Genetic Diversity Analysis of Different Populations of *Lutjanus kasmira* Based on SNP Markers

Fangcao Zhao¹,²,³, Liang Guo¹,³, Nan Zhang¹,³, Jingwen Yang¹,³, Kecheng Zhu¹,³, Huayang Guo¹,³, Baosuo Liu¹,³, Bo Liu¹,³, Dianchang Zhang¹,³,⁴,⁵,*, and Shigui Jiang¹,³,⁴,⁵

¹ Key Laboratory of South China Sea Fishery Resources Exploitation and Utilization, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Ministry of Agriculture and Rural Affairs, Guangzhou 510300, China
² School of Fisheries and Life Science, Dalian Ocean University, Dalian 116023, China
³ Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China
⁴ Guangdong Provincial Engineer Technology Research Center of Marine Biological Seed Industry, Guangzhou 510300, China
⁵ Tropical Aquaculture Research and Development Center, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Sanya 572018, China

* Correspondence: zhangdch@scsfri.ac.cn

Abstract: *Lutjanus kasmira* belongs to the family Lutjanidae. Over the past 20 years, the *L. kasmira* population in the South China Sea has been shrinking due to climate change, pressure from human activities, and inadequate food supplies. In this study, single nucleotide polymorphism (SNP) data obtained from restriction site-associated DNA sequencing (RAD-seq) were used to assess the genetic diversity of *L. kasmira* in Zhubi Dao (ZB) and Meiji Dao (MJ). The genome-wide nucleotide diversity (*π*) of the ZB population and MJ population was 0.02478 and 0.02154, respectively. The inbreeding coefficient (*Fis*) of the ZB population and MJ population was −0.18729 and 0.03256, respectively. The genetic differentiation (*Fst*) between the ZB and MJ subpopulations was 0.00255102. The expected heterozygosity (*He*) of individuals from ZB and MJ was 0.33585 and 0.22098, respectively. The observed heterozygosity (*Ho*) of individuals from the ZB population and MJ population was 0.46834 and 0.23103, respectively. Although the ZB and MJ populations did not have significant genetic differences, the genetic differentiation between them was confirmed using population structure, phylogenetic, and principal component analyses. These results indicated that the genetic diversity of the ZB and MJ populations was relatively low at the genome level, and that their genetic differences were small.

Keywords: *Lutjanus kasmira*; SNP; genetic diversity; genetic differentiation

1. Introduction

Single nucleotide polymorphisms (SNPs) are DNA sequence polymorphisms caused by the transformation or transposition of a single nucleotide at the genomic level [1]. SNP markers are a tool for studying the genetic structure of species [2]. Since any base of genomic DNA can be mutated, SNP molecular markers are ubiquitous in animals and plants [3]. SNPs have the advantages of a large number, a wide distribution, strong representativeness [4], good genetic stability, and convenience for high-throughput and highly automated detection and analysis [5]. The distribution of SNPs in the genome can comprehensively reflect a population’s genetic variation [4].

*L. kasmira* is a typical reef-dwelling fish found throughout the Indo-Pacific region, from Australia in the east to Japan in the north [6]. Individuals of this species are mainly distributed around the South China Sea islands, Taiwan waters, and in the southern portion of the East China Sea [7]. *L. kasmira* has a bright yellow surface and a reddish underside; the side of the body bears four blue longitudinal bands, with an indistinct black spot between
the third and fourth blue bands [8]. In recent years, the habitat of this species, i.e., coral reefs and mangroves, has been reduced and degraded due to anthropogenic destruction and environmental pollution [9]. Moreover, *L. kasmira* suffers from overfishing due to its delicious meat, high economic value, and spawning clusters [7]. At present, studies on *L. kasmira* mainly focus on its physiological ecology and molecular systematics [7–9]; there are no reports on the development of SNP markers and the genetic diversity level of *L. kasmira*. In this study, the whole-genome SNPs of *L. kasmira* were first discovered by restriction site-associated DNA sequencing (RAD-seq) technology [10] and then used to analyze the genetic diversity and structure of this fish species, providing a theoretical basis for the rational development and conservation of its germplasm resources.

