, our results also suggested that the stocks before and after release did not harbor the coincident genetic pool. Thus, the released individuals are likely to erode long-term genetic diversity of the wild stock through mating with wild individuals, and may have a lasting impact on population genetic structuring as time goes on. In this study, we only collected samples within the first year after release, and the released fishes did not reach sexual maturity. Therefore, more detailed surveys are required to elucidate the long-term genetic effects of *S. schlegelii* stock enhancement. Moreover, individuals sharing the same genetic architecture with the recipient population were recommended to use in subsequent enhancement activities to avoid adverse effects on the local gene pool.

## DATA AVAILABILITY STATEMENT

The mitochondrial dataset generated for this study can be found in the NCBI GenBank (https://www.ncbi.nlm.nih.gov/) with access numbers MN952238–MN952525. The microsatellite dataset analyzed in this study is included in the **Supplementary Data Sheet S1**.

## AUTHOR CONTRIBUTIONS

FY, HL, and LW conceived and designed the experiments. LW, ZW, and FY performed the experiments, analyzed the data, prepared the figures and tables, and wrote the draft of the manuscript. LW, YW, ML, and AS carried out the sampling work in the field. All authors revised the manuscript and approved the final version.

## FUNDING

This research was funded by the Major Agricultural Application Technology and Innovation Projects of Shandong Province (2017–2020), Special Program for Basic Research of the Ministry of Science and Technology of China (No. 2014FY110500), and the National Infrastructure of Fishery Germplasm Resource (No. 2017DKA30470).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2020.00094/full#supplementary-material

## REFERENCES


An, H., Kim, M. J., Park, K., Cho, K., Bae, B., Kim, J., et al. (2012). Genetic diversity and population structure in the heavily exploited korean rockfish, *Sebastes schlegelii*, in Korea. *J. World Aquac. Soc.* 43, 73–83. doi: 10.1111/j.1749-7345.2011.00544.x

An, H. S., Park, J. Y., and Kim, M. J. (2009). Isolation and characterization of microsatellite markers for the heavily exploited rockfish *Sebastes schlegelii*, and cross-species amplification in four related *Sebastes* spp. *Conserv. Genet.* 10, 1969–1972. doi: 10.1007/s10592-009-9870-8

Araki, H., and Schmid, C. (2010). Is hatchery stocking a help or harm?: evidence, limitations and future directions in ecological and genetic surveys. *Aquaculture* 308, S2–S11. doi: 10.1016/j.aquaculture.2010.05.036

Araki, H., Waples, R. S., Ardren, W. R., Cooper, B., and Blouin, M. S. (2007). Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. *Mol. Ecol.* 16, 953–966. doi: 10.1111/j.1365-294X.2006.03206.x

Bai, C. C., Liu, S. F., Zhuang, Z. M., Lin, L., Wang, H., and Chang, Y. Q. (2011). Isolation and characterization of microsatellite markers for the Korean rockfish, *Sebastes schlegelii*. *Genet. Mol. Res.* 10, 2065–2068. doi: 10.4238/vol10-3gmr1522

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

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

Champagnon, J., Elmberg, J., Guillemain, M., Gauthier-Clerc, M., and Lebreton, J. D. (2012). Conspecifics can be aliens too: a review of effects of restocking practices in vertebrates. *J. Nat. Conserv.* 20, 231–241. doi: 10.1016/j.jnc.2012.02.002

Chin, B. S., Nakagawa, M., Noda, T., Wada, T., and Yamashita, Y. (2013). Determining optimal release habitat for black rockfish, *Sebastes schlegelii*: examining growth rate, feeding condition, and return rate. *Rev. Fish. Sci.* 21, 286–298. doi: 10.1080/10641262.2013.837364

Clement, M., Posada, D. C. K. A., and Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659. doi: 10.1046/j.1365-294x.2000.01020.x

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.

Dick, S., Shurin, J. B., and Taylor, E. B. (2014). Replicate divergence between and within sounds in a marine fish: the copper rockfish (*Sebastes caurinus*). *Mol. Ecol.* 23, 575–590. doi: 10.1111/mec.12630

Ding, K., Zhang, H., Zhang, X. M., Song, N., and Gao, T. X. (2014). Comparative analysis of genetic variation in cultured and wild populations of black rockfish (*Sebastes schlegelii*) based on mitochondrial DNA control region. *J. Fish. China* 38, 769–777.

Excoffier, L., and Lischer, H. E. (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. doi: 10.1111/j.1755-0998.2010.02847.x

Gao, T., Ding, K., Song, N., Zhang, X., and Han, Z. (2018). Comparative analysis of multiple paternity in different populations of viviparous black rockfish, *Sebastes schlegelii*, a fish with long-term female sperm storage. *Mar. Biodivers.* 48, 2017–2024. doi: 10.1007/s12526-017-0713-4

Garibaldi, L. (2012). The FAO global capture production database: a six-decade effort to catch the trend. *Mar. Policy* 36, 760–768. doi: 10.1016/j.marpol.2011.10.024

Gold, J. R., Ma, L., Saillant, E., Silva, P. S., and Vega, R. R. (2008). Genetic effective size in populations of hatchery-raised red drum released for stock enhancement. *Trans. Am. Fish. Soc.* 137, 1327–1334. doi: 10.1577/t07-192.1

Goudet, J. (2001). *FSTAT, a Program To Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.* 3). Lausanne: Lausanne University.

Grant, W. S., Jasper, J., Bekkevold, D., and Adkison, M. (2017). Responsible genetic approach to stock restoration, sea ranching and stock enhancement of marine fishes and invertebrates. *Rev. Fish Biol. Fish.* 27, 615–649. doi: 10.1007/s11160-017-9489-7

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysisprogram for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hamasaki, K., Toriya, S., Shishidou, H., Sugaya, T., and Kitada, S. (2010). Genetic effects of hatchery fish on wild populations in red sea bream Pagrus major (Perciformes, Sparidae) inferred from a partial sequence of mitochondrial DNA. *J. Fish Biol.* 77, 2123–2136. doi: 10.1111/j.1095-8649.2010.02826.x

Han, C. H., Ma, H. T., Jiang, H. B., Liu, Y., Han, H. Z., and Wang, F. (2016). Isolation of microsatellite markers for black rockfish *Sebastes schlegelii* and genetic diversity of wild and cultured populations. *Oceanol. Limnol. Sin.* 47, 213–220.

Jia, C. F. (2014). *Development of Microsatellite DNA Markers and Population Genetics Assessment for the Black Rockfish, Sebastes schlegelii*. Shanghai: Shanghai Ocean University.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jones, O. R., and Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol. Ecol. Resour.* 10, 551–555. doi: 10.1111/j.1755-0998.2009.02787.x

Kang, K. M., and Shin, H. O. (2006). Movement ranges and routes of black rockfish *Sebastes schlegelii* in summer and autumn from acoustic telemetry. *Fish. Aquat. Sci.* 9, 91–96. doi: 10.5657/fas.2006.9.2.091

Kitada, S. (2018). Economic, ecological and genetic impacts of marine stock enhancement and sea ranching: a systematic review. *Fish Fish.* 19, 511–532. doi: 10.1111/faf.12271

Laikre, L., Schwartz, M. K., and Waples, R. S. (2010). Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends Ecol. Evol.* 25, 520–529. doi: 10.1016/j.tree.2010.06.013

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Leber, K. M. (2013). "Marine fisheries enhancement, coming of age in the new millennium," in *Sustainable Food Production*, eds. P. Christou, R. Savin, B. A. Costa-Pierce, I. Misztal, and C. B. A. Whitelaw (New York, NY: Springer Science), 1139–1157.

Lorenzen, K., Beveridge, M., and Mangel, M. (2012). Cultured fish: integrative biology and management of domestication and interactions with wild fish. *Biol. Rev.* 87, 639–660. doi: 10.1111/j.1469-185x.2011.00215.x

Lü, H., Zhang, X., Xi, D., and Gao, T. (2014). Use of calcein and alizarin red S for immersion marking of black rockfish *Sebastes schlegelii* juveniles. *Chin. J. Oceanol. Limnol.* 32, 88–98. doi: 10.1007/s00343-014-3022-9

Luikart, G., and Cornuet, J. M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.* 12, 228–237. doi: 10.2307/2387479

Luo, G., and Zhang, Z. D. (2014). Current situation of the development in stock enhancement of China. *China Fish.* 12, 37–39.

Nagasawa, T., and Domon, K. (1997). The early life history of kurosoi, *Sebastes schlegelii* (Scorpaenidae), in the sea of Japan. *Ichthyol. Res.* 44, 237–248. doi: 10.1007/bf02678703

Nakagawa, M. (2008). Studies on the stock enhancement technology of the black rockfish *Sebastes schlegelii*. *Bull. Fish. Res. Agency* 25, 223–287.

Nakagawa, M., Okouchi, H., Adachi, J., Hattori, K., and Yamashita, Y. (2007). Effectiveness of stock enhancement of hatchery-released black rockfish *Sebastes schlegelii* in Yamada Bay-evaluation by a fish market survey. *Aquaculture* 263, 295–302. doi: 10.1016/j.aquaculture.2006.10.02

Nihira, A., and Takashima, Y. (1999). Growth and movement of the rockfish, *Sebastes schlegelii*, released in the coastal waters of Kashimanada. *Bull. Ibaraki. Pref. Fish. Exp. Stn.* 37, 1–8. doi: 10.1016/j.jhazmat.2019.120861

Noda, T., Nakagawa, M., Nagakura, Y., and Okouchi, H. (2012). Migration of released black rockfish *Sebastes schlegelii* in the Miyako Bay analyzed using fish market survey data. *Aquac. Sci.* 60, 65–71.

Peakall, R., and Smouse, P. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Romo, M. D. M. O.-V., Aritaki, M., and Taniguchi, N. (2006). Pedigree analysis of recaptured fish in the stock enhancement program of spotted halibut *Verasper variegates*. *Fisheries Science* 72, 48–52. doi: 10.1111/j.1444-2906.2006.01115.x

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359

Ryman, N., Jorde, P. E., and Laikre, L. (1995). Supportive breeding and variance effective population size. *Conserv. Biol.* 9, 1619–1628. doi: 10.1046/j.1523-1739.1995.09061619.x

Ryman, N., and Laikre, L. (1991). Effects of supportive breeding on the genetically effective population size. *Conserv. Biol.* 5, 325–329. doi: 10.1016/0006-3207(92)90802-t

Saitoh, K., and Rousset, F. (2008). Genepop: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Saitou, N., and Nei, M. (1987). The neighbour-joining method: a new method for reconstructing evolutionary trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Slatkin, M., and Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.

