Muscular Coloration Diversity and Microsatellite Genetic Differentiation for the Clam Paphia amabilis (Philippi, 1847)

Suiteng Wang 1*, Xingxing Wu 1*, Feng Yu 1*, Tingxin Liao 1, Qiongzheng Li 2, Bang Luo 2* Yan Wang 2*

1 State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan Aquaculture Breeding Engineering Research Centre, College of Marine Science, Hainan University, Haikou, 570228, Hainan, China
2 Guangxi Academy of Fishery Science, Nanning, 530021, China

* These authors contributed equally to this work

Corresponding author email: ywang@hainanu.edu.cn; 76959887@qq.com

International Journal of Marine Science, 2021, Vol.11, No.1 doi: 10.5376/ijms.2021.11.0001

Abstract  Paphia amabilis is an important fishery shellfish widely distributing along Southeast Asia coast. The muscles of its foot and siphons are commonly colored bright red-orange due to the presence of C37 skeletal carotenoids. However, 10.5%-18.6% of yellow-muscled individuals (YM) were found among the four P. amabilis wild samples from Vietnam (Co To island, CT) and China (Beihai, BH/BHA, Guangxi Province; and Zhanjiang, ZJ, Guangdong Province), the north of Beibu Gulf, South China Sea. Twenty genomic microsatellite markers were developed and 12 were employed to analyze the population genetic differentiation of three geographical populations (CT, BH, and ZJ, n=32/poolulation), as well as YM and RM (red-muscled) subpopulations. Medium levels of genetic diversity were detected in all three populations with the number of alleles ranging from 4-21 (8.4±4.5) for CT, 4-16 (7.6±3.3) for BH and 4-18 (7.5±3.7) for ZJ. The average observed/expected heterozygosities were 0.66/0.66 for CT, 0.62/0.63 for BH, and 0.55/0.62 for ZJ, respectively. Total 4 out of 36 tests of microsatellites deviated from Hardy–Weinberg equilibrium (HWE, P<0.05 after Bonferroni correction), all in ZJ population due to the presence of null alleles suggested by MICRO-CHECK. Exact test indicated highly significant differentiations (Fst) between Vietnamese and two Chinese populations (0.099/0.123 between CT and BH/ZJ). No allelic of any microsatellite marker was found significantly associated with muscular coloration of P. amabilis, consequently no convergent microsatellite genetic differentiation presents between YM and RM subpopulations, supporting that epigenetic factor might play an important role in carotenoids accumulation of P. amabilis.

Keywords  Paphia amabilis; Microsatellite; Population genetic structure; Muscular color variance; Genetic diversity

The clam Paphia amabilis (Philippi, 1847) is a popular marine fishery bivalve (Mollusca: Veneridae). It is widely distributed along Southeast Asia coast, from Korea and Japan to Philippines, and along Chinese and Vietnamese coasts to Siam Gulf, Malaysia and Andaman Islands (Maoka et al., 2008; Li et al., 2014). In China, natural populations of P. amabilis were mainly found rich in Fujian, Guangdong, Guangxi and Hainan Provinces in Southern China in the last two decades (Xu, 1987; Zhang and Qi, 1960). However, in recent years, due to heavy fishing pressure caused by increasing need in consumption (Wang et al., 2013), they are only found in limited regions of Guangdong and Guangxi Provinces and it is necessary to develop the aquaculture of this clam (Li et al., 2014).

The muscles of foot, inhalant and exhalant siphons of P. amabilis are mostly colored bright red orange (Figure 1-1) due to high C37 skeletal carotenoid content, which is derived from the algae it feed on (Maoka et al., 2008). However, we found that some wild P. amabilis have yellow or yellow orange muscles (YM, Figure 1-2; Figure 1-3), making them less desirable than the red-muscled individuals (RMs) for the consumers. Besides, during the long time of anhydrous transportation at around 20°C, we found that the red feet of RMs obviously protruded out of their shell, while YMs closed their shell tightly. At 10-12 h post- transportation, quite a few YMs died but none RMs did.
The color variance in marine bivalve is probably due to the differences in total carotenoid content (Zheng et al., 2010) caused by genetic factor or environmental factors (Zheng et al., 2012a; Borodina and Soldatov, 2016). Recent studies showed that there are significant differences in the qualitative composition and seasonal variation of carotenoids in shellfishes with similar habitats and filter feeding patterns (Borodina and Zadorozhny, 2020), indicate that genetic and metabolic characteristics play important roles in the concentration of carotenoids. In order to establish effective molecular markers for selective breeding of the red muscle strains, the distribution of muscle color in natural populations and in male and female, as well as the genetic differentiation of neighboring geographical populations and subpopulations of red/yellow muscle, should be explored. In our earlier efforts to develop genetic markers for conservation and aquaculture of this species, 12 microsatellite markers have been identified based on one wild population (Li et al., 2014). In this study, (1) we developed more genomic microsatellite markers and analyzed the genetic structure of three geographic populations, and the genetic differentiation of YMs and RMs subpopulations; (2) the distribution of different muscular coloration in the natural populations, as well as female and male subpopulations of *P. amabilis* in the north of Beibu Bay, were investigated. Our aim is to assist aquatic products breeding in the selection of parental founder population, and to explore the mode of inheritance of muscular coloration in *P. amabilis*.

