

## Article

# Population Genomics of Commercial Fish *Sebastes schlegelii* of the Bohai and Yellow Seas (China) Using a Large SNP Panel from GBS

Beiyan Zhu <sup>1</sup>, Tianxiang Gao <sup>1</sup> , Yan He <sup>2</sup>, Yinquan Qu <sup>1,\*</sup>  and Xiumei Zhang <sup>1</sup>

<sup>1</sup> Fishery College, Zhejiang Ocean University, Zhoushan 316022, China; ziyu9911@163.com (B.Z.); gaotianxiang0611@163.com (T.G.); xiumei1227@163.com (X.Z.)

<sup>2</sup> MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao 266003, China; yanhe@ouc.edu.cn

\* Correspondence: qyquan@njfu.edu.cn

**Abstract:** *Sebastes schlegelii* is one of the most commercially important marine fish in the northwestern Pacific. However, little information about the genome-wide genetic characteristics is available for *S. schlegelii* individuals from the Bohai and Yellow Seas. In this study, a total of 157,778, 174,480, and 188,756 single-nucleotide polymorphisms from Dalian (DL), Yantai (YT), and Qingdao (QD) coastal waters of China were, respectively, identified. Sixty samples (twenty samples per population) were clustered together, indicating shallow structures and close relationships with each other. The observed heterozygosity, expected heterozygosity, polymorphism information content, and nucleotide diversity ranged from 0.14316 to 0.17684, from 0.14035 to 0.17145, from 0.20672 to 0.24678, and from  $7.63 \times 10^{-6}$  to  $8.77 \times 10^{-6}$ , respectively, indicating the slight difference in genetic diversity among *S. schlegelii* populations, and their general genetic diversity was lower compared to other marine fishes. The population divergence showed relatively low levels (from 0.01356 to 0.01678) between *S. schlegelii* populations. Dispersing along drifting seaweeds, as well as the ocean current that flows along the western and northern coasts of the Yellow Sea and southward along the eastern coast of China might be the major reasons for the weak genetic differentiation. These results form the basis of the population genetic characteristics of *S. schlegelii* based on GBS (Genotyping by Sequencing). In addition to basic population genetic information, our results provide a theoretical basis for further studies aimed at protecting and utilizing *S. schlegelii* resources.

**Keywords:** *Sebastes schlegelii*; population genomics; GBS; SNPs; genetic differentiation; genetic diversity



**Citation:** Zhu, B.; Gao, T.; He, Y.; Qu, Y.; Zhang, X. Population Genomics of Commercial Fish *Sebastes schlegelii* of the Bohai and Yellow Seas (China) Using a Large SNP Panel from GBS. *Genes* **2024**, *15*, 534. <https://doi.org/10.3390/genes15050534>

Academic Editor: Manuel Vera Rodriguez

Received: 17 March 2024

Revised: 22 April 2024

Accepted: 23 April 2024

Published: 24 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

*S. schlegelii* belongs to the genus *Neohispaniscus* (Scorpaenidae), widely distributed in the northwestern Pacific [1]. In China, *S. schlegelii* mainly inhabits the Bohai Sea to the Yellow Sea and the East Sea of China, usually spending the majority of its life amid reef-associated water, gravel, and caves [2]. *S. schlegelii* is an economical fish species with delicious and nutritious flesh, which is widely appreciated by consumers in China, Republic of Korea, and Japan [3], but it also has an ecological value as a demersal carnivorous fish [4]. Adults and those in the early-life stages constitute an important link in the transfer of energy and matter within the pelagic and benthic trophic webs as well as in the benthopelagic coupling of the region [5]. Due to the advantages of fast growth, a low temperature tolerance, and high stress resistance, *S. schlegelii* is an important marine culture and commercial fishing fish in the northwestern Pacific [6]. The aquaculture production of *S. schlegelii* can reach up to 33,020 metric tons per year [7].

Given the increasingly prominent issues of overfishing and changes in the environment of inshore fishing grounds, the fishery resources of *S. schlegelii* have been gradually decreased, and the genetic diversity of wild populations has been seriously affected [8].

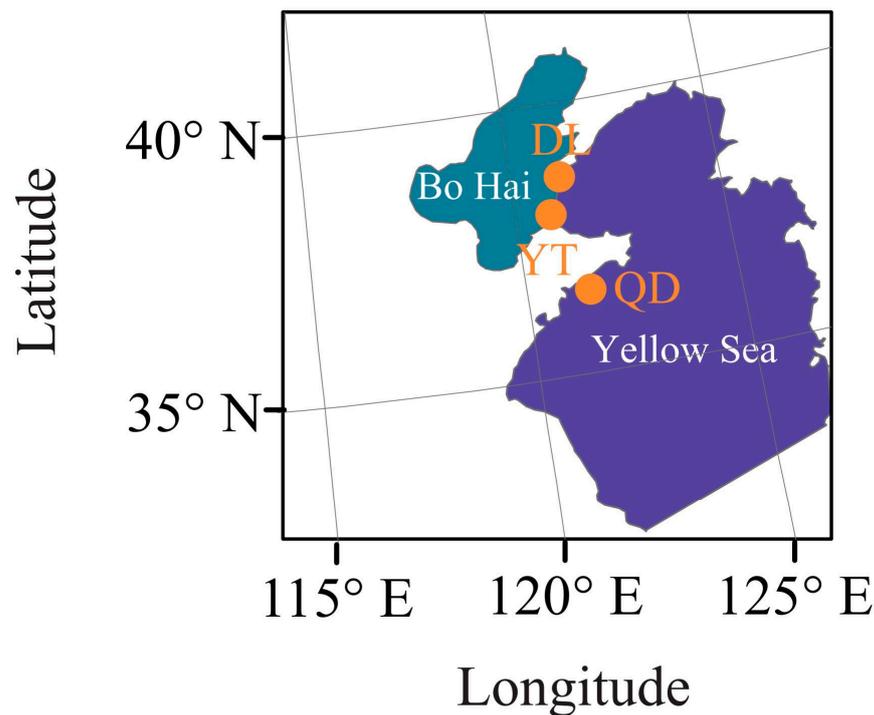
There are important economic and social benefits to breeding high-quality varieties with fast growth rates and high disease resistance from wild populations for artificial breeding. For instance, the quantity of wild fishery resources can be increased to a certain extent by adding artificially bred seedlings upon the transfer to natural waters for stock enhancement [9], so as to restore fishery resources, protect biological diversity, maintain ecological balance, and promote the sustainable development of fishery [10]. Hence, stock enhancement has become an important measure for the recovery and management of offshore fishery resources in China [11]. However, the inflow of artificially bred seedlings may affect the genetic structure, diversity, and divergence of the original wild population, and unscientific breeding may also have a negative impact on the wild resource population [12].

For the process of artificial breeding with seedlings of marine fishes, one of the key issues is to ensure the genetic diversity of parents or seedlings for stock enhancement. Therefore, molecular biotechnology has been widely used to ensure the genetic population structure, diversity, and divergence of the wild populations [13–15]. Previous studies have explored the genetic diversity and structure of *S. schlegelii* among different populations (Bohai Sea, Yellow Sea, and Japan Sea) based on various genetic markers. For instance, Zhang et al., found no significant population genetic structure of *S. schlegelii* corresponding to sampling locations in China, Japan, and Republic of Korea using a mitochondrial DNA sequence (mtDNA) control region [16,17]. Another two investigations using microsatellites and amplified fragment length polymorphism (AFLP) markers both showed no significant genetic differentiation among populations of *S. schlegelii* from China and Japan, indicating a single gene pool among populations [8,18]. To date, the population structure and genetic divergence based on genome-wide markers of the *S. schlegelii* species remain poorly reported. In our study, we sequenced the genomes of *S. schlegelii*, including 60 accessions from three coastal waters (Yantai, Dalian, and Qingdao) of China. Population structure, genetic diversity, and genetic divergence of the *S. schlegelii* populations were investigated. Candidate genomic loci or genes involved in some important functions, such as the regulation of DNA recombination, ribosomal large subunit assembly, and mitochondrial genome maintenance, have been identified in *S. schlegelii* through genomic selective sweep analyses. The single-nucleotide polymorphism (SNP) panel and derived genomic analyses could serve as a reference for fishery resource protection, improved species breeding, and the stock enhancement of *S. schlegelii*.