Szpiech, Z. A., Jakobsson, M., and Rosenberg, N. A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24, 2498–2504. doi: 10.1093/bioinformatics/btn478

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Wang, J. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genet. Res.* 89, 135–153. doi: 10.1017/S0016672307008798

Wang, J. (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Mol. Ecol.* 18, 2148–2164. doi: 10.1111/j.1365-294x.2009.04175.x

Wang, J. (2011). COANCESTRY: a program for simulating, estimating and analyzing relatedness and inbreeding coefficients. *Mol. Ecol. Res.* 11, 141–145. doi: 10.1111/j.1755-0998.2010.02885.x

Wang, W. Q., Zhang, Y., Liu, M. X., Wu, Z. H., Wang, L. J., Zi, J., et al. (2012). Microsatellite analysis of genetic diversity in four wild stocks of black rockfish *Sebastes schlegelii*. *Mar. Sci.* 36, 10–16.

Waples, R. S., and Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol. Appl.* 3, 244–262. doi: 10.1111/j.1752-4571.2009.00104.x

Ward, R. D. (2006). The importance of identifying spatial population structure in restocking and stock enhancement programmes. *Fish. Res.* 80, 9–18. doi: 10.1016/j.fishres.2006.03.009

Xu, B. D., and Jin, X. S. (2005). Variations in fish community structure during winter in the southern Yellow Sea over the period 1985-2002. *Fish. Res.* 71, 79–91. doi: 10.1016/j.fishres.2004.07.011

Yasuike, M., Noda, T., Fujinami, Y., and Sekino, M. (2013). Tri-, tetra-and pentanucleotide-repeat microsatellite markers for the Schlegel's black rockfish *Sebastes schlegelii*: the potential for reconstructing parentages. *Conserv. Genet. Resour.* 5, 577–581. doi: 10.1007/s12686-012-9856-6

Yoshida, K., Nakagawa, M., and Wada, S. (2005). Multiplex PCR system applied for analysing microsatellite loci of Schlegel's black rockfish, *Sebastes schlegelii*. *Mol. Ecol. Note* 5, 416–418. doi: 10.1111/j.1471-8286.2005.00945.x

Zhang, H., Yanagimoto, T., Zhang, X., Song, N., and Gao, T. (2016). Lack of population genetic differentiation of a marine ovoviviparous fish *Sebastes schlegelii* in Northwestern Pacific. *Mitochondrial DNA* 27, 1748–1754. doi: 10.3109/19401736.2014.963797

Zhang, Y., Xu, Q., Alós, J., Liu, H., Xu, Q., and Yang, H. (2015). Short-term fidelity, habitat use and vertical movement behavior of the black rockfish *Sebastes schlegelii* as determined by acoustic telemetry. *PLoS One.* 10:e0134381. doi: 10.1371/journal.pone.0134381

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 30 June 2020
doi: 10.3389/fmars.2020.00516

# Genetic Population Structure of the Hard Clam *Meretrix meretrix* Along the Chinese Coastlines Revealed by Microsatellite DNA Markers

*Yingying Ye[1,2]*, Chengrui Yan[1], Wansuk Senanan[3], Baoying Guo[1,2], Kaida Xu[4] and Zhenming Lü[1,2]*

*[1] National Engineering Research Center for Marine Aquaculture, Zhejiang Ocean University, Zhoushan, China, [2] National Engineering Laboratory of Marine Germplasm Resources Exploration and Utilization, Zhejiang Ocean University, Zhoushan, China, [3] Department of Aquatic Science, Faculty of Science, Burapha University, Chon Buri, Thailand, [4] Scientific Observing and Experimental Station of Fishery Resources for Key Fishing Grounds, MOA, Key Laboratory of Sustainable Utilization of Technology Research, Marine Fisheries Research Institute of Zhejiang, Zhoushan, China*

OPEN ACCESS

**Edited by:**
*Christian Marcelo Ibáñez, Andres Bello University, Chile*

**Reviewed by:**
*Zhang Yuehuan, South China Sea Institute of Oceanology (CAS), China*
*Cristian B. Canales-Aguirre, Universidad de Los Lagos, Chile*
*Pablo A. Oyarzun, Andres Bello University, Chile*

***Correspondence:**
*Yingying Ye*
*yeyy@zjou.edu.cn*

**Specialty section:**
*This article was submitted to Marine Evolutionary Biology, Biogeography and Species Diversity, a section of the journal Frontiers in Marine Science*

**Received:** *04 November 2019*
**Accepted:** *04 June 2020*
**Published:** *30 June 2020*

**Citation:**
*Ye Y, Yan C, Senanan W, Guo B, Xu K and Lü Z (2020) Genetic Population Structure of the Hard Clam Meretrix meretrix Along the Chinese Coastlines Revealed by Microsatellite DNA Markers. Front. Mar. Sci. 7:516. doi: 10.3389/fmars.2020.00516*

The hard clam *Meretrix meretrix* is ecologically and economically important in the coastal regions of China. We evaluated the genetic diversity and population structure among eight *M. meretrix* samples from the Yellow Sea (YS) and South China Sea (SCS) using nine microsatellite DNA loci. Both conventional and model-based population genetic analyses suggested significant genetic divergence between YS and SCS regions (pairwise $F_{ST}$ values ranging from 0.014 to 0.056). Samples within each region were not genetically different, except for Zhanjiang which clearly differed from other the four SCS samples. Membership coefficients, estimated by STRUCTURE, suggested some genetic admixture of the two genetic clusters in ZJ. Population genetic structure was detected in SCS region. We detected moderate levels of genetic variation in all eight samples (mean $A$ = 16.111–22.111, mean $A_r$ = 14.512–19.029, mean $H_o$ = 0.736–0.843, mean $H_e$ = 0.823–0.868) and two genetic clusters (mean $A$ = 27.167–29.833, mean $A_e$ = 8.834–9.471, mean $A_r$ = 26.032–27.005, mean $H_o$ = 0.824–0.839, and mean $H_e$ = 0.821–0.850). Low levels of $N_e$ estimates were detected in *M. meretrix* populations. None of the genetic populations had signs of recent genetic bottlenecks. Knowledge on genetic variation and population structure of *M. meretrix* populations along the Chinese coasts will support the aquaculture management and conservation of *M. meretrix*, and will provide insights for stock selection in selective breeding programs for these species and delineating management units.

**Keywords: genetic diversity, population structure, microsatellite, *Meretrix meretrix*, management strategies**

## INTRODUCTION

The hard clam *Meretrix meretrix* (Linnaeus, 1758) is one of the largest clam species among the Venus-shells (Veneridae). It is an important marine bivalve with nutritive and medical value, and widely distributed in Korea, Japan, China, Vietnam, Thailand, India, and Saudi Arabian Gulf (Tang et al., 2006; Wang et al., 2006; Xie et al., 2012). Due to its high commercial value, the

price has increased from ¥10 to more than ¥100 per kilogram of some areas in China, the clam has widely been cultured in China since 1992 (Ho and Zheng, 1994; Chen et al., 2004; Li et al., 2011). With the advances of breeding technologies, aquaculture of *M. meretrix* in China has been rapidly expanding, cultural production has reached 350,000–400,000 tons per year since the beginning of the 21st century (Chen et al., 2004). The aquaculture industry still relies heavily on natural spat (Wang et al., 2011). With the rapid increase in demand, natural stocks of *M. meretrix* have declined dramatically. For the past few years, over-exploitation, damage to natural habitat and bacteria or viruses was serious problems that restrict the culture industry (Tang et al., 2006; Wang et al., 2019). In 2003, about 10,000 tons of adults and juveniles died in Rudong County, Jiangsu Province, resulting in a tremendous economic loss of ¥100,000,000 (Zhan, 2003). A large number of individuals died suddenly indicated that a bottleneck may occur by a sharp reduction in the size of the population (Lande, 1988). Over the last decade, selective breeding programs were developed to improve the *M. meretrix* industry, particularly to provide cultivated seed with better performance than wild seed (Liu et al., 2006). However, inappropriate breeding programs could lead to reduction of genetic variability, as a result of inbreeding associated with the use of a small number of parents to establish and maintain a strain (Yu and Li, 2007). Therefore, the genetic conservation of *M. meretrix* is essential for the sustainable management of natural resources and to increase aquaculture production (Lu et al., 2011a). Lu et al. (2011a) developed 33 SSR loci for *M. meretrix* and performed further the parentage determination for this species (Lu et al., 2011b). Zhu et al. (2012) conducted the genetic analysis among four strains of different shell colors of *M. meretrix* using microsatellite markers without population genetic analysis along China coast. Knowledge of genetic diversity and population structure is critical for understanding population dynamics, which is important for delineating management units and maintaining a sustainable fishery (Thorpe et al., 2000). It can be used to guide the selection of high-quality parents.

Population genetic structure of marine organisms results from multiple processes, numerous factors may affect gene flow in planktonic life stages by promoting or limiting either the mean or the variance of larval dispersal distance, and these effects may be more widespread than previously thought (Benzie, 2000). Factors effecting gene flow can be divided roughly into biological and physical categories, although the interactions between them may be most important (Jackson, 1986). In the present study, we employed polymorphic microsatellite loci to make an initial exploration of the factors that contribute to population genetic structure of *M. meretrix* in the Yellow Sea (YS) and South China Sea (SCS). The fishing activities are abundant in both seas and the cultivation of *M. meretrix* in YS has overwhelming superiority along Chinese coast based on the sandy ecological environment which is easy for clam inhabiting. Our objectives were to characterize the genetic diversity within and between populations, identify management units for this clam species, and estimate levels of gene flow among them. The results can aid the design of spatial aquaculture management and conservation strategies for hard clam inhabiting the YS and SCS.