2. Materials and Methods

2.1. Sample Collection and DNA Extraction

*L. kasmira* samples were collected from Meiji Dao (MJ) and Zhubi Dao (ZB), which are separated by approximately 200 km (Figure 1). A total of 30 *L. kasmira* samples were collected, comprising 14 samples from MJ and 16 samples from ZB (Table 1). Tissue was collected by clipping the fin strip of the experimental samples. The tissue samples were fixed in 75% anhydrous ethanol and stored at room temperature. A Tissue DNA Extraction CZ Kit (Mobio, Guangzhou, China) was used to extract genomic DNA from the pterygiophore of each fish according to the manufacturer’s protocol.

![Figure 1](image)

**Figure 1.** Locations of the 30 collected *Lutjanus kasmira*.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Sample Location</th>
<th>Length (mm)</th>
<th>Sampling Date</th>
<th>Sample ID</th>
<th>Sample Location</th>
<th>Length (mm)</th>
<th>Sampling Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZB_SDDD_1</td>
<td>Zhubi Dao</td>
<td>136</td>
<td>2019</td>
<td>MJ_SDDD_1</td>
<td>Meiji Dao</td>
<td>153</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_2</td>
<td>Zhubi Dao</td>
<td>146</td>
<td>2019</td>
<td>MJ_SDDD_2</td>
<td>Meiji Dao</td>
<td>179</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_3</td>
<td>Zhubi Dao</td>
<td>150</td>
<td>2019</td>
<td>MJ_SDDD_3</td>
<td>Meiji Dao</td>
<td>153</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_4</td>
<td>Zhubi Dao</td>
<td>128</td>
<td>2019</td>
<td>MJ_SDDD_4</td>
<td>Meiji Dao</td>
<td>164</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_5</td>
<td>Zhubi Dao</td>
<td>133</td>
<td>2019</td>
<td>MJ_SDDD_5</td>
<td>Meiji Dao</td>
<td>123</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_6</td>
<td>Zhubi Dao</td>
<td>137</td>
<td>2019</td>
<td>MJ_SDDD_6</td>
<td>Meiji Dao</td>
<td>154</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_7</td>
<td>Zhubi Dao</td>
<td>146</td>
<td>2019</td>
<td>MJ_SDDD_7</td>
<td>Meiji Dao</td>
<td>126</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_8</td>
<td>Zhubi Dao</td>
<td>148</td>
<td>2019</td>
<td>MJ_SDDD_8</td>
<td>Meiji Dao</td>
<td>137</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_9</td>
<td>Zhubi Dao</td>
<td>152</td>
<td>2019</td>
<td>MJ_SDDD_9</td>
<td>Meiji Dao</td>
<td>146</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_10</td>
<td>Zhubi Dao</td>
<td>139</td>
<td>2019</td>
<td>MJ_SDDD_10</td>
<td>Meiji Dao</td>
<td>125</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_11</td>
<td>Zhubi Dao</td>
<td>120</td>
<td>2019</td>
<td>MJ_SDDD_11</td>
<td>Meiji Dao</td>
<td>143</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_12</td>
<td>Zhubi Dao</td>
<td>149</td>
<td>2019</td>
<td>MJ_SDDD_12</td>
<td>Meiji Dao</td>
<td>135</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_13</td>
<td>Zhubi Dao</td>
<td>132</td>
<td>2019</td>
<td>MJ_SDDD_13</td>
<td>Meiji Dao</td>
<td>154</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_14</td>
<td>Zhubi Dao</td>
<td>125</td>
<td>2019</td>
<td>MJ_SDDD_14</td>
<td>Meiji Dao</td>
<td>144</td>
<td>2019</td>
</tr>
<tr>
<td>ZB_SDDD_15</td>
<td>Zhubi Dao</td>
<td>165</td>
<td>2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZB_SDDD_16</td>
<td>Zhubi Dao</td>
<td>143</td>
<td>2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Sample information for the 30 *L. kasmira* collected in Zhubi Dao and Meiji Dao.
2.2. RAD-Seq Library Construction and Sequencing