## 2. Materials and Methods

### 2.1. Sample Collection and DNA Extraction

To explore the population structure and genetic diversity of the *S. schlegelii* population, we sequenced 60 samples from 2 geographic areas distributed in the Yellow Sea (20 samples for Qingdao (QD) with sampling codes QD1 to QD20) and Bohai Sea (20 samples for Yantai (YT) with sampling codes YT1 to YT20, and Dalian (DL) with sampling codes DL1 to DL20, respectively) of the northwestern Pacific covering the DL (38°70' N, 121°16' E), YT (37°88' N, 120°76' E), and QD (36°21' N, 120°77' E) coastal waters of China (Figure 1). DL and YT coastal waters are located in the north temperate zone, with no extreme heat in summer and no severe cold in winter. The annual variation in water temperature in QD coastal waters is greater than that in YT and DL coastal waters, in which the climate is characterized by cold and dry winters and warm and humid summers. The sampling locations are in close proximity to each other (Figure 1). Fresh muscles on the right side of the back of the fish were preserved in liquid nitrogen and stored at −80 °C for further experiment. Total DNA was extracted from muscle tissues using a standard phenol–chloroform method [19]. The DNA concentration and purity (OD260/OD280 ratio) were determined by using a NanoDrop spectrophotometer 2000C (Thermo Scientific, Waltham, MA, USA), and DNA integrity was confirmed by 1.0% agarose gel electrophoresis. The qualified DNA precipitation was dissolved in sterilized ddH<sub>2</sub>O for further library construction [20].



**Figure 1.** Collection locations of three populations of *S. schlegelii*.

### 2.2. GBS Library Construction, Illumina Sequencing, and Read Mapping

The restriction endonucleases HaeIII and EcoRI were used to digest the genome DNA. First, a P1 adapter, sequencing primer, and barcode were added to enzyme digestion products together with a ligase. Second, a P2 adapter was added to the sheared and pooled DNA fragments with a ligase. Third, the adaptor-ligated DNA sequences were amplified using polymerase chain reaction (PCR). The AxyPrep DNA gel extraction kit (AxyGEN, Union City, CA, USA) was selected to obtain the products with a 200 bp fragment size, and 1 ng/μL concentration was prepared for subsequent analyses. The insert sizes were detected by an Agilent 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA) and the effective concentration of the libraries was quantified by qPCR (effective concentration > 2 nm).

The required fragments (200–400 bp) were selected for library construction. All libraries were sequenced based on the Illumina (San Diego, CA, USA), HiSeq™2500 (2 × 150 bp read length). GBS libraries and the sequencing itself were carried out by Novogene Co., Ltd. (Beijing, China). Finally, raw data for each sample were generated, and the quality of raw reads was checked by FastQC v0.11.5 [21]. The poor-quality sequences (<average Q30), short sequences (<50 bp), primer sequences, and adapter sequences were removed and excluded from the raw reads.

High-quality sequencing data from each sample were aligned to a reference genome (the assembled genome can be obtained by assembly ID CNA0000824 [22]) by using BWA v0.7.17 (Burrows-Wheeler Aligner) software with the settings “mem-t 4-k 32-M” [23]. The alignment reads with Sequence Alignment/Map (SAM) format were extracted and converted into Binary Alignment/Map (BAM) format, and then sorted according to genomic positions by SAMtools v1.9 with a maximum of 1000 reads at a position per BAM file [24]. BAM files had PCR duplicates removed using Picard v2.24.0 (MarkDuplicates) with the settings “REMOVE\_DUPLICATES = true”, and unmapped reads and secondary alignments were removed using SAMtools with the “-r” option. Additionally, a minimum mapping quality (Q) score threshold of “-q 30” was applied to filter out low-quality alignments in BAM files using SAMtools.

### 2.3. Variant Calling, Population Structure, Diversity, and Divergence

The variant calling was carried out by using GATK v.4.0.3.0 [25] following the best practices workflow [26]. The general variants were identified for each individual using GATK HaplotypeCaller, and then combined by GenotypeGVCFs function to a single variant calling file. This two-step approach was carried out to ensure variant accuracy, which included re-genotyping and quality recalibration in the combined vcf file. SNPs were then identified using samtools/bcftools with default parameters based on alignments of all Illumina short reads. SNPs were filtered according to the following parameters: (1) SNPs were only present in one of the two pipelines (SAMtools/BCFtools and GATK); (2) SNPs with a read depth more than 1000 or less than 5; (3) non-biallelic SNPs; (4) SNPs with a missing rate more than 40%; (5) SNPs in repeat regions; (6) SNPs were removed if the distance with nearby variant sites was less than 5 bp.

A phylogenetic gene tree was constructed based on SNPs in the single-copy genes regions. Two popular programs (RAxML v8.2.11 [27] with the GTRCAT model and IQ-Tree v1.6.12 [28] with the self-estimated best substitution model) were applied to construct the maximum likelihood (ML) tree. The PCA (principal component analysis) was performed using GCTA software v1.94.1 [29]. The ancestral population structure among individuals from the Bohai and Yellow Seas was estimated from the number of ancestral population units ( $K = 2-4$ ) by using ADMIXTURE software v.1.3.0. Admixture analysis was performed following the parameter standard errors, which estimated by bootstrapping (bootstrap = 200). To investigate the genetic diversity and divergence among the two geographical areas, we evaluated the observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), polymorphism information content ( $PI_C$ ), nucleotide diversity ( $\pi$ ), and population divergence ( $F_{ST}$ ) based on the filtered SNP information using the program Stacks v2 and PLINK v1.90 [30,31].

### 2.4. Selection Pressure Analysis

To identify selective sweeps, Tajima's D (which accounts for allelic variation) with a 100 kbp sliding window on the larger SNP data set was estimated using VCFtools v0.1.17 [32]. Three populations (QD, DL, and YT) from two geographic areas (Bohai and Yellow Seas) were, respectively, selected for selective analysis. The gene ontology (GO) functional annotation analysis was performed using the clusterProfiler R package v3.14.0 [33] with  $pvalueCutoff = 0.05$ ,  $pAdjustMethod = "BH"$ , and  $TERM2GENE = term2gene$ .

## 3. Results

### 3.1. Sequencing and Variants Calling

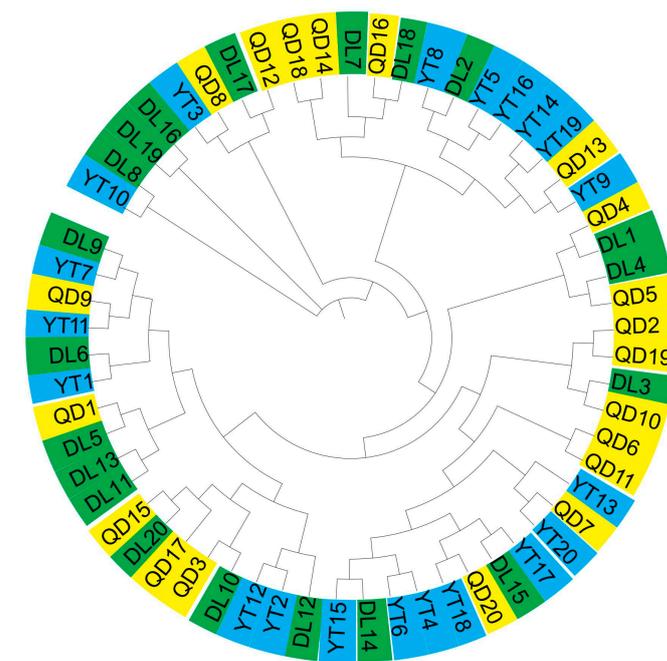
In this study, 12.81 Gb of clean data (1,841,342 reads, 87.12% of raw reads with 2,113,568 reads) was generated after strict filtration and normalization with an average of 213.48 Mb of clean data per sample. A total of 199,811, 222,675, and 239,152 variants were, respectively, identified from DL, YT, and QD populations, including 157,778, 174,480, and 188,756 SNPs, as well as 42,033, 48,195, and 50,396 InDels (Insert/Deletion). Among them, SNPs from 60 samples were employed for analyses of genetic diversity and divergence and genome-wide selection pressure. After filtration, 149,434 SNPs were applied to the population structure analyses of the three populations (QD, DL, and YT) (Table S1). The ratio of transition to transversion ( $T_S/T_V$ ) in SNPs was 2.48. We also identified that the three populations contained 107,709, 120,457, and 129,432 SNPs in intergenic regions, 88,180, 98,059, and 105,455 in introns, and 4444, 4810, and 4948 SNPs in exons, including 2377, 2690, and 2662 being synonymous, respectively (Table 1).