1 Results
1.1 Muscle coloration

The muscles of *P. amabilis*'s feet and siphons mostly have bright red color (Figure 1-1; Table 1). The variety from 10.5% to 18.6% of YMs was, however, found in CT, ZJ and BH/BHA populations (Figure 1-3; Table 1), and, 5 clams (8.3%) with intermediate muscle colour (between red and yellow, Figure 1-2; Table 1) were observed only in the spring BH population.

Chi-Square tests revealed that no significant difference of the sex ratio (about 1:1) and the percentage of YMs among those four populations (*p*>0.05, Table 1). And, although the ratio of YM females was lower than that of YM males, but not significant in spring populations (BH, CT and ZJ, *p*>0.05), whereas significant in autumn BHA specimens (*p*<0.05, Table 1).

Table 1 Percentage and the sex ration of the RM and YM clams in each population

| Pop                  | CT    | ZJ    | BH*   | BHA   |
|----------------------|-------|-------|-------|-------|
| Total No.            | 57    | 43    | 60    | 60    |
| ♀:♂                  | 24:33 | 23:20 | 26:34 |       |
| YMs No. (percentage) | 6 (10.5%) | 8 (18.6%) | 11 (18.3%) | 10 (16.7%) |
| ♀:♂ in RM clams      | 23:28 | 21:14 |       |       |
| ♀:♂ in YM clams      | 1:5   | 2:6   |       | 1:9   |

Note: *The muscle color of 5 individuals in BH were between red and yellow (Figure 1); *Data not tested. Values in bold represented significant differences at *p*<0.05
1.2 Microsatellite markers of *P. amabilis*

In total, twenty primer pairs amplified consistently microsatellites in the 3 individual DNA pools were selected for further population genotyping (Table 2). The GenBank accession numbers, primer sequences, motifs, the optimized annealing temperatures, and total allele number based on all samples from three geographical populations were listed in Table 3. For the three populations, 19 out of 20 loci (excluding *Pam5*) were polymorphic with the total number of alleles ranging from 3 for *Pam03* to 24 for *Pam30* (Table 2). Since *Pam23* and *Pam38* were monomorphic in at least one population, and *Pam33*, *Pam37* and *Pam38* were only amplified in less than half clams, these 4 loci were excluded from further population genetics analysis (Table 2).

| Pop | Loci | *F*<sub>ST</sub> | *Pam06* | *Pam07* | *Pam08* | *Pam11* | *Pam12* | *Pam16* | *Pam17* | *Pam19* | *Pam24* | *Pam25* | *Pam30* | *Pam35* | Average |
|-----|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| BH  | n    | 30     | 31     | 28     | 30     | 30     | 31     | 31     | 29     | 31     | 32     | 28     | 30.2±1.3 |
|     | a    | 9      | 4      | 7      | 8      | 5      | 7      | 4      | 9      | 7      | 10     | 16     | 5      | 7.6±3.3 |
|     | as   | 228-266| 142-174| 210-244| 269-293| 302-322| 097-119| 134-146| 266-312| 226-242| 191-241| 236-276| 168-182 |
| CT  | n    | 32     | 32     | 31     | 30     | 31     | 32     | 30     | 32     | 31     | 32     | 32     | 31.3±0.8 |
|     | a    | 6      | 6      | 6      | 8      | 4      | 8      | 7      | 5      | 9      | 11     | 11     | 21     | 6      | 8.4±4.5 |
|     | as   | 246-262| 136-170| 220-246| 269-293| 310-330| 099-129| 130-146| 288-332| 248-253| 230-284| 216-182 |
| ZJ  | n    | 30     | 31     | 28     | 32     | 30     | 31     | 32     | 31     | 29     | 32     | 32     | 30.6±1.4 |
|     | a    | 8      | 6      | 8      | 6      | 7      | 4      | 9      | 5      | 9      | 18     | 5      | 7.5±3.7 |
|     | as   | 206-266| 154-178| 210-244| 269-293| 300-322| 097-129| 134-146| 274-320| 230-248| 195-241| 226-270| 168-182 |