RAD-seq libraries were prepared as described by Yangkun Wang and Yan Hu [11]. First, genomic DNA samples were digested by the restriction endonuclease EcoRI. Then, P1 adapter was added to both ends of the digested genome fragment, which contains a complementary sequence that binds to the primer for PCR amplification and a complementary sequence that binds to the primer for Illumina sequencing. The barcode used for sample tracking and the corresponding restriction enzyme cutting site were also part of the P1 adapter. Next, the sequence of the added P1 adapter was interrupted. Through agarose gel detection, the target band was selected in the range of 400–500 bp. The DNA fragment was then attached to the P2 adapter. The multiplexed libraries were sequenced on the Illumina 1.9 platform with an average sequencing depth of 1.5× for each individual.

2.3. Sequence Alignment and SNP Genotyping

Stacks (version 2.55, Julian Catchen and Nicolas Rochette, Urbana, IL, USA) software is widely used in RAD-seq data analysis [12]. First, raw reads were processed using the process_radtags program [11]. Through raw data processing, each clean read was assigned to a sample to ensure the quality of the sequencing data for use in subsequent analysis. Ustacks begins by clustering reads from a single sample obtained from the process_radtags program to produce stacks [13]. Ustacks in the Stacks package was applied to cluster the reads for each sample, with the same stack representing one enzyme cutting site (locus). The clustering parameter -m was set to 3 [14], and the loci and locus sequencing depth of each sample were statistically analyzed. Next, we used Cstacks to merge the loci of all samples and allowed up to 2 mismatches between different sample loci to obtain the catalogue consensus sequence of each locus [14]. Stacks was used next. The minor allele frequency (MAF) was set at 0.05 [15]. Loci in each sample deviating from the catalogue consensus sequence were identified, and then populations were filtered to obtain the SNPs. Only the genotype of the first SNP for each tag and loci shared by more than 75% of individuals were used for further analysis. The loci not in Hardy–Weinberg equilibrium (HWE) were removed [16], and the selected loci were used to analyze genetic diversity and population genetic differentiation.

2.4. Genetic Diversity and Population Structure

The genetic diversity parameters of MJ and ZB were calculated by Stacks-2.55 [15], including observed heterozygosity (Ho), expected heterozygosity (He), genetic differentiation (Fst), the inbreeding coefficient (Fis), and nucleotide diversity (π).

Genetic variation was determined by population structure analysis and genetic differentiation assessment. The R package poppr [17] was used to calculate a distance matrix and draw a heatmap [18]. The R packages igraph and poppr were used to cluster multi-locus genotypes (MLGs) to generate minimum spanning networks (MSNs) [19]. A phylogenetic tree was constructed by PHYLIP (version 3.695, Phylip Team, Washington, DC, USA) software [20] to analyze the relationships between samples. Compared with Structure, ADMIXTURE software can be used to estimate the ancestral components of individuals faster [21]. Therefore, ADMIXTURE (version 1.3.0, David H. Alexander et al., Los Angeles, CA, USA) software [22] was used to consider various numbers of putative clusters (ranging from 1 to 10) and compute the ten-fold cross-validation error to select the most suitable number of clusters [23]. The PCA module of GCTA (version 1.93.2, Jian Yang et al., Hangzhou, Zhejiang, China) software [24] was used for principal component analysis (PCA).

3. Results

3.1. Genomic Data Statistics and SNP Discovery

Raw reads were obtained by RAD-seq of 30 individuals of L. kasmira. After trimming barcodes and filtering, a total of 286,264,230 high-quality clean reads were obtained from 30 individuals. The number of clean reads per individual ranged from 6,672,063 to 21,514,652, with an average of 9,542,141. In all the libraries, the GC content was sta-
ble between 39% and 40%, and the Phred score was 20 ≥ 93.64% (Table 2). A total of 9822 high-quality SNPs were screened out for further analysis after the selection of SNPs with an MAF < 0.05 and in HWE.