**Table 1.** A summary of the SNPs and indels identified in the three populations of *S. schlegelii*.

	DL	QD	YT	
Samples	20	20	20	
SNPs	157,778	174,480	188,756	
Indels	42,033	48,195	50,396	
Intergenic	107,709 (68.27%)	120,457 (69.04%)	129,432 (68.57%)	
Intron	88,180	98,059	105,455	
Exon	4444	4810	4948	
	Synonymous	2377 (53.49%)	2690 (55.93%)	2662 (53.80%)
	Non-synonymous	2067 (46.51%)	2120 (44.07%)	2316 (46.20%)
Splicing	628	644	703	
Upstream	31,285	36,050	38,776	
Downstream	31,507	35,416	37,971	

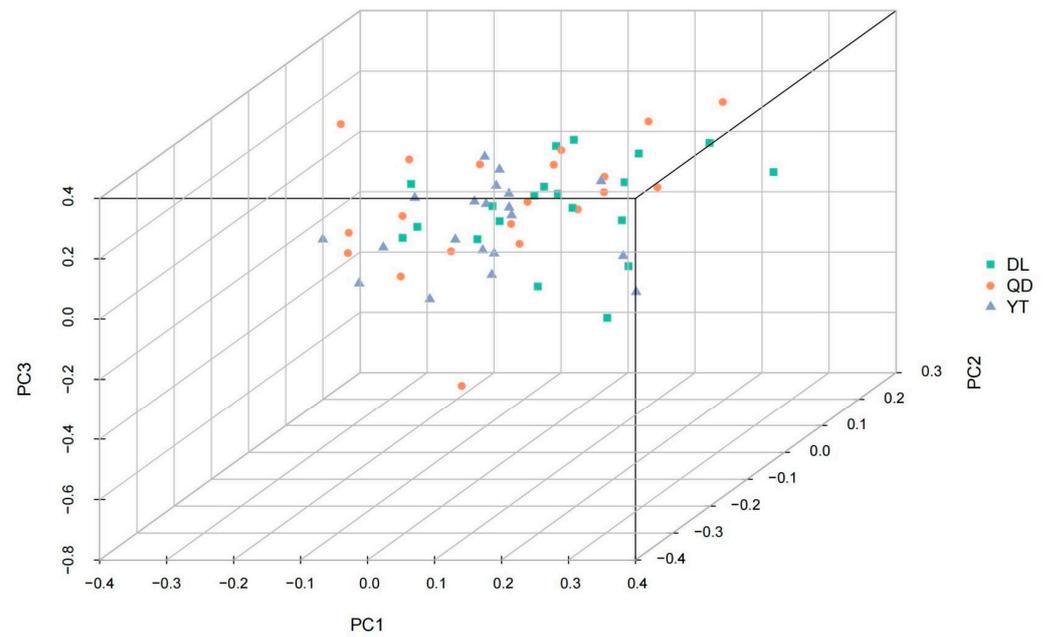
### 3.2. Population Structure of *S. schlegelii*

A phylogeny of these *S. schlegelii* individuals collected from two geographic areas was constructed (Figure 2). The 60 samples from different populations were clustered together, indicating shallow structures and close relationships with each other. The topology could not be generally defined by geographic localities; for instance, an amount of individuals from DL showed mixed components from QD and YT.

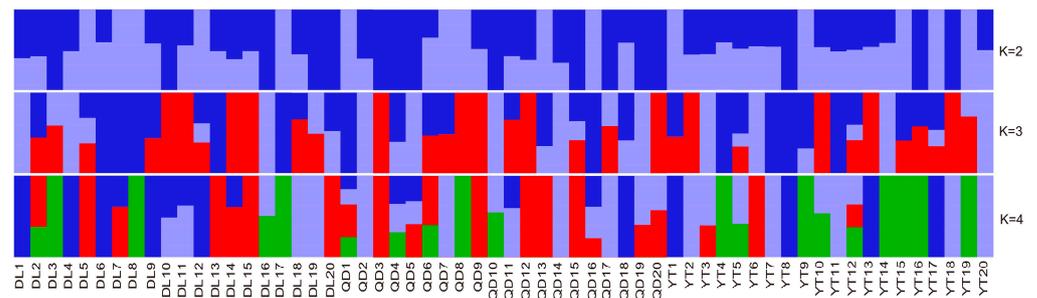
**Figure 2.** Phylogenetic relationship of 60 *S. schlegelii* individuals.

PCA was applied to cluster individuals into different geographic populations according to their characteristics based on the SNP differences at the genome level (Figure 3). None of the three principal components (PC1, PC2, and PC3) were able to distinguish among the DL, YT, and QD populations.

The population structure among two *S. schlegelii* geographical areas was estimated from population sizes  $K = 2-4$  by using ADMIXTURE analyses (Figure 4). The population structure was determined and supported the phylogenetic topology and PCA analyses. The structure among DL, YT, and QD populations showed no obvious genetic differentiation among different geographical populations, indicating gene flow among these populations.



**Figure 3.** PCA analyses of the three *S. schlegelii* populations. Each sample is represented by a triangle.



**Figure 4.** ADMIXTURE plot for *S. schlegelii* showing the distribution of  $K = 2, 3,$  and  $4$  genetic clusters among them. Each color represents a genetic cluster.

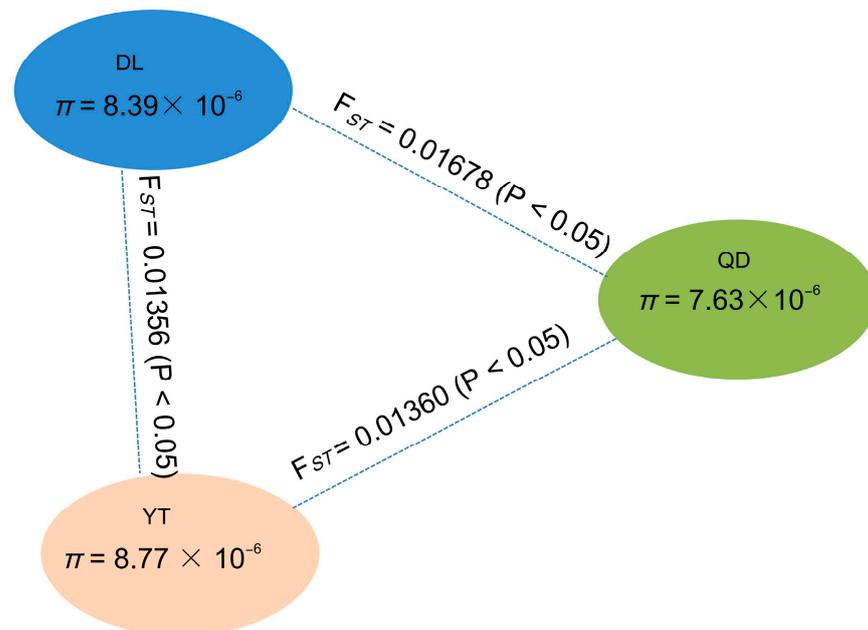
### 3.3. Genetic Diversity and Divergence of *S. schlegelii* Population

To investigate the genetic diversity and divergence among the two geographical areas (QD, DL, and YT), we observed a relatively high level of genetic diversity in the YT population ( $\pi = 8.77 \times 10^{-6}$ ) relative to DL ( $\pi = 8.39 \times 10^{-6}$ ) and QD ( $\pi = 7.63 \times 10^{-6}$ ). The observed heterozygosity ( $H_o$ ) measures the proportion of individuals in a population that are heterozygous at a given locus. In this study, the  $H_o$  for the DL, QD, and YT populations was 0.14316, 0.15809, and 0.17684, respectively (Table 2). The expected heterozygosity ( $H_e$ ) for the DL, QD, and YT populations was 0.14035, 0.15420, and 0.17145, respectively (Table 2). The polymorphism information content ( $PIC$ ) ranged from 0 to 0.5, with higher values indicating greater informativeness [34].  $PIC$  values for the DL, QD, and YT populations were 0.20672, 0.22616, and 0.24678, respectively (Table 2), indicating that the SNP markers used in this study are moderately informative. Additionally, the differences in the three parameters ( $H_o$ ,  $H_e$ , and  $PIC$ ) among the three populations were not obvious, indicating that there was little variation in the level of observed heterozygosity between the populations. These results suggest that the populations have similar levels of genetic diversity.