Note: * Loci excluded for population genetic statistical analysis. n: number of individuals, a: number of alleles, Rs: allele richness per locus, as: allele size range in base pairs, Ho: observed heterozygosity; He: expected heterozygosity; Fis: inbreeding coefficient, HW: probability of concordance with HWE. Values in bold represent significant deviation from HWE after Bonferroni correction (initial α = 0.05/12 = 0.0042); † Loci showed evidence for null alleles in the population, suggested by MICRO-CHEKER, may result in deviation from HWE.

1.3 Genetic structure and diversity in three geographical populations and musculature coloration groups

The statistics of microsatellite genetic variation in the three populations was listed in Table 2. Six loci (*Pam7, 8, 14, 16, 25, 36*) deviated from HWE in ZJ (*p*<0.05 after Bonferroni correction) in the direction of heterozygote deficiency, likely due to the presence of null alleles suggested by MICRO-CHEKER. Three loci (*Pam14, 20, 36*) deviated from HWE in BH, two (*Pam20, 36*) of them and other two loci (*Pam19, 25*) show signs of null alleles. In CT population, two loci (*Pam20 and Pam36*) deviated from HWE, and one locus (*Pam19*) show evidence for
null alleles (Table 2). Loci Pam14, 20, 36 were excluded from further population genetics analysis because of their deviation from HWE ($p<0.05$, after Bonferroni correction) in at least 2 populations.

Table 3 Information of 20 P. amabilis microsatellite markers

| Locus/GenBank no. | Primer Sequence (5’-3’) | Motif | Ta (°C) | Total No. of allele |
|-------------------|--------------------------|-------|---------|-------------------|
| Pam05             | F:aggagctcagattgtgttctc | (AC)$_5$ | 60 | 1 |
| KF740483          | R:taggagtaaagttgagggagt |       |       |                   |
| Pam06             | F:agtcagttgagcaactcttg | (TG)$_7$ | 60 | 11 |
| KF740484          | R:taacgctgtagtggcctac |       |       |                   |
| Pam07             | F:caaatcaacaggtgtcttcc | (AC)$_9$(AG)$_9$ | 48 | 9 |
| KF740485          | R:gttttcattttgcccttc |       |       |                   |
| Pam08             | F:ggtgacagcaggtctaaacc | (TG)$_6$ | 60 | 12 |
| KF740486          | R:gcgactcttttctttgtgctg |       |       |                   |
| Pam11             | F:aaaccaacacaaagaacacaa | (ACCG)$_9$(AGC)$_9$... | 60 | 11 |
| KF740487          | R:ccaaagaaaaagttggacacag | (CACA)$_9$...(AAAAAC)$_4$ | 60 | 5 |
| Pam12             | F:aagttggctgctacaatgcag | (TA)$_7$ | 60 | 7 |
| KF740488          | R:ccattattgcccgctttc |       |       |                   |
| Pam14             | F:tatttttgtgtaccccttgc | (TG)$_8$ | 50 | 14 |
| KF740489          | R:gccaggattttaaaagcatga |       |       |                   |
| Pam16             | E:ccattttgctaatgcaagcag | (TG)$_5$ | 60 | 8 |
| KF740490          | R:acagacgcacacacacacaa |       |       |                   |
| Pam17             | F:ttgtaattgcacccgacag | (GACA)$_4$ | 60 | 3 |
| KF740491          | R:tttagcctacaaagagaacacag |       |       |                   |
| Pam19             | F:ctgatagtgtcctcaatacaca | (AG)$_9$ | 60 | 11 |
| KF740492          | R:ccgcgttcatttacattatcg |       |       |                   |
| Pam20             | F:tgaaattgtccctcctacacta | (CT)$_9$ | 60 | 10 |
| KF740493          | R:acgcctttgtaagatgtagaa |       |       |                   |
| Pam23             | F:aacctgactagtggctcaca | (GA)$_4$(GA)$_4$ | 60 | 3 |
| KF740495          | R:tcgcacaacaaaattctgcca |       |       |                   |
| Pam24             | F:ccgagagaaagtctcagatga | (GA)$_9$(GA)$_6$ | 60 | 11 |
| KF740496          | R:ctgtagctgtgctacatct |       |       |                   |
| Pam25             | F:cttcagcagctgttatactct | (TC)$_9$ | 60 | 17 |
| KF740497          | R:atccagacgctctacattcg |       |       |                   |
| Pam26             | F:cttgcaacaatttctgctca | (AG)$_4$(GA)$_6$ | 60 | 24 |
| KF740498          | R:gcctgcgctaatttctgctagc |       |       |                   |
| Pam27             | F:ggctgacaagaatagtctgagc | (GA)$_3$ | 60 | 17 |
| KF740500          | R:acgcgtgttacagacagaag |       |       |                   |
| Pam28             | F:gcactattatgctgctgatcag | (AG)$_3$ | 60 | 6 |
| KF740501          | R:ctgtaaggaggtcacaacaa |       |       |                   |
| Pam29             | F:gcaggtttggaggtgagtgtgc | (AG)$_2$ | 60 | 20 |
| KF740502          | R:acagcagatcaacaaagactca |       |       |                   |
| Pam30             | F:ccccaggcttataatggaaac | (CA)$_6$...(CA)$_3$ | 60 | 7 |
| KF740503          | R:gtcattttgtcacaatttcc |       |       |                   |
| Pam31             | F:gagggttaagtgaagctgatg | (CT)$_3$ | 60 | 12 |
| KF740504          | R:ggcactattatggtgcttcca |       |       |                   |