# MATERIALS AND METHODS

## Sample Collections and Genomic DNA Preparation

A total of 374 individuals of *M. meretrix* were collected from eight coastal localities in the YS (Lianyungang: LYG, Yancheng: YC, and Nantong: NT) and SCS (Zhangzhou: ZZ, Zhuhai: ZH, Zhanjiang: ZJ, Haikou: HK, and Beihai: BH) areas (**Figure 1** and **Table 1**). All individuals were collected between October 2014 and May 2015. To avoid confusion among specimens of other coexisting clam species, all the individuals were identified by experienced fisheries researchers according to the identification key of bivalve mollusks. The feature of apex of shell toward the front without radial grain was used to identify *M. meretrix* from similar clams. Tissues from adductor muscle were dissected from fresh specimens and preserved in 95% ethanol and stored frozen at −20°C prior to genetic analysis.

Extraction of the total genomic DNA from each individual was performed using the salt-extraction procedure as described in Aljanabi and Martinez (1997) with slight modifications (i.e., 15 μl of 20 mg/ml protenase K and 10 μl of 10 mg/ml RNase were added and mixed well before overnight incubate). The quantity and quality of the isolated DNA was examined by NanoDrop 2000 spectrophotometer (Thermo Scientific) and by electrophoresis in agarose gel stained with SYBR Safe (Invitrogen Corp., Carlsbad, CA, United States); once extracted, DNA was stored in 1 × TE buffer, then quantified and diluted to approximately 20 ng/μl for further polymerase chain reactions (PCRs).

## Microsatellite Loci

We selected nine microsatellite loci developed for *M. meretrix* based on their genetic diversity information (i.e., high level of heterozygosity and number of alleles were more than 10). These loci included the MM08, MM10, MM11, MM14, MM15, MM19, MM22, MM25, and MM26 (Lu et al., 2011a,b). Forward primer of each pair was labeled with a fluorescent dye (6-FAM or HEX; Applied Biosystems, Foster City, CA, United States) at the 5'end.

Polymerase chain reaction was performed in a final volume of 15 μl containing 20–50 ng template DNA, 1.5 pmol of each forward and reverse primer, 0.25 mM dNTPs (Fermentas), 1 × PCR buffer (Fermentas), 3.75 mM $MgCl_2$, and 1 unit *Taq* DNA polymerase (Fermentas). Cycling conditions for all assays

**TABLE 1 |** Collection detail of clam *M. meretrix* samples.

| Sample name (abbr.) | Geographic locations | Sea region | Sample size | Collection date |
|---|---|---|---|---|
| Lianyungan (LYG) | Jiangsu (34°42′ N, 119°22′ E) | YS | 45 | May 2015 |
| Yancheng (YC) | Jiangsu (33°42′ N, 120°33′ E) | YS | 47 | Nov. 2014 |
| Nantong (NT) | Jiangsu (31°50′ N, 121°51′ E) | YS | 48 | Oct. 2014 |
| Zhangzhou (ZZ) | Fujian (24°02′ N, 117°52′ E) | SCS | 48 | Apr. 2015 |
| Zhuhai (ZH) | Guangdong (22°16′ N, 113°36′ E) | SCS | 47 | Oct. 2014 |
| Zhanjiang (ZJ) | Guangdong (20°55′ N, 110°34′ E) | SCS | 46 | Oct. 2014 |
| Haikou (HK) | Hainan (20°03′ N, 110°10′ E) | SCS | 47 | Nov. 2014 |
| Beihai (BH) | Guangxi (21°27′ N, 109°04′ E) | SCS | 46 | Oct. 2014 |

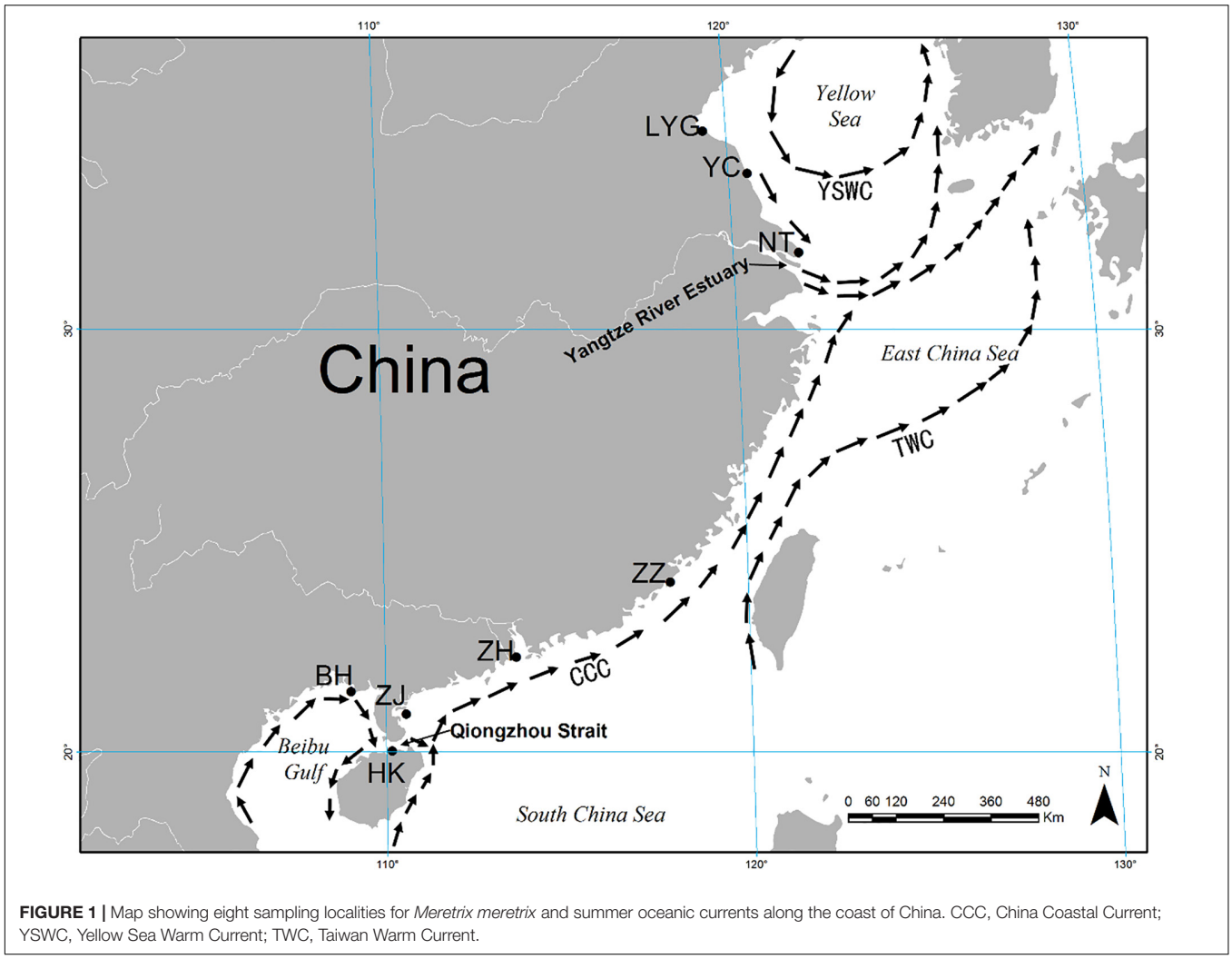


**FIGURE 1 |** Map showing eight sampling localities for *Meretrix meretrix* and summer oceanic currents along the coast of China. CCC, China Coastal Current; YSWC, Yellow Sea Warm Current; TWC, Taiwan Warm Current.

included an initial denaturation at 94°C for 3 min; 35 cycles of 94°C for 30 s (denaturation), 51–58°C for 30 s (annealing), and 72°C for 30 s (elongation) and a final elongation at 72°C for 5 min. PCR products were sent to a commercial genetic analysis service (Generay Biotech Co., Ltd., Shanghai, China) for electrophoresis and genotyping on an ABI Prism 3730 automatic DNA sequencer (Applied Biosystems). The alleles were scored using Gene Mapper software v5.0 (Applied Biosystems, Forest City, CA, United States) with a ROX 500 size standard, using an internal control for allele calling; each allele was coded by its size in nucleotides (base pairs).

## Microsatellite Analyses

Before population genetic data analysis, we used MICRO-CHECKER v2.2.3 software (Van Oosterhout et al., 2004) to double-check the effect of null alleles and allele scoring errors. We removed individuals with more than two missing loci and loci with null alleles present across most of the individuals (the final number of individuals used and microsatellite loci see section "Results").

We took two approaches to detect population genetic structure. First, we treated each sample as a separate unit and we performed conventional population genetic analysis [e.g., analysis of molecular variance (AMOVA), pairwise $F_{ST}$, and genetic distance measures]. Second, we performed a model-based clustering analysis. This latter approach does not presume genetic differentiation based on geographic locations, but instead it seeks for sets of multi-locus genotypes that minimize the departure from Hardy Weinberg equilibrium (HWE) and linkage disequilibrium (LD) (within set number $K$ genetic clusters).

For the conventional approach, first we partitioned the overall genetic variation to three sources of variation, namely, among samples, among individuals within samples, and within individuals using AMOVA implemented in ARLEQUIN v3.5 (Excoffier and Lischer, 2010). $P$-values were obtained through 1,000 permutations. The pairwise $F$-statistics ($F_{ST}$) value and an exact-test $P$-value were estimated to assess the level of genetic divergence using ARLEQUIN v3.5 (10,000 permutations) (Excoffier and Lischer, 2010). The multiple simultaneous tests of $F_{ST}$ values were adjusted using the sequential Bonferroni procedure (Rice, 1989). We generated a dendrogram illustrating

relationships among samples based on Cavalli-Sforza and Edwards (1967) genetic distance ($D_C$) by using the neighbor-joining algorithm (Saitou and Nei, 1987). The dendrograms based on multiple genetic distance matrixes (1,000 bootstrap replications), computed in the Microsatellite Analyzer software (Dieringer and Schlötterer, 2003) were constructed using NEIGHBOR and CONSENSE implemented in the PHYLIP v3.69 software package (Felsenstein, 2004). The heat map of pairwise $F_{ST}$ values and the plot of the NJ tree were constructed by the R package ggtree (Yu et al., 2017). The consensus tree was visualized by MEGA v7.0 (Kumar et al., 2016).