**Table 2.** Population genetic diversity parameters of *S. schlegelii*.

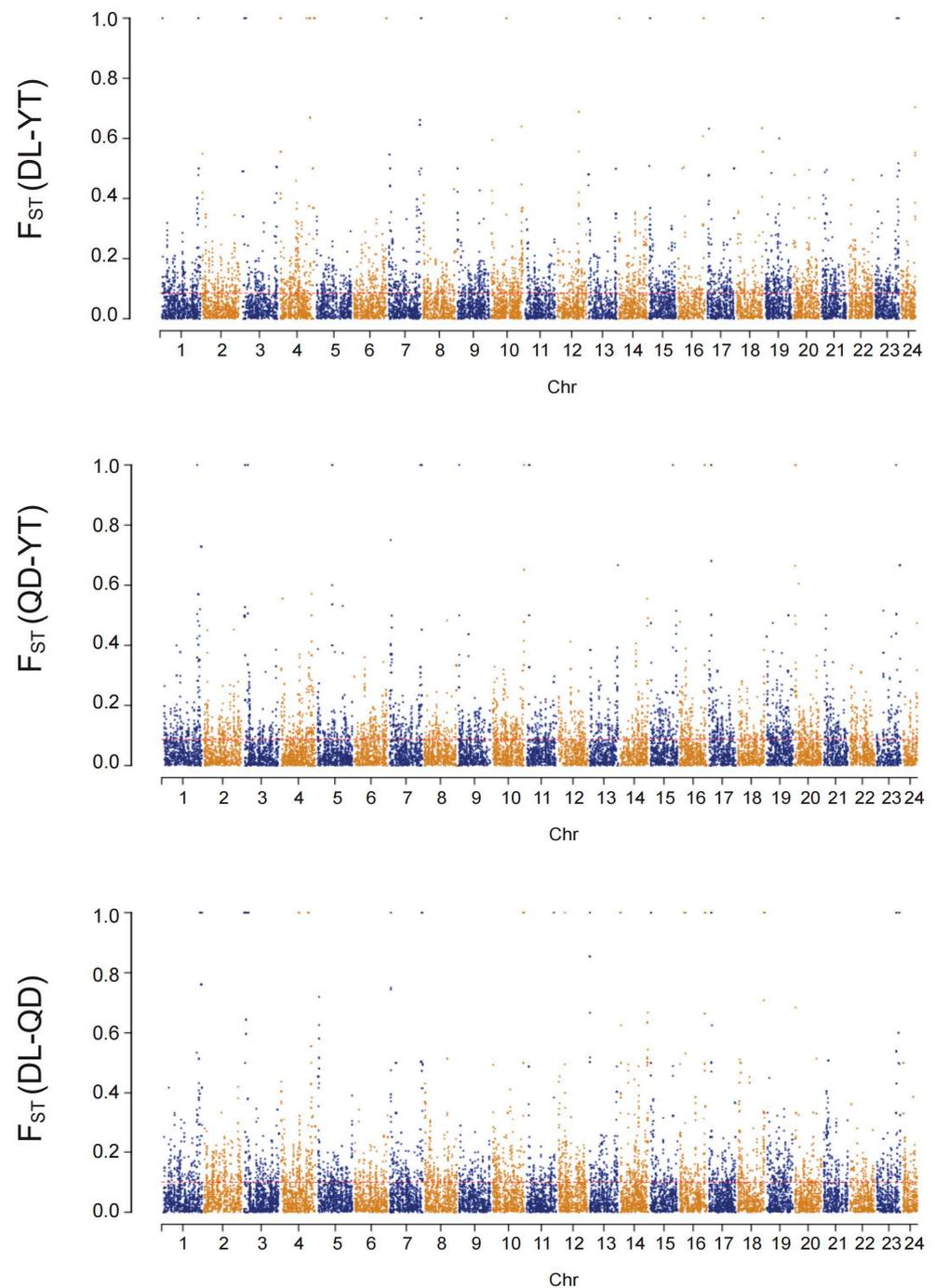
Population	Observed Heterozygosity ( $H_o$ )	Expected Heterozygosity ( $H_e$ )	Polymorphism Information Content (PIC)
DL	0.14316	0.14035	0.20672
QD	0.15809	0.15420	0.22616
YT	0.17684	0.17145	0.24678

A low but significant genetic differentiation among all populations was detected based on Global  $F_{ST}$  value using SNP loci ( $F_{ST} = 0.01667$ ,  $p < 0.05$ ). The pairwise  $F_{ST}$  values between the two populations ranged from 0.01356 to 0.01678 (Figure 5), and the genetic differentiation between the two populations reached a significant level ( $p < 0.05$ ). The highest pairwise  $F_{ST}$  value (0.01678) was observed between the DL and QD populations, while the lowest value (0.01356) was found for the DL and YT populations. The pairwise  $F_{ST}$  value between the YT and QD populations was 0.01360. The results indicate that the pairwise  $F_{ST}$  values in all three pairwise populations showed low levels of genetic differentiation ( $F_{ST} < 0.05$ ).



**Figure 5.** A summary of nucleotide diversity ( $\pi$ ) and population divergence ( $F_{ST}$ ) across the three populations. The values in parentheses represent the measures of nucleotide diversity of each group, and the values between pairs indicate population divergence.