Table 2. Statistics for the genomic data obtained from 30 *L. kasmira* individuals.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Number of Clean Reads</th>
<th>Bases (Q20%)</th>
<th>GC%</th>
<th>Sample ID</th>
<th>Number of Clean Reads</th>
<th>Bases (Q20%)</th>
<th>GC%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZB_SDDD_1</td>
<td>7,974,070</td>
<td>99.32</td>
<td>39</td>
<td>MJ_SDDD_1</td>
<td>9,265,891</td>
<td>95.32</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_2</td>
<td>6,867,444</td>
<td>96.12</td>
<td>40</td>
<td>MJ_SDDD_2</td>
<td>13,500,657</td>
<td>96.75</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_3</td>
<td>11,342,742</td>
<td>96.4</td>
<td>39</td>
<td>MJ_SDDD_3</td>
<td>9,120,598</td>
<td>99.30</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_4</td>
<td>6,831,672</td>
<td>95.40</td>
<td>40</td>
<td>MJ_SDDD_4</td>
<td>12,363,573</td>
<td>99.64</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_5</td>
<td>10,947,352</td>
<td>96.32</td>
<td>40</td>
<td>MJ_SDDD_5</td>
<td>8,022,139</td>
<td>98.20</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_6</td>
<td>8,673,002</td>
<td>98.89</td>
<td>39</td>
<td>MJ_SDDD_6</td>
<td>10,972,365</td>
<td>97.67</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_7</td>
<td>8,040,810</td>
<td>96.63</td>
<td>39</td>
<td>MJ_SDDD_7</td>
<td>14,935,273</td>
<td>94.03</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_8</td>
<td>7,251,357</td>
<td>97.60</td>
<td>39</td>
<td>MJ_SDDD_8</td>
<td>10,483,102</td>
<td>96.43</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_9</td>
<td>8,440,580</td>
<td>96.30</td>
<td>40</td>
<td>MJ_SDDD_9</td>
<td>9,245,798</td>
<td>94.60</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_10</td>
<td>7,441,155</td>
<td>94.56</td>
<td>39</td>
<td>MJ_SDDD_10</td>
<td>6,732,763</td>
<td>98.75</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_11</td>
<td>8,497,499</td>
<td>93.21</td>
<td>39</td>
<td>MJ_SDDD_11</td>
<td>9,673,994</td>
<td>95.30</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_12</td>
<td>6,672,063</td>
<td>95.63</td>
<td>39</td>
<td>MJ_SDDD_12</td>
<td>9,046,087</td>
<td>98.00</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_13</td>
<td>9,912,874</td>
<td>99.11</td>
<td>39</td>
<td>MJ_SDDD_13</td>
<td>6,814,092</td>
<td>95.19</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_14</td>
<td>7,340,574</td>
<td>98.65</td>
<td>39</td>
<td>MJ_SDDD_14</td>
<td>7,280,377</td>
<td>96.32</td>
<td>39</td>
</tr>
<tr>
<td>ZB_SDDD_15</td>
<td>21,514,652</td>
<td>99.30</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZB_SDDD_16</td>
<td>11,059,675</td>
<td>97.21</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2. Genetic Diversity and Population Structure

The genetic diversity parameters were calculated according to the original SNPs, and the statistical results were obtained. The genome-wide nucleotide diversity (π) of the ZB population and MJ population was 0.02154 and 0.02478, respectively. The inbreeding coefficient (Fis) of the ZB population and MJ population was −0.18729 and 0.03256, respectively. The genetic differentiation (Fst) between the ZB and MJ subpopulations was 0.00255102. The expected heterozygosity (He) of individuals from ZB and MJ was 0.33585 and 0.22098, respectively. The observed heterozygosity (Ho) of individuals from the ZB population and MJ population was 0.46834 and 0.23103, respectively (Table 3).