The model-based Bayesian clustering procedure in the program STRUCTURE v2.3.4 (Pritchard et al., 2000) was used to investigate the most likely number of clusters ($K$) within and among the eight locations and the level of admixture within a cluster. We initially simulated likelihood values for a range of predetermined $K$ values (i.e., 1–10; in our case) and examined the distribution of $\Delta K$, an *ad hoc* statistic based on the rate of change in the log probability of the data between successive $K$ values (Evanno et al., 2005). If population structure exists, the likelihood values of successive $K$ values should show some drastic changes. The most likely $K$ value is the modal value of this $\Delta K$ distribution. We performed 20 independent runs for each value of $K$ (the number of putative populations) from 1 to 10; each with 300,000 MCMC iterations, following a burn-in period of 50,000. We used the admixture model with correlated allele frequencies, and default parameter settings. The $\Delta K$ statistics were determined by STRUCTURE HARVESTER v0.6.94 (Earl and vonHoldt, 2012). Then, we examined the averaged proportion of membership of each cluster against our samples.

We analyzed genetic diversity within each sample as well as within each genetically discrete unit identified by the Bayesian approach ($K$). We assessed the microsatellite variability in terms of the number of alleles per locus ($A$), effective number of alleles per locus ($A_e$), observed ($H_o$) and expected ($H_e$) heterozygosities, and inbreeding coefficient ($F_{is}$) using the program GENALEX v6.5 (Peakall and Smouse, 2012). To standardize the number of alleles across all samples based on equal number of individuals per sample, allelic richness ($A_r$) was estimated based on the smallest sample size ($n$ = 40 and 135 for an assessment of the samples and the genetic clusters, respectively), using the program FSTAT v2.9.3 (Goudet, 2001).

GENEPOP v4.0 (Rousset, 2008) was used to determine for departure from HWE by using a Markov chain approximation (Guo and Thompson, 1992) with parameters set at 10,000 dememorizations, 100 batches, and 5,000 iterations per batch. We also used GENEPOP v4.0 (Rousset, 2008) to test for non-random association of alleles among loci (genotypic LD). Probability thresholds for the HWE and LD tests were adjusted for the number of simultaneous tests using the sequential Bonferroni correction (Rice, 1989).

The effective population size ($N_e$) in each sample based on the LD method was estimated by using NeESTIMATOR v2.0 (Do et al., 2014). We used 0.01 as the lowest allele frequency and putative 95% confidence intervals were calculated by using a parametric method (Do et al., 2014). The two-phase model (TPM) with 90% single-step mutations and 10% multiple-step mutations (1,000 replications) and the mode-shift test (Luikart et al., 1998) based on an L-shaped distribution of allele frequency under mutation-drift equilibrium were used to test for recent bottlenecks in each sample by using BOTTLENECK v1.2.02 (Piry et al., 1999). The significant heterozygote excess at each locus was used to determine the one-tailed Wilcoxon signed-rank test (Luikart and Cornuet, 1998).

# RESULTS

## Variation of Markers

All nine microsatellite loci examined were polymorphic in all individuals. A total of 324 alleles were observed across nine loci in eight *M. meretrix* samples (**Table 2**). The total number of alleles per locus ranged from 18 (MM08 and MM22) to 63 (MM14). At a locus level, the average number of alleles per locus ranged from 9.250 (MM22) to 29.625 (MM14), the average effective number of alleles per locus across all samples ranged from 2.130 (MM22) to 16.951 (MM19), and the mean allelic richness ranged from 7.967 (MM22) to 24.055 (MM14). The observed and expected heterozygosities averaged across samples ranged from 0.410 (MM22) to 0.970 (MM25) and from 0.500 (MM22) to 0.912 (MM14), respectively.

The observed genotype frequencies were tested for agreement with HWE (**Table 2**). After sequential Bonferroni correction for multiple tests (Rice, 1989), of the 72 sample-locus cases (eight samples × nine loci), 21 cases (LYG, YC, ZZ, ZH, and ZJ at MM08; LYC at MM10; ZJ at MM11; YC, ZZ, ZJ, and HK at MM15; all the samples except NT at MM19; YC, NT, and ZH at MM22) showed significant deviations from HWE ($P < 0.001$, after Bonferroni correction = 0.05/72). Locus MM19 had the highest frequency of HWE departures (seven out of eight tests at this locus). Most of the 21 deviations displayed significant heterozygote deficiencies ($H_o < H_e$). Micro-checker analysis (Van Oosterhout et al., 2004) for the presence of null alleles suggested that the possible presence of null alleles at MM19 in all eight samples, MM08 in seven samples (except NT), and MM15 in seven samples (except LYG). Therefore, we removed genotypes at MM08, MM15, and MM19 from further analyses that required HWE. After removing MM08, MM15, and MM19, no LD was found among locus pairs after Bonferroni correction.

## Genetic Differentiation Among Samples

Both conventional and Bayesian approaches suggested genetic divergence between two large coastal areas, YS (LYG, YC, and NT) and SCS (ZZ, ZH, ZJ, HK, and BH) (**Figures 2**, **3**). AMOVA revealed that variation among-regions contributed 6% ($P = 0.01$), among-samples within regions contributed 1% ($P = 0.01$), and within-population accounted for 93% ($P = 0.01$) of genetic variation. Pairwise $F_{ST}$ values between YS and SCS ranged from 0.014 to 0.056 (**Figure 2**). The lowest and highest divergences between YS and SCS populations were ZJ-NT and LYG-BH, respectively. Of 28 pairwise $F_{ST}$ values, 22 values were statistically different from zero ($P < 0.002$; $P$-value after adjusting for multiple comparisons = 0.05/28). Three samples from the YS were not genetically different. Pairwise $F_{ST}$ values revealed

**TABLE 2 |** Allelic variability at nine microsatellite loci in clam *M. meretrix* along the coasts of China.

| Pop. | | MM08 | MM10 | MM11 | MM14 | MM15 | MM19 | MM22 | MM25 | MM26 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LYG | *N* | 37 | 40 | 45 | 45 | 44 | 40 | 43 | 45 | 45 | 42.667 |
| | *A* | 11 | 14 | 19 | 38 | 36 | 22 | 9 | 14 | 19 | 20.222 |
| | $A_e$ | 5.901 | 8.377 | 6.716 | 19.104 | 18.009 | 16.244 | 3.196 | 8.672 | 6.661 | 10.320 |
| | $A_r$ | 10.799 | 13.384 | 16.585 | 32.013 | 30.773 | 20.691 | 7.877 | 13.123 | 16.728 | 17.997 |
| | $H_o$ | 0.541 | 0.650 | 0.933 | 0.978 | 0.932 | 0.500 | 0.628 | 0.956 | 0.600 | 0.746 |
| | $H_e$ | 0.831 | 0.881 | 0.851 | 0.948 | 0.944 | 0.938 | 0.687 | 0.885 | 0.850 | 0.868 |
| | $F_{is}$ | 0.349 | 0.262 | −0.097 | −0.032 | 0.013 | 0.467 | 0.086 | −0.080 | 0.294 | 0.140 |
| YC | *N* | 43 | 47 | 46 | 47 | 46 | 45 | 47 | 47 | 46 | 46.000 |
| | *A* | 9 | 14 | 23 | 35 | 32 | 25 | 11 | 14 | 21 | 20.444 |
| | $A_e$ | 4.037 | 8.305 | 7.185 | 16.798 | 18.241 | 16.071 | 2.624 | 8.106 | 5.475 | 9.649 |
| | $A_r$ | 8.419 | 13.027 | 19.233 | 29.188 | 27.185 | 22.389 | 9.858 | 12.742 | 18.215 | 17.806 |
| | $H_o$ | 0.535 | 0.936 | 0.978 | 0.957 | 0.696 | 0.511 | 0.383 | 1.000 | 0.891 | 0.765 |
| | $H_e$ | 0.752 | 0.880 | 0.861 | 0.940 | 0.945 | 0.938 | 0.619 | 0.877 | 0.817 | 0.848 |
| | $F_{is}$ | 0.289 | −0.064 | −0.136 | −0.018 | 0.264 | 0.455 | 0.381 | −0.141 | −0.090 | 0.104 |
| NT | *N* | 46 | 48 | 48 | 48 | 47 | 48 | 46 | 48 | 47 | 47.333 |
| | *A* | 13 | 15 | 21 | 37 | 28 | 31 | 9 | 14 | 24 | 21.333 |
| | $A_e$ | 5.377 | 8.084 | 8.811 | 19.361 | 15.835 | 18.656 | 2.469 | 7.972 | 6.565 | 10.348 |
| | $A_r$ | 11.391 | 13.362 | 17.967 | 29.581 | 23.848 | 26.751 | 8.035 | 12.498 | 20.27 | 18.189 |
| | $H_o$ | 0.761 | 0.896 | 0.938 | 1.000 | 0.787 | 0.854 | 0.435 | 1.000 | 0.915 | 0.843 |
| | $H_e$ | 0.814 | 0.876 | 0.887 | 0.948 | 0.937 | 0.946 | 0.595 | 0.875 | 0.848 | 0.858 |
| | $F_{is}$ | 0.065 | −0.022 | −0.058 | −0.054 | 0.160 | 0.097 | 0.269 | −0.143 | −0.079 | 0.026 |
| ZZ | *N* | 47 | 45 | 48 | 48 | 48 | 39 | 47 | 48 | 48 | 46.444 |
| | *A* | 14 | 19 | 17 | 24 | 23 | 30 | 10 | 15 | 15 | 18.556 |
| | $A_e$ | 6.979 | 8.581 | 10.402 | 7.642 | 6.286 | 21.574 | 2.145 | 4.655 | 9.253 | 8.613 |
| | $A_r$ | 12.147 | 16.472 | 15.581 | 18.589 | 17.726 | 27.754 | 8.067 | 12.914 | 13.969 | 15.913 |
| | $H_o$ | 0.596 | 0.956 | 0.854 | 0.958 | 0.583 | 0.513 | 0.511 | 0.979 | 0.833 | 0.754 |
| | $H_e$ | 0.857 | 0.883 | 0.904 | 0.869 | 0.841 | 0.954 | 0.534 | 0.785 | 0.892 | 0.835 |
| | $F_{is}$ | 0.305 | −0.082 | 0.055 | −0.103 | 0.306 | 0.462 | 0.043 | −0.247 | 0.066 | 0.090 |
| ZH | *N* | 47 | 46 | 47 | 47 | 47 | 40 | 47 | 47 | 47 | 46.111 |
| | *A* | 14 | 24 | 14 | 21 | 15 | 32 | 9 | 22 | 17 | 18.667 |
| | $A_e$ | 6.102 | 11.376 | 9.109 | 7.450 | 6.214 | 18.605 | 1.540 | 8.513 | 9.481 | 8.710 |
| | $A_r$ | 12.38 | 21.003 | 12.563 | 16.876 | 13.223 | 28.57 | 7.071 | 18.535 | 14.79 | 16.112 |
| | $H_o$ | 0.532 | 0.913 | 0.894 | 0.894 | 0.617 | 0.650 | 0.277 | 0.979 | 0.872 | 0.736 |
| | $H_e$ | 0.836 | 0.912 | 0.890 | 0.866 | 0.839 | 0.946 | 0.351 | 0.883 | 0.895 | 0.824 |
| | $F_{is}$ | 0.364 | −0.001 | −0.004 | −0.032 | 0.265 | 0.313 | 0.211 | −0.109 | 0.025 | 0.115 |
| ZJ | *N* | 46 | 46 | 46 | 46 | 46 | 42 | 46 | 46 | 46 | 45.556 |
| | *A* | 15 | 17 | 22 | 32 | 30 | 32 | 9 | 18 | 24 | 22.111 |
| | $A_e$ | 5.927 | 10.554 | 13.696 | 14.346 | 15.616 | 18.766 | 1.656 | 9.260 | 12.747 | 11.396 |
| | $A_r$ | 12.904 | 15.71 | 19.631 | 25.793 | 24.52 | 27.761 | 8.063 | 16.192 | 20.691 | 19.029 |
| | $H_o$ | 0.630 | 0.935 | 0.957 | 0.935 | 0.739 | 0.714 | 0.326 | 0.978 | 0.913 | 0.792 |
| | $H_e$ | 0.831 | 0.905 | 0.927 | 0.930 | 0.936 | 0.947 | 0.396 | 0.892 | 0.922 | 0.854 |
| | $F_{is}$ | 0.242 | −0.033 | −0.032 | −0.005 | 0.210 | 0.246 | 0.177 | −0.097 | 0.009 | 0.080 |
| HK | *N* | 44 | 47 | 47 | 47 | 47 | 34 | 44 | 47 | 47 | 44.889 |
| | *A* | 14 | 13 | 22 | 27 | 27 | 20 | 12 | 15 | 20 | 18.889 |
| | $A_e$ | 5.661 | 4.942 | 13.149 | 8.562 | 9.667 | 15.013 | 1.852 | 7.591 | 12.005 | 8.716 |
| | $A_r$ | 12.819 | 11.863 | 18.854 | 20.856 | 21.479 | 19.614 | 10.177 | 13.964 | 17.46 | 16.343 |
| | $H_o$ | 0.705 | 0.894 | 0.915 | 0.851 | 0.660 | 0.441 | 0.432 | 0.936 | 0.894 | 0.747 |
| | $H_e$ | 0.823 | 0.798 | 0.924 | 0.883 | 0.897 | 0.933 | 0.460 | 0.868 | 0.917 | 0.834 |
| | $F_{is}$ | 0.144 | −0.120 | 0.010 | 0.036 | 0.264 | 0.527 | 0.061 | −0.078 | 0.025 | 0.097 |
| BH | *N* | 43 | 45 | 46 | 46 | 43 | 31 | 45 | 45 | 46 | 43.333 |
| | *A* | 14 | 13 | 20 | 23 | 22 | 18 | 5 | 12 | 18 | 16.111 |
| | $A_e$ | 5.231 | 5.518 | 12.824 | 11.137 | 10.688 | 10.678 | 1.556 | 7.019 | 12.231 | 8.542 |