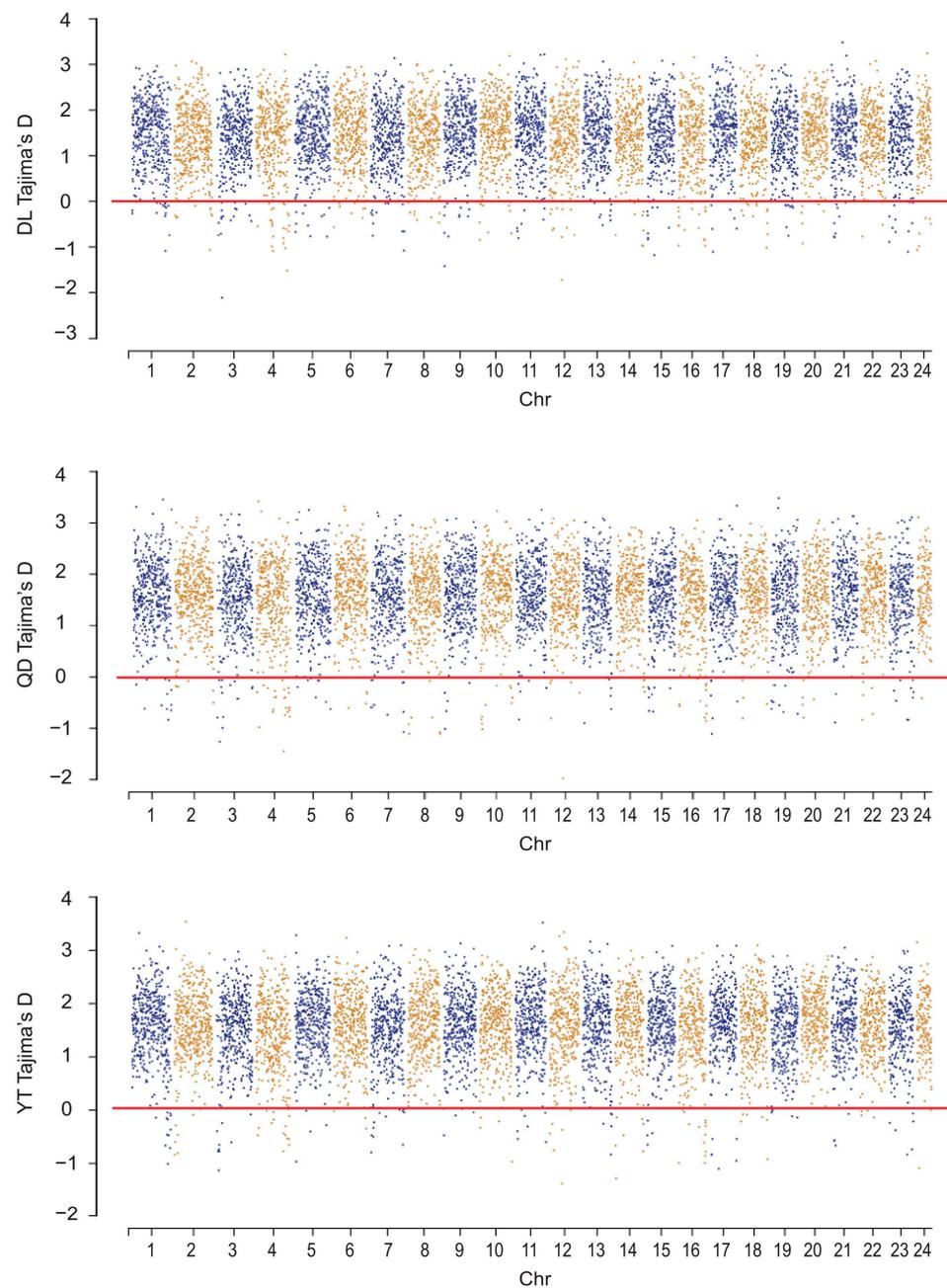
In addition, the number of loci at high levels of differentiation ( $F_{ST} \geq 0.15$ ) in  $F_{ST}$  values for pairwise populations (between DL and QD, between DL and YT, and between QD and YT) were 1770, 1325, and 1374, respectively. Comparatively, the number of loci at moderate differentiation levels ( $0 < F_{ST} < 0.15$ ) for pairwise populations were 11,743, 10,839, and 11,115, respectively (Figure 6).



**Figure 6.** Population divergence ( $F_{ST}$ ) between the QD, DL, and YT populations. Mean  $F_{ST}$  values are plotted in sliding windows of 100 kb with a 10 kb step size.

### 3.4. Genome-Wide Selection Pressure Analysis

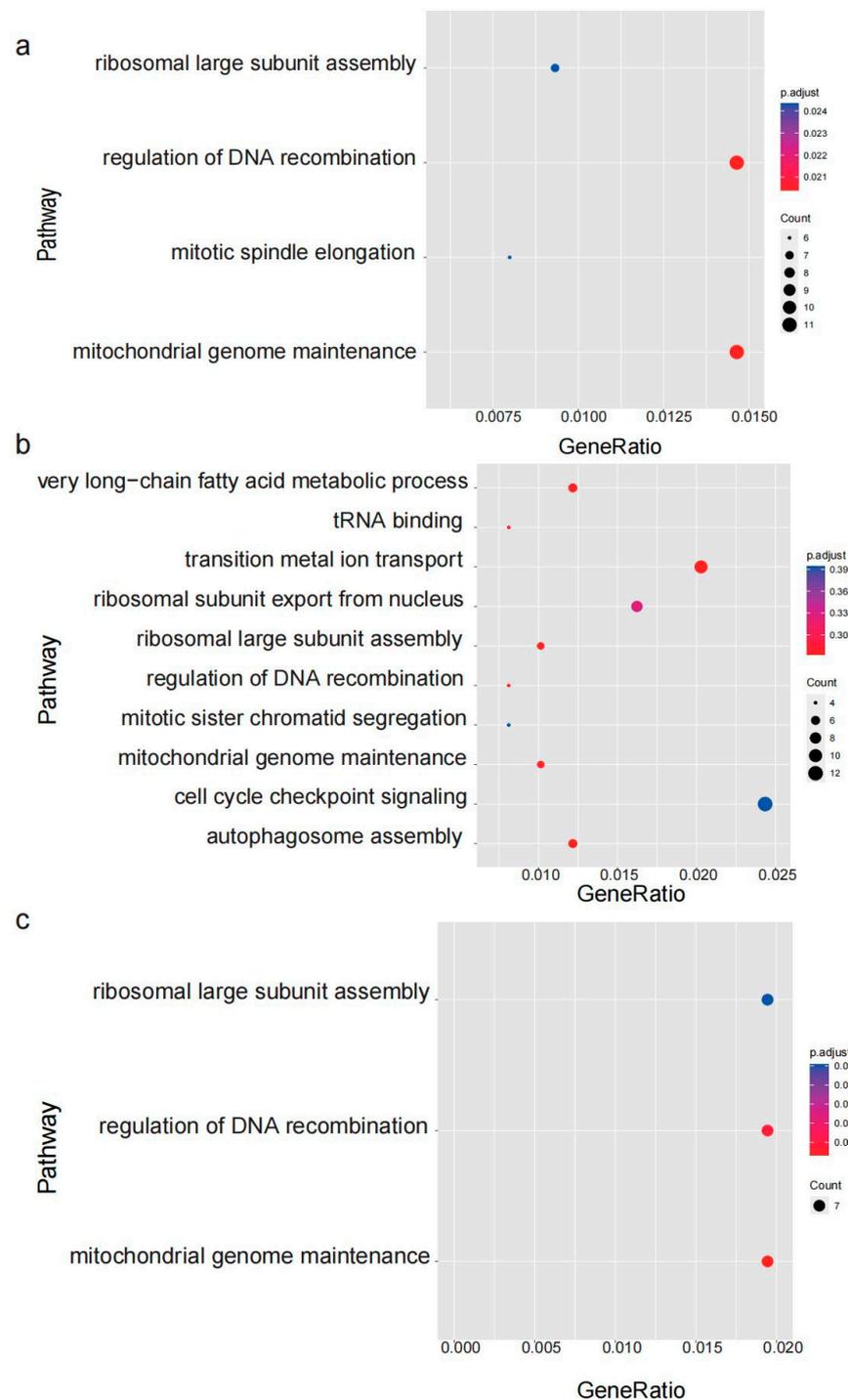
To compare the signatures of selective sweeps among the three *S. schlegelii* populations, a total of 225, 158, and 118 regions were analyzed under purifying selection (Tajima's  $D < 0$ ) based on a global database (149,434 SNPs) for DL, QD, and YT populations, respectively (Figure 7). These selectively swept regions were evenly distributed in 24 chromosomes of *S. schlegelii*, with a number of them showing high levels of selective sweeps (Figure 7).



**Figure 7.** Distribution of selective sweep regions in *S. schlegelii* genome.

These swept genomic regions overlapped with 774 protein-coding genes in DL, 506 genes in QD, and 370 genes in YT populations. Gene ontology functional enrichment analysis of the swept genes also revealed that the number of 15, 33, and 7 genes, respectively, in the DL, QD, and YT populations were involved in important functions (Figure 8). Interestingly, none of the swept genes were shared with each other among the three populations, yet these genes were enriched in the same pathways such as the regulation of DNA recombination, ribosomal large subunit assembly, and mitochondrial genome maintenance. In addition, the function of swept genes among the different populations also showed differences. For instance, some swept genes in the DL population were enriched in mitotic spindle elongation associated with sexual maturity, which may contribute to the breeding of superior varieties. In the QD population, some genes were involved in very long-chain fatty acid metabolic processes associated with energy metabolism, which may

provide a theoretical foundation for future research on the conservation and sustainable utilization of *S. schlegelii* resources.