Table 3. Comparison of population genetic parameters in the ZB population and MJ population.

<table>
<thead>
<tr>
<th>Population ID</th>
<th>π</th>
<th>Ho</th>
<th>He</th>
<th>Average_Fis</th>
<th>Fst</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZB population</td>
<td>0.02478</td>
<td>0.46834</td>
<td>0.33585</td>
<td>−0.18729</td>
<td>0.00255102</td>
</tr>
<tr>
<td>MJ population</td>
<td>0.02154</td>
<td>0.23103</td>
<td>0.22098</td>
<td>0.03256</td>
<td></td>
</tr>
</tbody>
</table>

The heatmap showed positive correlations between the individuals of *L. kasmira*. In general, the correlations between individuals within the ZB population were the strongest, followed by those between the ZB population and MJ population, and the correlations between individuals within the MJ population were the weakest (Figure 2). The POP revealed the genetic distances between the 30 individuals, clustering MLGs based on genetic distance. The distances between individuals from the ZB population and MJ population were within the range of 0.002–0.004, indicating relatively short distances, consistent with other analysis results. Each MLG is a node, and the distance between the nodes represents the genetic distance between individuals (Figure 3). The phylogenetic tree revealed no significant clustering of the 30 *L. kasmira* samples (Figure 4). The population structure of *L. kasmira* was analyzed by Admixture software. Assuming that the number of groups (K value) is 1–10, the cross-validation error is maximal when K = 7 (Figure 5a,b). Therefore, the 30 samples of *L. kasmira* collected in this study were not from the same population. The genetic backgrounds of the ZB and MJ populations were complex, and there was admixture between the populations, which might be explained by the coexistence of artificial breeding and wild resources. The cross-validation results of the clustering showed that when K = 1
Figure 5a), the error rate of cross-validation was minimal, indicating that the optimal clustering number was 1. It is speculated that \textit{L. kasmira} individuals in the ZB and MJ populations came from the same primitive ancestor (Figure 5b). The results of PCA showed that the 30 \textit{L. kasmira} individuals were tightly clustered (Figure 6a,b) and the genetic heterogeneity between individuals was small, which was consistent with the phylogenetic tree results.

The heatmap showed positive correlations between the individuals of \textit{L. kasmira}. In general, the correlations between individuals within the ZB population were the strongest, followed by those between the ZB population and MJ population, and the correlations between individuals within the MJ population were the weakest (Figure 2). The POP revealed the genetic distances between the 30 individuals, clustering MLGs based on genetic distance. The distances between individuals from the ZB population and MJ population were within the range of 0.002~0.004, indicating relatively short distances, consistent with other analysis results. Each MLG is a node, and the distance between the nodes represents the genetic distance between individuals (Figure 3). The phylogenetic tree revealed no significant clustering of the 30 \textit{L. kasmira} samples (Figure 4). The population structure of \textit{L. kasmira} was analyzed by Admixture software. Assuming that the number of groups (\(K\) value) is 1–10, the cross-validation error is maximal when \(K = 7\) (Figure 5a,b). Therefore, the 30 samples of \textit{L. kasmira} collected in this study were not from the same population. The genetic backgrounds of the ZB and MJ populations were complex, and there was admixture between the populations, which might be explained by the coexistence of artificial breeding and wild resources. The cross-validation results of the clustering showed that when \(K = 1\) (Figure 5a), the error rate of cross-validation was minimal, indicating that the optimal clustering number was 1. It is speculated that \textit{L. kasmira} individuals in the ZB and MJ populations came from the same primitive ancestor (Figure 5b). The results of PCA showed that the 30 \textit{L. kasmira} individuals were tightly clustered (Figure 6a,b) and the genetic heterogeneity between individuals was small, which was consistent with the phylogenetic tree results.