*(Continued)*

**TABLE 2 |** Continued

| Pop. | | MM08 | MM10 | MM11 | MM14 | MM15 | MM19 | MM22 | MM25 | MM26 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $A_r$ | 12.229 | 11.699 | 17.687 | 19.542 | 18.904 | 18 | 4.586 | 11.528 | 16.431 | 14.512 |
| | $H_o$ | 0.628 | 1.000 | 0.913 | 0.913 | 0.814 | 0.355 | 0.289 | 0.933 | 0.783 | 0.736 |
| | $H_e$ | 0.809 | 0.819 | 0.922 | 0.910 | 0.906 | 0.906 | 0.357 | 0.858 | 0.918 | 0.823 |
| | $F_{is}$ | 0.224 | −0.221 | 0.010 | −0.003 | 0.102 | 0.608 | 0.191 | −0.088 | 0.148 | 0.108 |
| Total | *A* | 18 | 27 | 37 | 63 | 50 | 52 | 18 | 23 | 36 | |
| Mean | *N* | 44.125 | 45.500 | 46.625 | 46.750 | 46.000 | 39.875 | 45.625 | 46.625 | 46.500 | |
| | *A* | 13.000 | 16.125 | 19.750 | 29.625 | 26.625 | 26.250 | 9.250 | 15.500 | 19.750 | |
| | $A_e$ | 5.652 | 8.217 | 10.237 | 13.050 | 12.570 | 16.951 | 2.130 | 7.724 | 9.302 | |
| | $A_r$ | 11.636 | 14.565 | 17.263 | 24.055 | 22.207 | 23.941 | 7.967 | 13.937 | 17.319 | |
| | $H_o$ | 0.616 | 0.897 | 0.923 | 0.936 | 0.728 | 0.567 | 0.410 | 0.970 | 0.838 | |
| | $H_e$ | 0.819 | 0.869 | 0.896 | 0.912 | 0.906 | 0.939 | 0.500 | 0.865 | 0.882 | |
| | $F_{is}$ | 0.248 | −0.035 | −0.031 | −0.026 | 0.198 | 0.397 | 0.178 | −0.123 | 0.050 | |

*Number of samples (N), number of alleles (A), effective number of alleles ($A_e$), allele richness ($A_r$), heterozygosity observed ($H_o$), heterozygosity expected ($H_e$), $F_{is}$ values, and probability of significant deviation from Hardy–Weinberg equilibrium (P) are given for each population and locus. Values underlined indicate statistical significance (P < 0.001, after Bonferroni correction = 0.05/72).*

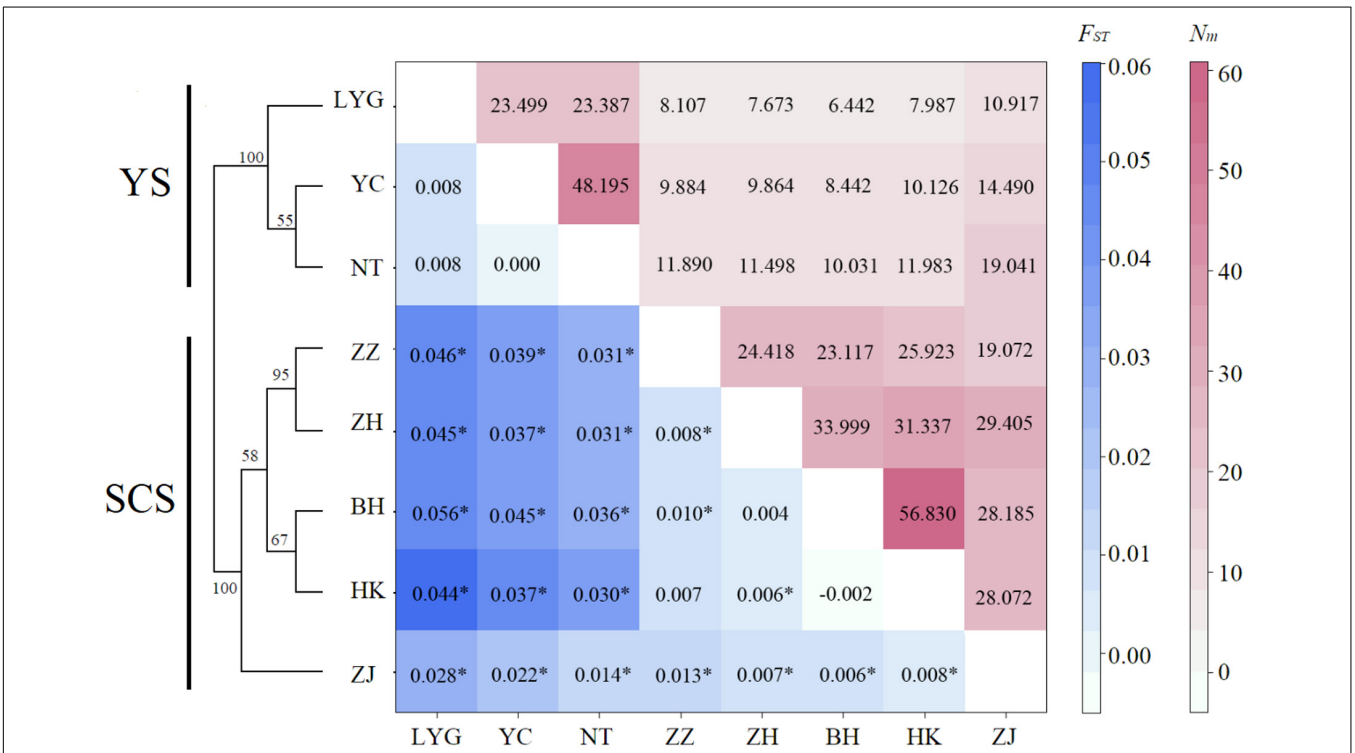


**FIGURE 2 |** Matrix showing pairwise differentiation estimates ($F_{ST}$) between eight *M. meretrix* samples based on six microsatellite loci (below the diagonal), and estimated number of migrants per generation ($N_m$) among samples (above the diagonal). Values with * indicate statistical significant difference from zero ($P < 0.002$, *P*-values adjusted for multiple comparisons using Bonferroni correction = 0.05/28). Neighbor-joining dendrogram based on Cavalli-Sforza and Edwards' genetic distance ($D_C$) values among eight *M. meretrix* samples from Chinese coastal areas based on six microsatellite loci Bootstrap values supporting each node indicate the percentage of phylogenetic tree sharing that node among 1,000 bootstrap replicates.

remarkable differentiation between YS samples and SCS samples. For the SCS samples, population subdivision was observed, with ZJ genetically distinct from the others. Pairwise $F_{ST}$ analyses also suggested that ZZ–ZH, ZZ–BH, and ZH–HK pairs were genetically different. The estimated number of migrant per generation ($N_m$), a measure of gene flow among samples, ranged from 6.442 to 19.041 (among YS and SCS) and from 19.072 to 56.830 (within YS and SCS), which suggested considerable gene flow within each coastal region (**Figure 3B**).