**Figure 8.** GO enrichment analysis of genes under selective sweeps in three population of *S. schlegelii*. (a) DL population; (b) QD population; and (c) YT population.

## 4. Discussion

### 4.1. Genetic Differentiation Analysis

*S. schlegelii*, as an important economic fish species, is widely distributed in the northern sea area of China [22]. Previous studies have revealed that no significant morphological variations and no obvious genetic differentiation were detected among the above three *S.*

*schlegelii* populations (DL, QD, and YT) based on analyses of microsatellite DNA loci [8] and the mitochondrial DNA control region [16]. In this study, we investigated the genetic differentiation  $F_{ST}$  of *S. schlegelii* populations in the northwestern Pacific using GBS-based SNP data belonging to genome-wide markers which achieved deeper coverage than previous markers.  $F_{ST}$  values between 0.15 and 0.25 indicate high differentiation,  $F_{ST}$  values between 0.05 and 0.15 indicate moderate differentiation and  $F_{ST} < 0.05$  indicate no differentiation [35]. Analysis of genetic differentiation among above three populations based on Global  $F_{ST}$  revealed low genetic differentiation. In addition, all of the pairwise  $F_{ST}$  values between the populations indicated low levels of genetic divergence ( $F_{ST} < 0.05$ ), consistent with the previously reported lack of population genetic differentiation of *S. schlegelii* based on mtDNA control region sequences [16,17]. This result was also in accordance with the results of Peng et al., who used comparative mtDNA loop analysis to study the population structure and genetic diversity of the *S. schlegelii* populations in the Bohai Sea (Qingdao, Jinzhou, and Dalian), where there was no significant population genetic structural differences and low population genetic differentiation ( $F_{ST} = -0.0113-0.0061$ ) within the geographical range of the *S. schlegelii* population [36]. In addition, the pairwise  $F_{ST}$  values obtained were also similar to those obtained by Cao et al. (wild populations:  $F_{ST} < 0.05$ ) with a similar sampling design and 2b-RAD sequencing method [37]. In addition, the pairwise  $F_{ST}$  level of *S. schlegelii* populations ( $F_{ST} = 0.01356, 0.01360, \text{ and } 0.01678$ ) was less than many other marine organisms, such as *Clupea pallasii* ( $F_{ST} = 0.064-0.106$ ) [38], *Acanthopagrus latus* (the highest  $F_{ST}$  was 0.4288) [39], and *Brachionichthys hirsutus* (overall  $F_{ST} = 0.043$ ) [40]. This could confirm that *S. schlegelii* populations showed relatively low levels of genetic divergence among these two areas (QD, YT, and DL).

We hypothesized that the low levels of genetic differentiation among geographic populations might be due to two factors. First, it may be attributable to the swimming capability of *S. schlegelii* larvae. They live on the surface of the ocean before the adult stage and can disperse along drifting seaweeds and eat plankton with high dispersal potential, which leads to extensive gene flow between populations of *S. schlegelii* in the northwestern Pacific [18]. Second, during the main spawning seasons (April to June each year) of *S. schlegelii*, the Yellow Sea Warm Current is characterized by the influx of relatively warmer water from the East China Sea into the Yellow Sea; the China Coastal Current, flowing southward along the eastern coast of China, transports water and materials, interacting with the Yellow Sea Warm Current [41]. Sea currents can carry plankton and a large number of *S. schlegelii* larvae to further areas, which suggests that the movement of *S. schlegelii* occurs frequently between geographic populations [42]. However, this hypothesis needs to be tested with future studies.

#### 4.2. Genetic Diversity and Selection Pressure Analysis

The role of genetic diversity in the world is related to the survival, reproduction, fitness, and evolution of species [43]. Parameters such as  $H_o$ ,  $H_e$ ,  $PIC$ , and  $\pi$  based on SNPs are often used as criteria to evaluate population genetic diversity [44]. In our study, the  $H_o$  (0.14316–0.17684),  $H_e$  (0.14035–0.17145), and  $PIC$  (0.20672–0.24678) of *S. schlegelii* populations (DL, QD, and YT) showed an obviously lower level compared to the results ( $H_o = 0.2220-0.2271$ ,  $H_e = 0.2280-0.2297$ , and  $PIC = 0.989-0.911$ ) of *Pampus echinogaster* near the QD and DL coastal waters of China based on SNPs markers in a previous study [45]. This finding indicated that the number of loci possibly play a key role in genetic diversity. Analyses of  $\pi$  ( $7.63 \times 10^{-6}-8.77 \times 10^{-6}$ ) within all three populations of *S. schlegelii* revealed generally lower levels than many other marine organisms, such as *Apostichopus Japonicas* in Yantai ( $7.92 \times 10^{-3}$ ) [46], *Harpadon nehereus* (0.68833) [47], and *Sebastes marmoratus* (0.136) [48]. The above outcomes demonstrate that all three *S. schlegelii* populations with low levels of genetic diversity have poor capabilities to adapt to changing environmental conditions. Compared with our results, the slight difference in genetic diversity levels among the three populations might be largely caused by gene flow, which is consistent with the above genetic differentiation level.

However, the structure of *S. schlegelii* showed no obvious genetic differentiation among different geographical populations. The high level of genetic divergence (pairwise  $F_{ST}$  values  $> 0.15$ ) between populations indicates obvious genetic differentiation among different populations at outlier loci, which may be due to ecological adaptation or environmental pressures. A previous study demonstrated that negative effects of climate change on the distribution patterns of *S. schlegelii* would result in varying degrees of habitat reduction [49]. In our study, the selective swept genes among the populations may have differed, but their main functions were all oriented toward biological processes, such as DNA recombination, ribosomal large subunit assembly, and mitochondrial genome maintenance, which may contribute to their environmental adaptability [50–53]. Unique swept genes were discovered in each population. For instance, in the DL population, the C-reactive protein gene (*Ssc\_10008587*) enriched in the mitotic spindle elongation term may promote host defense via agglutination, phagocytosis, and complement fixation with calcium-dependent phosphorylcholine binding [54]; in the QD population, the gene (*Ssc\_10015638*) could be mediated by G proteins which activate adenylyl cyclase and are involved in very long-chain fatty acid metabolic processes associated with energy metabolism.

#### 4.3. Population Structure and Environmental Adaptation

The PCA plot displayed a mixed distribution of individuals, indicating that they could not be clearly separated by three eigenvectors, which further supported the phenomenon that there is no significant genetic difference among individuals from different geographic populations. These results were also consistent with the topology of the phylogenetic tree and the ADMIXTURE plot. Furthermore, no obvious geographical barrier existed among different sampling locations. Therefore, the free-floating eggs and larvae could be easily transported by strong coastal currents during the breeding season. The frequent gene flow and drift caused a low level of genetic differentiation among populations along the coast of China. Our results may provide a theoretical foundation for future environmental changes and for implementing adaptive management measures of *S. schlegelii* resources.

## 5. Conclusions

In conclusion, our study demonstrated that *S. schlegelii* populations showed low levels of genetic differentiation and genetic diversity among three populations of QD, DL, and YT in the Bohai and Yellow Seas. This phenomenon can guide the selection of populations with distant genetic relationships for propagation and release, which can promote genetic exchange within the population, increase the genetic diversity of the gene pool, and thereby enhance the adaptability and survival capacity of the population. Using selection pressure analysis, we found that the selective swept genes were mostly enriched in some biological processes involved in environmental adaptation. These findings help provide a theoretical basis for further studies aimed at protecting and utilizing *S. schlegelii* resources.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15050534/s1>, Table S1. The number of SNPs per filtering step.

**Author Contributions:** Y.Q., T.G. and X.Z. planned and designed the experiments. B.Z., Y.Q. and T.G. collected the samples and performed the experiments. Y.Q. analyzed and interpreted the sequencing data. B.Z., Y.Q. and Y.H. wrote the manuscript. X.Z. provided funding for sequencing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Natural Science Foundation of China (32072966) and the Province Key Research and Development Program of Zhejiang (2021C02047). The funding bodies had no role in the design of the study, in the collection, analysis, and interpretation of the data, and in the writing of manuscript.