Medium levels of polymorphism were observed in all three populations with the number of alleles /allele richness ranging from 4-16 (7.6±3.3) /3.9-15.2 (7.1±3.1) for BH, 4-21 (8.4±4.5) /3.8-19.2 (7.8±4.1) for CT, and 4-18 (7.5±3.7) /4.0-15.9 (6.9±3.2) for ZJ, respectively (Table 2). The average observed /expected heterozygosity were 0.62±0.17 /0.63±0.19 for BH, 0.66±0.14 /0.66±0.16 for CT, and 0.55±0.22 /0.62±0.19 for ZJ, respectively (Table 3). No significant differences of diversity level (allele number /richness and observed /expected heterozygosities) were found between any population pair (Kruskal–Wallis test, $p>0.05$). The inbreeding coefficient $F_{IS}$ varied
among markers from 0.16 (Pam12) to 0.37 (Pam35) in BH, from -0.21 (Pam06, 24) to 0.21 (Pam19) in CT, and from -0.22 (Pam35) to 0.52 (Pam08) in ZJ (Table 3). Average $F_{IS}$ based on all 12 loci were not high, with the highest (0.12) for ZJ, and the lowest (-0.02) for CT (Table 3).

The global multilocus $F_{ST}$ based on 12 loci was estimated to be 0.026 ($p < 0.001$, Table 4), indicating genetic differentiation among three populations. The pairwise $F_{ST}$ value (-0.002) and Nei’s genetic distance (0.067) were low between BH and ZJ populations (Table 4), but high between CT and the two Chinese populations ($F_{ST}$, 0.036/0.032 for BH /ZJ, and Nei’s genetic distance, 0.099/0.123 for BH/ZJ). The exact tests showed the same genetic pattern, suggesting significantly genetic differences between CT and BH/ZJ populations (Table 4). The consequent Neighbour-joining dendrogram (Figure 2), constructed based on Nei’s $D_A$ genetic distance (Nei et al., 1983), recognized two clusters. The two Chinese populations (BH and ZJ) clustered together to form cluster I, and the Vietnamese CT population separately formed cluster II.

Table 4 Pairwise Nei’s genetic distance (below the diagonal) and $F_{ST}$ (above the diagonal) between $P$. amabilis populations. HS indicates highly significant genetic differences between the populations ($p<0.05$)

| Pop  | CT   | BH    | ZJ     |
|------|------|-------|--------|
| CT   | -    | 0.036* | 0.032* |
| BH   | 0.099| -     | -0.002 |
| ZJ   | 0.123| 0.067 | -      |

Figure 2 Neighbour-joining tree of 3 $P$. amabilis populations based on the Nei’s genetic distances