Figure 2. The distance matrix heatmap of 30 individuals of \textit{L. kasmira} created using the R package poppr.

Figure 3. The minimum spanning network of 30 \textit{L. kasmira} individuals.
Figure 4. Phylogenetic tree of 30 L. kasmira individuals based on SNP loci created using the neighbor-joining method.

Figure 5. The group structure diagrams. (a) The cross-validation error for L. kasmira according to the admixture value K; (b) results of Bayesian cluster analysis of L. kasmira based on SNP loci using ADMIXTURE software for K = 1–10 clusters.

4. Discussion

RAD-seq is a simplified genome sequencing technology based on whole-genome restriction sites developed on the basis of second-generation sequencing [11]. The number of RAD markers developed by this method is 10 times higher than that of traditional molecular marker development technology [11], with high accuracy and high data utilization [25]. RAD-seq shortens the marker development cycle compared to that of traditional markers and reduces experimental costs [26]. The technology can also screen for SNPs on a large scale in species without a reference genome [27]. SNPs are a new type of DNA molecular marker with broad application prospects [28]. SNPs occur at a high frequency and have a high marker density in most genomes. Compared with simple sequence repeat (SSR) markers, SNPs have higher genetic stability and are easier to automate [29].
L. kasmira populations of this species. Based on a study of quality DNA from wild individuals [33]. By referring to the study of 29 Rhinopithecus populations. A and B represent the MJ population and ZB population, respectively.

MJ populations. A and B represent the MJ population and ZB population, respectively. A and B represent the MJ population and ZB population, respectively; (a) Principal component analysis (PCA) plot of 30 L. kasmira individuals based on all SNP loci between the ZB and MJ populations. A and B represent the MJ population and ZB population, respectively; (b) three-dimensional PCA clustering map of 30 L. kasmira individuals based on all SNP loci between the ZB and MJ populations. A and B represent the MJ population and ZB population, respectively.

4. Discussion

RAD-seq is a simplified genome sequencing technology based on whole-genome restriction sites developed on the basis of second-generation sequencing [11]. The number of RAD markers developed by this method is 10 times higher than that of traditional molecular marker development technology [11], with high accuracy and high data utilization [25]. RAD-seq shortens the marker development cycle compared to that of traditional markers and reduces experimental costs [26]. The technology can also screen for SNPs on a large scale in species without a reference genome [27]. SNPs are a new type of DNA molecular marker with broad application prospects [28]. SNPs occur at a high frequency and have a high marker density in most genomes. Compared with simple sequence repeat (SSR) markers, SNPs have higher genetic stability and are easier to automate [29]. RAD-seq technology can also be used to construct linkage maps [30]. For example, Palaiokostas et al. used RAD-seq technology to construct the first linkage map of Dicentrarchus labrax based on high-density SNPs [31].

In previous studies, genetic diversity was studied mainly by SSR analysis and D-loop sequence analysis [32], but these techniques have inherent difficulties in obtaining high-quality DNA from wild individuals [33]. By referring to the study of 29 Rhinopithecus roxellana by Zhang et al. [34], we used RAD-seq to detect SNP markers in 30 L. kasmira individuals in this study and further analyzed the genetic diversity levels of different populations of this species. Based on a study of L. kasmira in the South China Sea by microsatellite analysis [32], further exploration was performed. A comparison of the results of this study with those of other studies shows that the degree of variation of L. kasmira in the ZB and MJ areas is low. This reflects its ability to adapt to environmental change and respond to natural selection. Therefore, it is necessary to investigate the wild population as soon as possible and determine how to improve its genetic diversity. A larger sample size would produce more statistically robust results. More samples should be collected in different regions for more comprehensive genetic diversity analysis in the future.