**Institutional Review Board Statement:** This study received ethical approval from the Ethics Committee of Ocean University of China (Permit Number: 20141201) and adhered to all relevant ethical regulations for animal testing. This study was carried out in compliance with the ARRIVE guidelines (<https://arriveguidelines.org>) (accessed on 2 November 2023). All methods were performed in accordance with the relevant guidelines and regulations.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Sixty GBS genome-wide raw data produced by Illumina HiseqTM2500 have been deposited in the Genome Sequence Archive (SRA) database (<https://ngdc.cncb.ac.cn/gsa/browse/CRA008609>) (accessed on 10 February 2023) under the accession number CRA008609.

**Acknowledgments:** We thank Shengyong Xu and Fei Han for providing experimental assistance.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Song, W.; Xie, Y.; Sun, M.; Li, X.; Fitzpatrick, C.K.; Vaux, F.; O'Malley, K.G.; Zhang, Q.; Qi, J.; He, Y. A duplicated *amh* is the master sex-determining gene for *Sebastes rockfish* in the Northwest Pacific. *Open Biol.* **2021**, *11*, 210063. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, R.; Liu, H.; Zhang, Q.; Zhang, H.; Zhao, J. Trophic interactions of reef-associated predatory fishes (*Hexagrammos otakii* and *Sebastes schlegelii*) in natural and artificial reefs along the coast of North Yellow Sea. *China Sci. Total Environ.* **2021**, *791*, 148250. [[CrossRef](#)] [[PubMed](#)]
3. Zhang, M.; Cao, M.; Xiu, Y.; Fu, Q.; Yang, N.; Su, B.; Li, C. Identification of antimicrobial peptide genes in black rockfish *Sebastes schlegelii* and their responsive mechanisms to *Edwardsiella tarda* infection. *Biology* **2021**, *10*, 1015. [[CrossRef](#)] [[PubMed](#)]
4. Yin, L.; Chen, B.; Xia, B.; Shi, X.; Qu, K. Polystyrene microplastics alter the behavior, energy reserve and nutritional composition of marine jacobever (*Sebastes schlegelii*). *J. Hazard. Mater.* **2018**, *360*, 97–105. [[CrossRef](#)] [[PubMed](#)]
5. Ainley, D.G.; Sydeman, W.J.; Parrish, R.H.; Lenarz, W.H. Oceanic factors influencing distribution of young rockfish (*Sebastes*) in central California—A predators perspective. *Calif. Coop. Ocean. Fish. Investag. Rep.* **1993**, *34*, 133–139.
6. Zhang, Z.; Yu, Y.X.; Jiang, Y.; Wang, Y.G.; Liao, M.J.; Rong, X.J.; Wang, K.; Zhang, H.; Chen, J. First report of isolation and complete genome of *Vibrio rotiferianus* strain SSVR1601 from cage-cultured black rockfish (*Sebastes schlegelii*) associated with skin ulcer. *J. Fish. Dis.* **2019**, *42*, 623–630. [[CrossRef](#)] [[PubMed](#)]
7. Cao, M.; Yan, X.; Yang, N.; Fu, Q.; Xue, T.; Zhao, S.; Hu, J.; Li, Q.; Song, L.; Zhang, X.; et al. Genome-wide characterization of Toll-like receptors in black rockfish *Sebastes schlegelii*: Evolution and response mechanisms following *Edwardsiella tarda* infection. *Int. J. Bio Macromol.* **2020**, *164*, 949–962. [[CrossRef](#)] [[PubMed](#)]
8. Gao, T.; Han, Z.; Zhang, Z.; Luo, J.; Yanagimoto, T.; Zhang, H. Population genetic differentiation of the black rockfish *Sebastes schlegelii* revealed by microsatellites. *Biochem. Syst. Ecol.* **2016**, *68*, 170–177. [[CrossRef](#)]
9. Niu, J.; Wang, X.; Liu, P.; Liu, H.; Li, R.; Li, Z.; He, Y.; Qi, J. Effects of cryopreservation on sperm with cryodiluent in viviparous Black Rockfish (*Sebastes schlegelii*). *Int. J. Mol. Sci.* **2022**, *23*, 3392. [[CrossRef](#)]
10. Shen, F.; Zhang, Z.; Fu, Y.; Zhang, Z.; Sun, X.; Dong, J.; Ding, X.; Chen, M.; Zhang, X. Effects of food deprivation duration on the behavior and metabolism of black rockfish (*Sebastes schlegelii*). *Fishes* **2021**, *6*, 58. [[CrossRef](#)]
11. Rogers, S.I. Environmental factors affecting the distribution of sole (*Solea solea* L.) within a nursery area. *Neth. J. Sea Res.* **1992**, *29*, 153–161. [[CrossRef](#)]
12. Li, J.; Li, J.; Gong, P.; Guan, C. Effects of the artificial reef and flow field environment on the habitat selection behavior of *Sebastes schlegelii* juveniles. *Appl. Anim. Behav. Sci.* **2021**, *245*, 105492. [[CrossRef](#)]
13. Xu, T.; Zhang, X.; Ruan, Z.; Yu, H.; Chen, J.; Jiang, S.; Bian, C.; Wu, B.; Shi, Q.; You, X. Genome resequencing of the orange-spotted grouper (*Epinephelus coioides*) for a genome-wide association study on ammonia tolerance. *Aquaculture* **2019**, *512*, 734332. [[CrossRef](#)]
14. Xu, S.Y.; Zhao, L.L.; Xiao, S.J.; Gao, T.X. Whole genome resequencing data for three rockfish species of *Sebastes*. *Sci. Data* **2019**, *6*, 97. [[CrossRef](#)] [[PubMed](#)]
15. Yang, Y.; Wang, T.; Chen, J.; Wu, L.; Wu, X.; Zhang, W.; Luo, J.; Xia, J.; Meng, Z.; Liu, X. Whole-genome sequencing of brown-marbled grouper (*Epinephelus fuscoguttatus*) provides insights into adaptive evolution and growth differences. *Mol. Ecol. Resour.* **2022**, *22*, 711–723. [[CrossRef](#)] [[PubMed](#)]
16. Zhang, H.; Yanagimoto, T.; Zhang, X.; Song, N.; Gao, T. Lack of population genetic differentiation of a marine ovoviviparous fish *Sebastes schlegelii* in Northwestern Pacific. *Mitochondrial DNA Part A* **2016**, *27*, 1748–1754.
17. Zhang, H.; Zhang, Y.; Zhang, X.M.; Song, N.; Gao, T.X. Special structure of mitochondrial DNA control region and phylogenetic relationship among individuals of the black rockfish, *Sebastes schlegelii*. *Mitochondrial DNA* **2013**, *24*, 151–157. [[CrossRef](#)]
18. Hui, Z.; Xiumei, Z.; Zhiqiang, H.; Tianxiang, G. AFLP markers suggest low population genetic differentiation of the black rockfish *Sebastes schlegelii*. *Biochem. Syst. Ecol.* **2015**, *59*, 325–330.
19. Kramvis, A.; Bukofzer, S.; Kew, M.C. Comparison of hepatitis B virus DNA extractions from serum by the QIAamp blood kit, GeneReleaser, and the phenol-chloroform method. *J. Clin. Microbiol.* **1996**, *34*, 2731–2733. [[CrossRef](#)]