The Nei’s genetic distance matrix for muscular coloration YM groups (coded as ZJy, BHy and CTy, respectively) and RM groups (coded as ZJr, BHy and CTr, respectively, Table 5) of three geographic populations revealed that, CTy kept the largest genetic distance between all other groups (from 0.216 for CTr to 0.331 for BHy), while the smallest genetic distances were found between BHy and ZJr (0.076) /CTr (0.108). The Neighbour-joining phylogenetic tree for those muscular coloration groups (Figure 3) indicated that YM (ZJy, BHy) and RM (ZJr, BHy) groups from Chinese populations do not cluster any clade characterized by geographic population or muscular coloration, although the CTy and CTr groups of Vietnamese population remain outgroups due to geographic isolation. Consistently, paired $F_{ST}$ (Table 5) showed little differentiation between 4 muscular coloration groups of Chinese populations (from -0.012 for BHy vs. ZJy to 0.006 for ZJr vs. ZJy). The genetic differentiations of CTy vs. Chinese RM groups (CTy vs. ZJr 0.076, CTy vs. BHy 0.072) were the highest, and significantly higher than that between CTy vs. YM groups (CTy vs. BHy 0.052, CTy vs. ZJy 0.011). On the contrary, the genetic differentiations between CTr and Chinese RM groups (CTr vs. BHy 0.033, CTr vs. ZJr 0.034) were significantly higher than that of CTr and Chinese YM groups (CTr vs. BHy 0.018, CTr vs. ZJy 0.009). Within the same population CT, the genetic differentiation is lowest (CTr vs. CTy 0.004). Stable convergence of genetic differentiations was not detected in either RM or YM groups.

PCoA provided a visual representation of genetic similarity between six muscular coloration groups of three populations (Figure 4). The first and second principal coordinates explained 91.51% and 7.60% of the total variation, respectively. The two groups from BH (BHy and BHy) and the two groups from ZJ (ZJr and ZJr) were mainly spread over a large proportion of coord. 2, respectively, and interlacing with each other (Figure 5). For the two groups from CT population (CTr and CTy), they spread over a large proportion of the right side of coord. 1 in
other two-dimensional space, while CTy on the far side away from Chinese color-subpopulations. This result was consistent with the Neighbour-joining tree and the genetic differentiation data shown in Table 4.

No association was found between the frequencies of the genotypes and alleles of the 32 microsatellites (20 in this study, and other 12 developed previously, Li et al., 2014) and muscular color ($p$>0.05 with Bonferroni’s correction).

Table 5 Pairwise Nei’s genetic distance (below the diagonal) and $F_{ST}$ (above the diagonal) between RM (r) and YM (y) subpopulations

|       | CTr | CTy | BHr | BHy | ZJr | ZJy |
|-------|-----|-----|-----|-----|-----|-----|
| CTr  | -   | 0.004 | 0.033 | 0.018 | 0.034 | 0.009 |
| CTy  | 0.216 | -   | 0.072 | 0.052 | 0.076 | 0.011 |
| BHr  | 0.108 | 0.293 | -   | -0.010 | -0.003 | 0.003 |
| BHy  | 0.167 | 0.331 | 0.124 | -   | -0.009 | -0.012 |
| ZJr  | 0.135 | 0.326 | 0.076 | 0.122 | -   | 0.006 |
| ZJy  | 0.161 | 0.291 | 0.136 | 0.181 | 0.131 | -   |

Figure 3 Neighbour-joining tree of 6 P. amabilis colour groups based on the Nei’s genetic distance

Figure 4 PCoA plots of RM and YM subpopulations deriving from three geographical populations. CT (diamond), BH (triangle) and ZJ (round dot) are labelled in red and green, respectively

Figure 5 Map of sample sites
2 Discussion

General causes for microsatellites markers deviated from HWE include presence of substructures of the sampled population, inbreeding, or the presence of null alleles (Zouros and Foltz, 1984). In this study, null alleles seem to play the major role and 8 out of 11 tests can be explained by the presence of null alleles. Such deviations appear to be quite common in microsatellite markers developed for bivalves (e.g., Hedgecock et al., 2004; Wang et al., 2010). High frequency of null alleles may misestimate many types of population genetic analyses that rely on HWE (Waples, 2014), thus we exclude those loci deviated from HWE in more than one population for population genetic analyses. In addition, in ZJ population more loci were significantly deviated from HWE, probably because the population was located in the northeast corner of Beibu Gulf and the poor flow exchange restricted its panmixia, thus forming facility substrates. The level of genetic diversity of population largely depends on the mating system of a species, and the population evolutionary history, and on the level of environmental heterogeneity (Booy et al., 2000). It is critical for us to discover the genetic diversity level of *P. amabilis* in order to ensure the sustainable use of this aquatic resource. In this study, we found medium levels of genetic diversity in three *P. amabilis* natural populations in the north of Beibu Bay, South China Sea. The level of diversity was similar to that of other *in vitro* fertilization mollusks who also naturally live in this region, e.g., hard clam *Meretrix meretrix* (Ye et al., 2020), pearl oyster *Pinctada martensii* (Qiu et al., 2013), silver-lipped pear oyster *Pinctada maxima* (Wu et al., 2014) and moon scallop *Amusium pleuronectes* (Mahidol et al., 2007; Su et al., 2015). For three geographic populations in this study, no significant difference of diversity level was observed between any population pair, even though the highest observed values were found in Vietnamese population, which possibly because there is less fishing pressure in Co To Island. In addition, the habitat of *P. amabilis* in Co To Island is mud-sandy sediment, so the environmental conditions are more complex than that of the sandy sediment of Beihai and Zhanjiang.