The neighbor-joining phylogenetic tree also suggested two major clusters (the two major clusters bootstrap value 100; **Figure 2**). Samples from the YS formed one basal cluster and another cluster consisted of samples from the SCS. However, ZJ was clearly differentiated from the other four SCS samples.

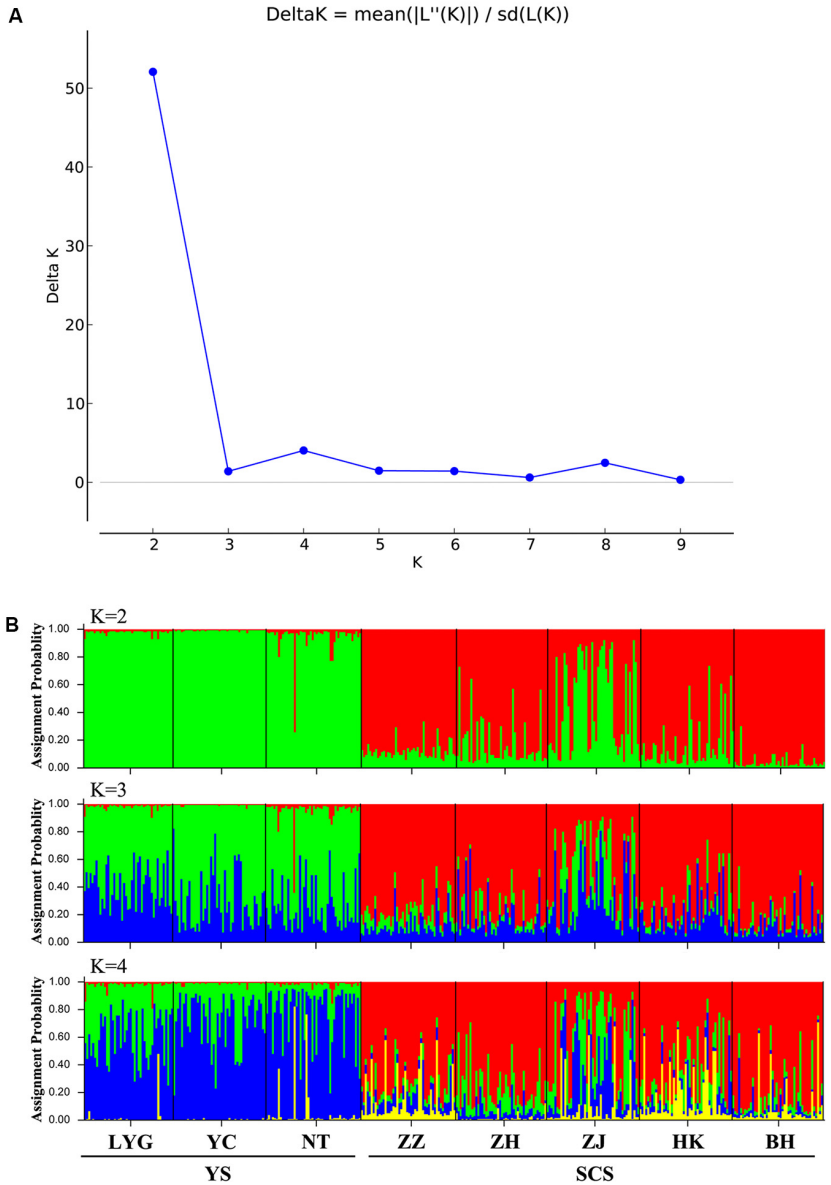


**FIGURE 3 | (A)** Magnitude of $\Delta K$ statistics as a function of the number of putative genetic clusters, $K$, for Chinese populations of *M. meretrix* based on six microsatellite loci. **(B)** Results of STRUCTURE analysis with membership coefficients of 374 *M. meretrix* individuals for $K$=2 collected from eight locations along the Chinese coast based on six polymorphic microsatellite loci.

The model-based STRUCTURE clustering analyses suggested two major genetic clusters (**Figure 3A**). The assignments of individuals within the clusters were consistent with geographic locations with samples in close proximity and within the same coastal areas being more genetically similar (**Figure 3B**). Within the SCS group, the ZJ sample had a distinct genetic composition from other samples and looked like a mixture of YS and SCS samples (**Figure 3B**).

## Genetic Diversity Within Populations

The two inferred clusters of genetic populations, YS (mean $A$ = 27.167, mean $A_e$ = 9.471, mean $A_r$ = 27.005, mean $H_o$ = 0.839, and mean $H_e$ = 0.850) and SCS (mean $A$ = 29.833, mean $A_e$ = 8.834, mean $A_r$ = 26.032, mean $H_o$ = 0.824, and mean $H_e$ = 0.821) had a comparable levels of genetic diversity (**Table 3**). Of 12 tests (two populations × six loci), five tests (YS at MM22; SCS at MM10, MM11, MM22, and MM25) significantly deviated from HWE ($P < 0.004$, after Bonferroni correction = 0.05/12). No LD was found among microsatellite locus pairs after Bonferroni correction.

The two inferred population clusters, YS and SCS, had similar effective population size ($N_e$) estimates with $N_e$ for YS and SCS being 93.4 (CI = 82.4–106.7) and 78 (CI = 70.0–87.6). Excluding ZJ, the $N_e$ estimate for SCS was 73.2 (CI = 64.7–83.2) (**Table 4**).

**TABLE 3 |** Allelic variability at six microsatellite loci in two genetic clusters of *M. meretrix*.

| Pop. | | MM10 | MM11 | MM14 | MM22 | MM25 | MM26 | Mean |
|---|---|---|---|---|---|---|---|---|
| YS | *N* | 135 | 139 | 140 | 136 | 140 | 138 | 138 |
| | *A* | 17 | 31 | 54 | 14 | 18 | 29 | 27.167 |
| | $A_e$ | 9.049 | 8.044 | 21.742 | 2.829 | 8.673 | 6.491 | 9.471 |
| | $A_r$ | 17.000 | 30.737 | 53.526 | 13.971 | 17.927 | 28.869 | 27.005 |
| | $H_o$ | 0.837 | 0.950 | 0.979 | 0.478 | 0.986 | 0.804 | 0.839 |
| | $H_e$ | 0.889 | 0.876 | 0.954 | 0.647 | 0.885 | 0.846 | 0.850 |
| | $F_{is}$ | 0.059 | −0.084 | −0.026 | 0.261 | −0.114 | 0.049 | 0.024 |
| SCS | *N* | 229 | 234 | 234 | 229 | 233 | 234 | 232.167 |
| | *A* | 27 | 30 | 50 | 18 | 23 | 31 | 29.833 |
| | $A_e$ | 8.530 | 12.779 | 10.100 | 1.744 | 7.911 | 11.939 | 8.834 |
| | $A_r$ | 24.262 | 26.343 | 42.680 | 15.530 | 20.568 | 26.807 | 26.032 |
| | $H_o$ | 0.939 | 0.906 | 0.910 | 0.367 | 0.961 | 0.859 | 0.824 |
| | $H_e$ | 0.883 | 0.922 | 0.901 | 0.427 | 0.874 | 0.916 | 0.821 |
| | $F_{is}$ | −0.064 | 0.017 | −0.010 | 0.140 | −0.100 | 0.062 | 0.008 |

*Number of samples (N), number of alleles (A), effective number of alleles ($A_e$), allele richness ($A_r$), heterozygosity observed ($H_o$), heterozygosity expected ($H_e$), $F_{is}$ values, and probability of significant deviation from Hardy–Weinberg equilibrium (P) are given for each population and locus. Values underlined indicate statistical significance, $P < 0.004$, after Bonferroni correction = 0.05/12.*

**TABLE 4 |** Estimates and 95% confidence intervals of effective population size ($N_e$) based on linkage disequilibrium and the detection of bottlenecks based on Wilcoxon's test for eight samples and the two genetic clusters of *M. meretrix* using six microsatellite loci.

| Pop. | $N_e$ estimates | 95% Confidence intervals | | Bottleneck test |
|---|---|---|---|---|
| | | Lower bound | Upper bound | TPM (*P*-value) |
| YS | 93.4 | 82.4 | 106.7 | 1.000 |
| SCS | 78.2 | 70.0 | 87.6 | 1.000 |
| SCS (without ZJ) | 73.2 | 64.7 | 83.2 | 1.000 |

*The allele frequency used in the Ne estimate was greater than 0.01.*

For the bottleneck tests, under the TPM, we did not detect significant heterozygote excess in other of the clusters (**Table 4**). Additionally, the mode-shift test showed a normal L-shaped distribution pattern of allele frequencies in all samples. The results implied the lack of bottleneck events in the recent histories of these populations.