In the phylogenetic tree based on RAD-seq data, the 30 samples of L. kasmira did not form obvious groups, and the population structure diagram also showed that they came from the same ancestor. Similar results were obtained with PCA, which were mutually confirmed with the results of the phylogenetic tree. The genome-wide nucleotide diversity (\(\pi\)) of the ZB population and MJ population was 0.02478 and 0.02154, respectively, indi-
cating high nucleotide diversity. The expected heterozygosity (He) of individuals from ZB and MJ was 0.33585 and 0.22098, respectively. However, the observed heterozygosity (Ho) of individuals from the ZB population and MJ population was 0.46834 and 0.23103, respectively. The He of the two populations was lower than the Ho, indicating a low degree of variation for *L. kasmira* in the two areas and a certain degree of heterozygote deficiency. The genetic differentiation (Fst) between the MJ and ZB subpopulations was 0.00255102. The Fst value was consistent with the results of the heatmap and MSN. The distance between individuals was short, and differentiation was not obvious. Although π was high, He and Ho were low. The overall analysis showed that the population structure of *L. kasmira* was simple and had a low degree of variation, which might lead to poor adaptability to the environment. Furthermore, the genetic diversity of the MJ population was lower than that of the ZB population.

These results may be due to overfishing and global warming, which have reduced the genetic diversity of wild populations. Therefore, there is an urgent need to protect the population genetic resources of *L. kasmira*, and we can improve its genetic diversity through fishing restriction protection measures and artificial breeding of superior varieties. It is suggested that effective population size analysis and population history tracing should be performed for *L. kasmira*.

5. Conclusions

In this study, SNP data obtained by RAD-seq technology were used to analyze the genetic diversity in two populations of *L. kasmira*, providing a new approach for genetic diversity evaluation. The results showed that the genetic diversity of the two populations was relatively low at the genome level. To ensure adequate survival of the species, it is necessary to protect existing diversity and take measures to improve genetic diversity. To promote the healthy and sustainable development of germplasm resources of *L. kasmira*, much attention should be given to improving and maintaining its population genetic diversity, to implementing methods to increase gene flow, and to breeding excellent varieties by applying molecular breeding techniques.

Finally, our results indicate that the RAD-seq technique can detect SNP markers and apply them to the research of genetic diversity, with the benefits of a huge number of markers, low costs, and simple automation. This study is useful for the conservation of aquatic germplasm resources as well as for the development and utilization of high-quality resources.

**Author Contributions:** Conceptualization, D.Z. and S.J.; methodology, F.Z. and L.G.; software, L.G.; validation, F.Z. and J.Y.; formal analysis, F.Z.; investigation, K.Z.; resources, N.Z.; data curation, F.Z.; writing—original draft preparation, F.Z.; writing—review and editing, D.Z.; visualization, F.Z.; supervision, H.G. and B.L. (Bo Liu); project administration, B.L. (Baosuo Liu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Financial Fund of the Ministry of Agriculture and Rural Affairs, P. R. of China (NHYYSWZZZYKZX2020) and National Marine Genetic Resource Center, and China-ASEAN Maritime Cooperation Fund.

**Institutional Review Board Statement:** All applicable international, national, and institutional guidelines for the care were followed by the authors.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patients to publish this paper.

**Data Availability Statement:** All data generated or analyzed during this study are included in this published article.

**Conflicts of Interest:** The authors declare no conflict of interest.
References


20. Felsenstein, J. *PHYLIP (Phylogeny Inference Package)*, version 3.6; Distributed by the author; Department of Genome Sciences, University of Washington: Seattle, DC, USA, 2005.


23. Wang, J.; Zhang, Z. Analysis of the Genetic Structure and Diversity of Upland Cotton Groups in Different Planting Areas Based on SNP Markers. *Gene* 2022, 809, 146042. [CrossRef]


33. Park, S.-H.; Scheffler, J.A.; Ray, J.D.; Scheffler, B.E. Identification of Simple Sequence Repeat (SSR) and Single Nucleotide Polymorphism (SNP) That Are Associated with the Nectariless Trait of *Gossypium hirsutum* L. *Euphytica* **2021**, *217*, 78. [CrossRef]