The average inbreeding coefficient *F*_φ based on all 12 loci were not high (-0.02 for CT to 0.12 for ZJ), suggesting that overall, the loss of heterozygosity in the three groups was not obvious, and the genetic composition of the population was normal and healthy (Carlsson et al., 2006). This study investigated the genetic population structure of the *P. amabilis* along the northern Beibai Bay coastline, covering a distance of nearly 400 km. The global *F*_ST estimate (0.026) and pairwise *F*_ST between CT and BJ/ZJ, revealed significant genetic variation between those populations, indicating the gene flow was limited at a large geographical scale (142/199 km, between CT and BH/ZJ), while it was strong at a small geographical scale (75 km, between BH and ZJ). Superficially, the genetic structure of *P. amabilis* was consistent with a model of isolation by distance, where all populations are connected by continuous migration but the rate of gene flow is greatest between neighbouring populations (Arnaud et al., 2000). In fact, no significant positive correlation between genetic differentiation (*F*_ST/genetic distance) and geographical distance was observed (Mantel test: *p* = 0.49/0.15). That could be attributed to the gene flow between population BH and ZJ by not only the planktrophic larvae migration but also human commercial transportation.

Patterns of larval dispersal and connectivity among marine invertebrate population are often estimated from life-history characteristics and the oceanographic conditions (Colson and Hughes, 2004; Cowen et al., 2006). The limited larval dispersal ability of the *P. amabilis* results from several life-history characteristics of this species, such as short planktonic larval stage (only 14–18 days, Li et al., 2012) and short spawning period (from April to May and from September to October, in total less than two months, Li et al., 2013). Furthermore, in the middle of the productive season of *P. amabilis*, Beibu Gulf currents alternate its direction (Yang et al., 2003), thus likely affecting the migration of *P. amabilis* larvae. In addition, the direction of coastal current near western Guangdong turns to the northeast (Yang et al., 2003) when the southwest summer monsoon becomes strong, which might promote the migration of *P. amabilis* larvae between ZJ and BH. Besides, larval settlement and metamorphosis of *P. amabilis* were influenced by many environmental factors, such as temperature and salinity (Li et al., 2012). All of which must affect the population dynamics of this species. Cowen et al. (2006) reported that typical larval dispersal distances for a variety of reef fish species are on the scale of only 10 to 100 kilometres using a high-resolution biophysical model for the Caribbean region, this distance scale is consistent with that of *P. amabilis* larval dispersal detected in this study.
Carotenoids are powerful antioxidants that affect many physiological functions (Soldatov et al., 2017), so they are essential nutrients for animals and humans (Zheng et al. 2012a). More than one-third of carotenoids found in nature are of marine origin (Maoka, 2011) and marine mollusks are receiving increasing attentions as important producers of natural carotenoids (Zhang et al., 2014). The muscles of P. amabilis’ foot and siphons are colored bright red because they contain a special C37 skeletal carotenoids (Maoka et al., 2008), making this clam not only has an attractive appearance but also be considered as healthy food. Female bivalves have greater carotene enrichment during either rapid growth stage or gonadal maturation (Zheng et al., 2012b; Borodina, 2018), and individuals with higher total carotenoid content (TCC) or carotenoid enrichment show greater resistance to stress (e.g., hypoxia, Zheng et al., 2012b; Gostyukhina and Andreenko, 2020). These are accordant with our findings in this study that, the percentages of YMs varied in four populations and in male/female groups, and the ratio of YM females is significantly lower than that of male only in autumn instead of spring samples, and, RMs showed stronger hypoxic tolerance during transportation. The structures of RMs’ foot might facilitate the exchange of gases, as in similar cases in the Black Sea invasive bivalve species Anadara kagoshimensis (Tokunaga, 1906), who is highly adaptable to changing environments and its foot not only ensures its active lifestyle but also has some physiological and metabolic features (Shulman, 2014).