# DISCUSSION

## Population Genetic Structure of *M. meretrix* in the Yellow Sea and South China Sea

The samples of *M. meretrix* from across the coast of China, excluding the East China Sea area, consists of two genetically distinct groups, the YS (LYG, YC, and NT) and SCS (ZZ, ZH, ZJ, HK, and BH) samples. This genetic division may be due to geographic isolation and oceanographic characteristics of these two areas. The planktonic larval duration (PLD) of *M. meretrix* is about 1 week with the major spawning period from April to September (Liu and Yang, 2005; Liu et al., 2006). Even though *M. meretrix* releases pelagic eggs and its pelagic larvae can be transported by strong oceanic currents, the distance between the two locations may serve as a geographical barrier. In addition, the freshwater outflow of the Yangtze River in spring and summer can decrease salinity levels at the river mouth located at the YS/East China Sea. The dilute plume of Yangtze River water can affect nutrient concentrations (Jiang et al., 2006), subsequently phytoplankton biomass (Lin et al., 2011) and hence survival of larvae probably during the reproductive season of *M. meretrix*, which may act as a dispersal barrier to larval transport between the two coastal areas. Other researchers had detected similar results in other mollusks across a similar geographic scale regardless of different PLD. These species include *Crassostrea gigas*, *Crassostrea ariakensis*, *Cellana toreuma*, *Scapharca broughtonii*, *Meretrix petechialis*, *Gomphina aequilatera*, and *Meretrix lamarckii* (Yu and Li, 2007; Xiao et al., 2010; Dong et al., 2012; Yu et al., 2015; Wang et al., 2017; Ye et al., 2018; Feng et al., 2020; respectively). For example, Dong et al. (2012) studied population genetics of *C. toreuma* (PLD similar to *M. meretrix*) along the coasts of China covering SCS, East China Sea, and YS. They found genetic differentiate of the YS and other populations. Both geographic distance and contemporary oceanographic conditions played important roles in shaping the north-south division of population structure in *C. toreuma*. The intense discharge from the Yangtze River strongly isolates the YS group from the East and SCS groups.

Interestingly, the results strongly suggested that ZJ is genetically distinct from the SCS population. This sample exhibits the lowest $F_{ST}$ values with samples from the YS compared to the other four samples from SCS. Membership coefficients, estimated by STRUCTURE, also suggested genetic admixture of the respective genetic clusters in this sample. It is possible the YS stock has been moved extensively to the SCS for aquaculture purposes. According to local farmers in Zhanjiang (ZJ), juveniles or broodstock used in this area typically are obtained from the largest breeding facility in China in Nantong (NT) located in the YS area. Hence, mixed individuals in ZJ cannot be divided truly from ZJ or NT leading a high membership in ZJ showed in Structure.

Low $F_{ST}$ values among samples within the YS and SCS may be explained by local oceanic currents. *M. meretrix* larvae may circulate within each coastal area on oceanic currents during the spawning season (i.e., April through September, Liu et al., 2006). There are two separate currents, the China Coastal Current (CCC) and the YS Warm Current (YSWC) (**Figure 1**). Summer ocean currents along the China coast (i.e., CCC and YSWC) are regular during April and September in the northern continental shelf of the SCS (Su, 2004). In the SCS, the CCC may facilitate the high level of gene flow among SCS samples. Similarly, under the influence of the YSWC, the three samples from the YS had relatively low and non-significant genetic differences ($F_{ST}$ values ranging from 0 to 0.008). Unfortunately, due to the scarcity of wild individuals by bedrock coast which is inhospitable for clam in the East China Sea region (Xu, 2009; Han, 2011), we failed to collect wild individuals there.

$F_{ST}$ values suggested low, but significant genetic differentiation among some samples within the SCS. Except

for ZJ, we detected low genetic differentiation between ZZ–ZH, ZZ–BH, and ZH–HK. This differentiation may be explained by the short PLD of this species (i.e., about 1 week). Similar to other species with short PLD, population genetic differentiation can occur over a small geographic scale, e.g., approximately 100 km in the surf clam *Mactra chinensis* in the YS region (Ni et al., 2015). The genetic differentiation in the SCS was inconsistent with geographic differences among of samples. The genetic homogeneity between BH–ZH and HK–ZZ may due to movements from BH to ZH and from HK to ZZ by aquaculture activities.

## Genetic Diversity of *M. meretrix* in the Yellow Sea and South China Sea Populations

The genetic diversity of a population is critical for its evolutionary potential. In the present study, six microsatellite markers revealed a high level of genetic diversity within hard clam *M. meretrix* within YS and SCS groups of populations (mean $A$ = 27.167–29.833, mean $A_e$ = 8.834–9.471, mean $A_r$ = 26.032–27.005, mean $H_o$ = 0.824–0.839, and mean $H_e$ = 0.821–0.850). The results suggest that wild populations of *M. meretrix* in China still contain considerable variation, which could be a valuable resource for a genetic breeding program. Also, this high variation can be arisen by the nature of the markers (i.e., microsatellite, highly polymorphics). The level of microsatellite variation detected in our study was a little higher than that detected in the parentage determination of *M. meretrix* by Lu et al. (2011b) (mean $A$, $H_o$, and $H_e$ were 13.9, 0.725, and 0.839, respectively, with three of five loci identical to those used in our study, i.e., MM11, MM14, and MM15). This result may because the parentage study used just a handful of families in genetic analysis. The genetic diversity in this study also was higher than that in the study of four strains of different shell colors and decorative patterns of *M. meretrix* using nine microsatellite markers (mean $A$ = 6.2–7.5, mean $H_o$ = 0.511–0.626, mean $H_e$ = 0.698–0.754) (Zhu et al., 2012). The first explanation may also reflect the genetic diversity from Zhu et al. (2012), due to sampling from same geographic location. The main object of Zhu et al. (2012) was to determinate the genetic difference between different color stains and the studied individuals are distributed in a smaller scale than our case that low genetic diversity is reasonable. The high level of genetic diversity in this study may be related to the wide distribution of *M. meretrix* studied. The lowest level of genetic diversity that we observed was in the BH sample (lowest levels of $A$, $A_e$, $A_r$, $H_o$, and $H_e$), which may be due to the Beibu Gulf being a semi-enclosed shallow water body with limited connectivity to the northern SCS through the Qiongzhou Strait (**Figure 1**).

Only small number of microsatellite tests (4/48 and 5/12 in all eight samples and two genetic clusters, respectively) exhibited significant deviations from HWE, which mostly was caused by heterozygote deficiency of microsatellite loci ($H_o < H_e$). We observed heterozygote deficiency at MM10 in LYG, MM11 in SCS, and MM22 in YC, NT, YS, and SCS. This may be due to the presence of null alleles at these two loci in those samples, and this pattern of deficiency did not occur at other loci. Null alleles are frequently reported at microsatellite loci in marine invertebrate species (Hedgecock et al., 2004). Null alleles can arise when mutations prevent primers from binding and can be avoided by redesigning primers (Callen et al., 1993). We also detected heterozygote excess at MM10 in SCS, MM11 in ZJ and MM25 in SCS ($H_o > H_e$).

## Management Implications

To ensure the long-term sustainability of wild marine stocks, management programs should recognize genetic diversity both within and among populations. This goal is also crucial for an aquaculture industry in China that relies entirely on the collection of natural seed. In this study, genetics data revealed two independent management units (i.e., YS and SCS groups). The two management units needs separate management strategies. Management actions, especially those relevant to translocation and propagation, should strive to avoid genetic mixing by managing each unit separately and by restricting transfers of seedstock or broodstock across units. Existing management practices may need to be adjusted.

Low levels of effective population size ($N_e$) were detected in *M. meretrix* populations. Low $N_e$ estimates indicate greater loss of genetic variation than what may be expected based on census size, especially of highly fecund and abundant species. The factors of low $N_e$ for *M. meretrix* may include heavy exploitation and annual mass mortality. Therefore, management should aim to increase $N_e$ for *M. meretrix* populations. Seedstock/broodstock should no longer be transferred from YS to the SCS. Our results and farmers' interviews suggested regular transfers of *M. meretrix* seeds/broodstock from the YS to the SCS. The National Marine Fisheries Service in China should monitor aquaculture activities and production, especially transplantation activities. Then, measures should be implemented to relieve some fishing pressure on this species and to apply appropriate mating schemes to seedstock used for restocking programs (Miller et al., 2003).

None of the samples and populations from *M. meretrix* had significant signs of recent genetic bottlenecks, with $P$-values always greater than 0.05 based on the TPM. Also, the lack of signs of bottleneck in our microsatellite data indicates that there was enough gene flow, in spite of this low level detected in this case, to maintain genetically diverse and genetically connected populations.

# CONCLUSION

In conclusion, we detected population genetic structure of *M. meretrix* in the SCS and the YS regions using nine microsatellite loci. We suggest separate management strategies for these two units to avoid genetic mixing by managing each unit separately and by restricting transfers of seedstock or broodstock across units.

## DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article and the **Supplementary Material**.

## ETHICS STATEMENT

All animal sampling and experiment procedures were reviewed and approved by the State Oceanic Administration of China and the Ethics Committee of Zhejiang Ocean University, and performed according to national laws and regulations.

## AUTHOR CONTRIBUTIONS

YY and WS designed the work, analyzed and acquired the data, interpreted the results, and wrote the manuscript. CY and KX collected samples and prepared genomic DNA, performed experimental works, and checked the data. BG and ZL designed the work, analyzed the data, wrote the manuscript, and provided funds for this study. All authors gave final approval for the publication of the article.

## FUNDING

This work was financially supported by the Fundamental Research Funds for Zhejiang Provincial Universities and Research Institutes (Nos. 2019J00021 and 2019J00038), the Natural Science Foundation of Zhejiang Province (No. LQ18D060004), and the Project of Bureau of Science and Technology of Zhoushan (Nos. 2020C21026 and 2019F12004).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2020.00516/full#supplementary-material

## REFERENCES

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Benzie, J. A. H. (2000). The detection of spatial variation in widespread marine species: methods and bias in the analysis of population structure in the crown of thorns starfish (Echinodermata: Asteroidea). *Mar. Genet.* 420:14.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C., et al. (1993). Incidence and origin of "null" alleles in the (AC) n microsatellite markers. *Am. J. Hum. Genet.* 52:922.

Cavalli-Sforza, L., and Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *Am. J. Hum. Genet.* 19:233.

Chen, L., Wang, X., and Chen, D. (2004). Research on the development of hard clam market in China mainland. *J. Shanghai Ocean Univ.* 3, 283–287.