20. Han, Z.Q.; Guo, X.Y.; Liu, Q.; Liu, S.S.; Zhang, Z.X.; Xiao, S.J.; Gao, T.-X. Whole-genome resequencing of Japanese whiting (*Sillago japonica*) provide insights into local adaptations. *Zool. Res.* **2021**, *42*, 548–561. [[CrossRef](#)]
21. de Sena, B.G.; Smith, A.D. Falco: High-speed FastQC emulation for quality control of sequencing data. *F1000Res* **2019**, *8*, 1874. [[CrossRef](#)]
22. He, Y.; Chang, Y.; Bao, L.; Yu, M.; Li, R.; Niu, J.; Fan, G.; Song, W.; Seim, I.; Qin, Y.; et al. A chromosome-level genome of black rockfish, *Sebastes schlegelii*, provides insights into the evolution of live birth. *Mol. Eco Resour.* **2019**, *19*, 1309–1321. [[CrossRef](#)] [[PubMed](#)]
23. Giannoulatou, E.; Park, S.H.; Humphreys, D.T.; Ho, J.W. Verification and validation of bioinformatics software without a gold standard: A case study of BWA and Bowtie. *BMC Bioinform.* **2014**, *15*, S15. [[CrossRef](#)] [[PubMed](#)]
24. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
25. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
26. Van-der, A.; Geraldine, A.; Brian, D.O. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*; O'Reilly Media: Sebastopol, CA, USA, 2020.
27. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
28. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
29. Yang, J.; Lee, S.H.; Goddard, M.E. Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [[CrossRef](#)]
30. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [[CrossRef](#)]
31. Meyermans, R.; Gorssen, W.; Buys, N.; Janssens, S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genom.* **2020**, *21*, 94. [[CrossRef](#)]
32. Danecek, P.; Auton, A.; Abecasi, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)] [[PubMed](#)]
33. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *OmcS* **2012**, *16*, 284–287. [[CrossRef](#)] [[PubMed](#)]
34. Serrote, C.M.L.; Reiniger, L.R.S.; Silva, K.B.; dos Santos Rabaiolli, S.M.; Stefanel, C.M. Determining the polymorphism information content of a molecular marker. *Gene* **2020**, *726*, 144175. [[CrossRef](#)] [[PubMed](#)]
35. Tang, X.T.; Ji, Y.; Chang, Y.W.; Shen, Y.; Tian, Z.H.; Gong, W.R.; Du, Y.Z. Population genetic structure and migration patterns of *Lirionomyza sativae* in China: Moderate subdivision and no bridgehead effect revealed by microsatellites. *Bull. Entomol. Res.* **2016**, *106*, 114–123. [[CrossRef](#)] [[PubMed](#)]
36. Peng, L.; Hua, W.; Chen, Y.; Wang, W.; Xue, Z. Comparative analysis of the population diversity of black rockfish (*Sebastes schlegelii*) in northern China. *Mol. Biol. Rep.* **2023**, *50*, 10015–10024. [[CrossRef](#)] [[PubMed](#)]
37. Cao, W.; Zhang, M.; Wu, N.; Han, H.; Zhong, R.; Yu, T.; Zheng, Y. Population genetics analysis of the black rockfish *Sebastes schlegelii* in Northern China based on 2b-RAD simplified genome sequencing. *Isr. J. Aquacul.-Bamid* **2023**, *75*, 1–10. [[CrossRef](#)]
38. Orlova, S.Y.; Rastorguev, S.; Bagno, T.; Kurnosov, D.; Nedoluzhko, A. Genetic structure of marine and lake forms of Pacific herring *Clupea pallasii*. *PeerJ* **2021**, *9*, e12444. [[CrossRef](#)] [[PubMed](#)]
39. Liu, C.; Zhang, J.; Liu, S.; Song, P.; Guan, Y.; Shan, B.; Lin, L. Genetic diversity of the yellowfin seabream, *Acanthopagrus latus* (Actinopterygii: Perciformes: Sparidae)—An enhancement species in Dongshan Bay. *Acta Ichthyol. Piscat.* **2021**, *51*, 281–287. [[CrossRef](#)]
40. Appleyard, S.A.; Lynch, T.P.; Green, M.A.; Encinas-Viso, F. Genetic diversity and restricted genetic connectivity in an endangered marine fish (*Brachionichthys hirsutus*) provides a model for conservation management in related and data-deficient species. *Mar. Freshw. Res.* **2021**, *72*, 1735–1745. [[CrossRef](#)]
41. Xu, T.; Sun, J.; Watanabe, H.K.; Chen, C.; Nakamura, M.; Ji, R.; Feng, D.; Lv, J.; Wang, S.; Bao, Z.; et al. Population genetic structure of the deep-sea mussel *Bathymodiolus platifrons* (Bivalvia: Mytilidae) in the Northwest Pacific. *Evol. Appl.* **2018**, *11*, 1915–1930. [[CrossRef](#)]
42. Hasson, E.; Eanes, W.F. Contrasting histories of three gene regions associated with in (3L) Payne of *Drosophila melanogaster*. *Genetics* **1996**, *144*, 1565–1575. [[CrossRef](#)]
43. Dobson, A.; Crawley, M. Pathogens and the structure of plant communities. *Trends Ecol. Evol.* **1994**, *9*, 393–398. [[CrossRef](#)] [[PubMed](#)]
44. Johnson, D.W.; Freiwald, J.; Bernardi, G. Genetic diversity affects the strength of population regulation in a marine fish. *Ecology* **2016**, *97*, 627–639. [[CrossRef](#)]
45. Li, Y.; Lou, F.R.; Liu, S.G.; Li, H.; Xiang, J.L.; Shan, B.B.; Lin, L.; Zhuang, X. Differentiation and temperature adaptation of *Pampus echinogaster* based on genome-wide SNPs. *Front. Mar. Sci.* **2022**, *9*, 936217. [[CrossRef](#)]

46. Guo, C.; Li, Y.; Xie, J.H.; Han, L.S.; Wang, Y.Q.; Zhang, X.L.; Wu, Y.; Song, J.; Chang, Y.; Ding, J. Revealing selection in breeding and genetic characteristics of economically important traits of new species of *Apostichopus japonicas* based on genome resequencing and GWAS analysis. *Front. Mar. Sci.* **2022**, *9*, 948882. [[CrossRef](#)]
47. Yang, T.Y.; Huang, X.; Jiang, Y.L. Reveal the population genetic characteristics of bombay duck (*Harpadon nehereus*) in coastal waters of China with Genotype-by-Sequencing technique. *J. Ocean. Univ. China* **2022**, *21*, 1373–1380. [[CrossRef](#)]
48. Xu, S.Y.; Yanagimoto, T.; Song, N.; Cai, S.S.; Gao, T.X.; Zhang, X.M. Population genomics reveals possible genetic evidence for parallel evolution of *Sebastes marmoratus* in the northwestern Pacific Ocean. *Open Biol.* **2019**, *9*, 190028. [[CrossRef](#)]
49. Chen, Y.L.; Shan, X.J.; Ovando, D.; Yang, T.; Dai, F.Q.; Jin, X.S. Predicting current and future global distribution of black rockfish (*Sebastes schlegelii*) under changing climate. *Ecol. Indic.* **2021**, *128*, 107799. [[CrossRef](#)]
50. Sarvari, P.; Rasouli, S.J.; Allanki, S.; Stone, O.A.; Sokol, A.M.; Graumann, J.; Stainier, D.Y. The E3 ubiquitin-protein ligase Rbx1 regulates cardiac wall morphogenesis in zebrafish. *Dev. Biol.* **2021**, *480*, 1–12. [[CrossRef](#)]
51. Mosor, M.; Ziolkowska-Suchanek, I.; Nowicka, K.; Dzikiewicz-Krawczyk, A.; Januszkiewicz-Lewandowska, D.; Nowak, J. Germline variants in MRE11/RAD50/NBN complex genes in childhood leukemia. *BMC Cancer* **2013**, *13*, 457. [[CrossRef](#)]
52. Wang, J.; Ji, X.; Liu, J.; Zhang, X. Serine/Threonine Protein Kinase STK16. *Int. J. Mol. Sci.* **2019**, *20*, 1760. [[CrossRef](#)] [[PubMed](#)]
53. Yu, J.; Zhong, B.; Zhao, L.; Hou, Y.; Wang, X.; Chen, X. Receptor-interacting serine/threonine-protein kinase 1 (RIPK1) inhibitors Necrostatin-1 (Nec-1) and 7-Cl-O-Nec-1 (Nec-1s) are potent inhibitors of NAD(P)H: Quinone oxidoreductase 1 (NQO1). *Free Radic. Biol. Med.* **2021**, *173*, 64–69. [[CrossRef](#)] [[PubMed](#)]
54. Black, S.; Kushner, I.; Samols, D. C-reactive protein. *J. Biol. Chem.* **2004**, *279*, 48487–48490. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.