In our earlier efforts to develop genetic markers for conservation and aquaculture of this clam, we found that Pam13, a microsatellite marker showing evidence of null allele, had significantly higher observed heterozygosity and higher frequency of allele 308 in YMs than in RMs (7 YMs and 25 RMs, one population, Li et al., 2014). However, in this study, this distinctness disappeared when it was genotyped with other geographic populations and more samples. It is speculated that this accidental association with muscle colour may be due in part to the presence of null allele. Overall, no specific alleles or genotype of total 32 microsatellite marker (20 in this study, and other 12 developed previously, Li et al., 2014) were found associated with muscular coloration, and no stable convergence of genetic differentiations were detected in the same color-muscle P. amabilis. Wang et al. (2015) identified a gene lysophosphatidylcholine acyltransferase 1 (LPCAT1) that is potentially crucial for carotenoid accumulation in Yesso scallop (Patinopecten yessoensis) adductor muscle. This gene shows significantly higher expression and lower methylation level at the 5’ end of the gene body in orange adductor muscle individual than in white ones. This epigenetic switch that controls accumulation of carotenoids in scallop adductor muscle may explain that carotenoid accumulation naturally occurs in female gonads of Yesso scallops (Wang et al., 2015), and implied that the strength of carotenoid accumulation maybe associated with the DNA methylation level. This might also be the reason why in our study we found 5 (8.3%) individuals with intermediate muscle color in BH population on spring. Probably, in P. amabilis there are individual/population differences in carotenoid accumulation and metabolic intensity, which are simultaneously affected by gene expression and environmental factors (e.g., the qualitative composition and annual pattern of succession of diatomic microalgae, Borodina and Soldatov 2016), or even by the gut microbiota (Liu et al., 2020).

This study provided the first molecular evidence of genetic divergence of P. amabilis in South Chinese Sea, which is helpful for the effective conservation and management of the species. These results will also be useful for the seeding production and for the genetic improvement of this clam. The significant genetic differentiation between the Chinese and Vietnamese populations implied the probability to develop particular strains for the benefit of aquaculture. In addition, our results also suggest that P. amabilis could be considered as an important animal model for studying the ways epigenetics and the environmental factors in controlling carotenoids accumulation in sea animals.

3 Materials and Methods
3.1 Animals and measure
57, 43 and 60 P. amabilis clams were collected in Co To Island, Vietnam (Coded for CT), the west of Zhanjiang, Guangdong province (ZJ) and Beihai, Guangxi province (BH) in China, in March 2013, respectively (Figure 5). Another 60 clams were sampled in Beihai in Oct. 2013 (BHA). Shell length, shell height, and shell width were measured for each clam with vernier caliper. The whole weight and shell weight were measured by electronic
balance, the gonadal tissues were observed under a microscope to determine the sex, and the foot and siphon’s color were recorded. The number of individuals, sex ratio, and muscle color differentiation of each population were listed in Table 1. Finally, 32 individuals were randomly selected for further genetic analysis from each population of CT, ZJ, and BH+BHA.

3.2 DNA extraction
DNA was extracted from ethanol-preserved adductor muscles using the modified phenol-chloroform method (Huang et al., 2014). The improved lysis buffer contained Tris-Cl 50 mM, EDTA 25 mM (pH 8), NaCl 150 mM, 1% SDS, and 0.15 μg/μL proteinase K.

3.3 Isolation of microsatellites and primer validation
Microsatellite markers were developed from an enriched genomic library deriving from 6 clams’ DNA. The genomic library enriched for CA-repeats was constructed using 200 ng DNA digested with Rsal (Bio-Red), a biotin labeled (CA)_{10} probe and a magnetic field (Li et al., 2010). Clones with inserts between 400 and 1000 bp were sequenced in both directions using an ABI3730 sequencer. After removing the sequences of linkers and T-vectors, DNA sequences were searched against each other (no other microsatellite sequence of _P. amabilis_ was found in GenBank till 31st Jan. 2021) using Vector NTI Advance 11.0.0 (http://www.invitrogen.com) to check for duplicates. MISA software (http://pgrc.ipk-gatersleben.de/misa/) was used to screen for sequences containing at least 6 di- , 5 tri- , 5 tetra- , 4 penta-, and 3 hexa-, hepta- and octa- nucleotide repeats. Qualified sequences with sufficient flanking regions were selected for primer design with Primer 3 (http://bioinformatics.umd.edu/ primer3 www.cgi) (Wang et al., 2010). Of the 60 sequenced clones, 48 yielded microsatellite sequences for primer design. The M13 (-21) universal leading sequence (5’-TGATAAACGACGCCCAGT-3’) was added to the 5’ end of each forward primer for economic fluorescent labelling of PCR fragments (Schuelke, 2000). Primer validation was performed following Yu et al. (2015), using 3 DNA pools, Premix Taq™ (TaKaRa), and an Eppendorf Master Cycler Gradient (Germany). Primer pairs amplified consistently in 3 DNA pools were selected for further genetic analysis (Table 2).