Dieringer, D., and Schlötterer, C. (2003). *Microsatellite analyser* (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* 3, 167–169. doi: 10.1046/j.1471-8286

Do, C., Waples, R. S., Peel, D., Macbeth, G., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol. Ecol. Resour.* 14, 209–214. doi: 10.1111/1755-0998.12157

Dong, Y. W., Wan, H. S., Han, G. D., Ke, C. H., Zhan, X., Nakano, T., et al. (2012). The impact of Yangtze River discharge, ocean currents and historical events on the biogeographic pattern of Cellana toreuma along the China coast. *PLoS One* 7:e36178. doi: 10.1371/journal.pone.0036178

Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 35: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Felsenstein, J. (2004). *PHYLIP (Phylogeny Inference Package) Version 36 Distributed by the author*. Seattle: University of Washington.

Feng, J., Fu, Z., Guo, Y., Li, J., Guo, B., Lv, Z., et al. (2020). Strong genetic differentiation of the clam *Meretrix lamarckii* in the China Sea revealed by mitochondrial DNA marker. *Mol. Biol. Rep.* 47, 693–702. doi: 10.1007/s11033-019-05177-9

Goudet, J. (2001). *FSTAT, A Program To Estimate And Test Gene Diversities And Fixation Indices (Version 2.9.3.)*. Available online at: https://www2.unil.ch/popgen/softwares/fstat.htm (accessed December, 2015).

Guo, S. W., and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48, 361–372.

Han, Q. (2011). *Research on the Status of Tidal Resources in China Using remote Sensing Technology*. Master dissertation, Nanjing normal university, Nanjing.

Hedgecock, D., Li, G., Hubert, S., Bucklin, K., and Ribes, V. (2004). Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *J. Shellf. Res.* 23, 379–386.

Ho, J. S., and Zheng, G. X. (1994). *Ostrincola koe* (Copepoda, Myicolidae) and mass mortality of cultured hard clam (*Meretrix meretrix*) in China. *Hydrobiologia* 284, 169–173. doi: 10.1007/bf00006888

Jackson, G. A. (1986). Interaction of physical and biological processes in the settlement of planktonic larvae. *Bull. Mar. Sci.* 39, 202–212.

Jiang, C. B., Zhang, L. J., and Wang, F. (2006). A study of $pCO_2$ in the surface water of the southern yellow sea in summer: the respective contribution of the upwelling and the yangtze river diluted water to the air-sea of $CO_2$ flux. *Period. Ocean Univers. China* 36, 147–152.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 70 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lande, R. (1988). Genetics and demography in biological conservation. *Science* 241, 1455–1460. doi: 10.1126/science.3420403

Li, H. J., Liu, W. D., Gao, X. G., Zhu, D., Wang, J., Li, Y. F., et al. (2011). Identification of host-defense genes and development of microsatellite markers from ESTs of hard clam *Meretrix meretrix*. *Mol. Biol. Rep.* 38, 769–775. doi: 10.1007/s11033-010-0165-4

Lin, J., Zhu, J., Zhang, J., Wu, H., and Luo, W. (2011). Distribution of phytoplankton biomass and its relationships with the environmental factors off the Changjiang (Yangtze River) Estuary. *J. Fish. China* 35, 74–87.

Linnaeus, C. (1758). *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, Cum Characteribus, Differentiis, Synonymis, Locis. Editio Decima, Reformata [10th Revised Edition]*, vol. 1, 824. Laurentius Salvius: Holmiae. Available online at: https://biodiversitylibrary.org/page/726886

Liu, B., Dong, B., Tang, B., Zhang, T., and Xiang, J. (2006). Effect of stocking density on growth, settlement and survival of clam larvae *Meretrix meretrix*. *Aquaculture* 258, 344–349. doi: 10.1016/j.aquaculture.2006.03.047

Liu, S., and Yang, A. (2005). *Healthy Aquaculture Technology of Marine Shellfish in China*. Beijing: China Ocean Press. 234–243.

Lu, X., Wang, H., Dai, P., and Liu, B. (2011a). Characterization of EST-SSR and genomic-SSR markers in the clam, *Meretrix meretrix*. *Conserv. Genet. Resour.* 3, 655–658. doi: 10.1007/s12686-011-9426-3

Lu, X., Wang, H., Liu, B., and Xiang, J. (2011b). An effective method for parentage determination of the clam (*Meretrix meretrix*) based on SSR and COI markers. *Aquaculture* 318, 223–228. doi: 10.1016/j.aquaculture.2011.05.017

Luikart, G., Allendorf, F., Cornuet, J., and Sherwin, W. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* 89, 238–247. doi: 10.1093/jhered/89.3.238

Luikart, G., and Cornuet, J. M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.* 12, 228–237. doi: 10.1046/j.1523-1739.1998.96388.x

Miller, L. M., Kapuscinski, A. R., and Hallerman, E. (2003). *Genetic Guidelines For Hatchery Supplementation Programs. Population Genetics: Principles And Applications For Fisheries Scientists*. Bethesda, MA: American Fisheries Society.

Ni, G., Li, Q., Ni, L., Kong, L., and Yu, H. (2015). Population subdivision of the surf clam Mactra chinensis in the East China Sea: Changjiang River outflow is not the sole driver. *PeerJ* 3:e1240. doi: 10.7717/peerj.1240

Peakall, R., and Smouse, P. E. (2012). GenAlEx 65: genetic analysis in excel population genetic software for teaching and research an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Piry, S., Luikart, G., and Cornuet, J. M. (1999). Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225. doi: 10.1111/j.1558-5646.1989.tb04220.x

Rousset, F. (2008). Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Su, J. (2004). Overview of the South China Sea circulation and its influence on the coastal physical oceanography outside the Pearl River Estuary. *Contin. Shelf Res.* 24, 1745–1760. doi: 10.1016/j.csr.2004.06.005

Tang, B., Liu, B., Wang, G., Zhang, T., and Xiang, J. (2006). Effects of various algal diets and starvation on larval growth and survival of *Meretrix meretrix*. *Aquaculture* 254, 526–533. doi: 10.1016/j.aquaculture.2005.11.012

Thorpe, J., Solé-Cava, A. M., and Watts, P. C. (2000). Exploited marine invertebrates: genetics and fisheries. *Hydrobiologia* 420, 165–184. doi: 10.1007/978-94-017-2184-4_16

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Wang, G., Liu, B., Tang, B., Zhang, T., and Xiang, J. (2006). Pharmacological and immunocytochemical investigation of the role of catecholamines on larval metamorphosis by β-adrenergic-like receptor in the bivalve *Meretrix meretrix*. *Aquaculture* 258, 611–618.

Wang, H., Huan, P., Lu, X., and Liu, B. (2011). Mining of EST-SSR markers in clam *Meretrix meretrix* larvae from 454 shotgun transcriptome. *Genes Genet. Syst.* 86, 197–205. doi: 10.1266/ggs.86.197

Wang, X., Kong, L., Chen, J., Matsukuma, A., and Li, Q. (2017). Phylogeography of bivalve *Meretrix petechialis* in the northwestern pacific indicated by mitochondrial and nuclear dna data. *PLoS One* 12:e0183221. doi: 10.1371/journal.pone.0183221

Wang, Y., Wang, M., Wang, B., Liu, M., Jiang, K., Hou, X., et al. (2019). A preliminary attempt to explore the potential functions of a tetraspanin gene (MmTSPAN) in the innate immunity of hard clam *Meretrix meretrix*: sequence features and expression profiles. *Fish Shellf. Immunol.* 88, 135–141. doi: 10.1016/j.fsi.2019.01.048

Xiao, J., Cordes, J. F., Wang, H., Guo, X., and Reece, K. S. (2010). Population genetics of *Crassostrea ariakensis* in Asia inferred from microsatellite markers. *Mar. Biol.* 157, 1767–1781. doi: 10.1007/s00227-010-1449-x

Xie, W., Chen, C., Liu, X., Wang, B., Sun, Y., Yan, M., et al. (2012). *Meretrix meretrix*: active components and their bioactivities. *Life Sci. J.* 9, 756–762.

Xu, J. (2009). *The CSS Evalution Of The Coastal Tourism Resources Of Zhejiang Province*. Master dissertation, Zhejiang Normal University, Jinhua.

Ye, Y., Fu, Z., Tian, Y., Li, J., Guo, B., Lv, Z., et al. (2018). Pelagic larval dispersal habits influence the population genetic structure of clam *Gomphina aequilatera* in China. *Genes Genom.* 40, 1213–1223. doi: 10.1007/s13258-018-0718-9

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. Y. (2017). Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210x.12628

Yu, H., Gao, S., Chen, A., Kong, L., and Li, Q. (2015). Genetic diversity and population structure of the ark shell *Scapharca broughtonii* along the coast of China based on microsatellites. *Biochem. Syst. Ecol.* 58, 235–241. doi: 10.1016/j.bse.2014.12.003

Yu, H., and Li, Q. (2007). Genetic variation of wild and hatchery populations of the Pacific oyster *Crassostrea gigas* assessed by microsatellite markers. *J. Genet. Genom.* 34, 1114–1122. doi: 10.1016/s1673-8527(07)60127-1

Zhan, W. Y. (2003). Study on the cause of death of *Meretrix meretrix* in large area in Rudong Jiangsu. *Mar. Fish. Res.* 24, 62–65.

Zhu, D., Dong, Y., Lin, Z., and Yao, H. (2012). Genetic analysis among four strains of different shell colors and decorative patterns of *Meretrix meretrix* using microsatellite markers. *J. Fish. China* 36, 202–209.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Advantages of publishing in Frontiers

**OPEN ACCESS**
Articles are free to read for greatest visibility and readership

**FAST PUBLICATION**
Around 90 days from submission to decision

**HIGH QUALITY PEER-REVIEW**
Rigorous, collaborative, and constructive peer-review

**TRANSPARENT PEER-REVIEW**
Editors and reviewers acknowledged by name on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** frontiersin.org/about/contact

**REPRODUCIBILITY OF RESEARCH**
Support open data and methods to enhance research reproducibility

**DIGITAL PUBLISHING**
Articles designed for optimal readership across devices

**FOLLOW US**
@frontiersin

**IMPACT METRICS**
Advanced article metrics track visibility across digital media

**EXTENSIVE PROMOTION**
Marketing and promotion of impactful research

**LOOP RESEARCH NETWORK**
Our network increases your article's readership