3.4 PCR conditions and genotyping
In total 96 clams from three geographic populations (CT, ZJ, and BH) were screened for variation by the selected microsatellite loci. PCR was conducted in a 10 μL solution contained ~50 ng template DNA and 1×Premix Taq™ (TaKaRa), and 0.2 pmol M13 tailed forward primer, 0.8 pmol reverse primer, 0.6 pmol M13 (-21) primer labelled with fluorescent dyes (FAM, VIC, NED, and PET; Applied Biosystems) (Wang et al., 2010). PCR was conducted on an Eppendorf Master Cycler (Germany). Cycling parameters were 94°C for 5 min followed by 30 cycles of 94°C for 30 s, annealing (temperatures indicated in Table 3) for 45 s, 72°C for 45 s, followed by 10 cycles of 94°C for 30 s, 53°C for 45 s, and a final extension at 72°C for 10 min. Amplified products were detected and sized on an ABI 3730 Prism Genetic Analyzer (Applied Biosystems) using GS-500LIZ as size standard. Allele scoring was performed with GeneMapper v3.7 (Yu et al., 2015). Samples failing to amplify the first time were re-amplified at least twice and we re-designed the primer pairs of 5 loci, which improved the amplification rate (Pam38) and avoided the multiple-loci amplification in two loci (Pam11 and Pam24).

3.5 Statistical analysis
GENEPOP on the web (http://genepop.curtin.edu.au/) was used to test for deviations from Hardy-Weinberg equilibrium (HWE) for each locus, and for genotypic linkage disequilibrium (LD, 1000 iterations) between all pairs of loci (exact tests, 1000 iterations), as well as to calculate the _F_ST for all populations and population pairs (10 100 permutations, Yu et al., 2015). Exact tests of population differentiation (Raymond and Rousset, 1995) and Mantel test (matrix of pairwise _F_ST/genetic distance and geographical distances, the shortest distance by sea between two populations) were also used GENEPOP on web. The ARLEQUIN 3.0 software (Excoffier et al., 2005) was used to calculate observed (_H_o_) and expected (_H_e_) heterozygosities. The MICRO-CHECKER 2.2.1 software (van Oosterhout et al. 2004) was used for identifying possible genotyping errors and null alleles (1000 randomizations). FSTAT 2.9.3.2 (Goudet, 1995) was used to calculate allelic richness and _F_is per locus and
sample. All tests were corrected for multiple comparisons by Bonferroni’s correction (Rice, 1989). Poptree 2 (Takezaki et al., 2010) was used to construct the Neighbour-joining trees based on the Nei’s D\textsubscript{A} genetic distances (Nei et al., 1983). GENALEX 6 (Peakall and Smouse, 2006) was used to do AMOVA (Analysis of Molecular Variance, Excoffier et al., 1992), estimating and comparing the genetic differentiation of muscular coloration groups among each geographic population.

In order to compare and visualize the genetic differentiations of RMs and Yms group in three geographical populations, each population was divided into two groups by muscular coloration. Neighbour-joining tree for muscular coloration groups was constructed and paired \( F_{ST} \) were calculated as aforesaid. A principal coordinate analysis (PCoA) was performed to display any pattern in the group genetic distances, using GENALEX 6 (Peakall and Smouse, 2006).

Chi-square tests were used to estimate the significances for the association of the frequencies of the genotypes and alleles of the microsatellites, as well as the sex ratio between RMs and YMs.

**Authors’ contributions**

SW, XW and FY contributed equally to this study. YW and BL were responsible for conceptualization, funding acquisition and resources. SW, XW, FY, TL and QL were responsible for experimental investigation. SW, XW and FY were mainly responsible for data processing and picture modification. YW and BL were responsible for the draft preparation. YW proofread the final manuscript before submission. All authors read and approved the final manuscript.

**Acknowledgements**

This study was funded by the National Key R & D Program of China (2018YFC1406500), the Hainan Provincial Key R & D Program (ZDYF2018108), and the earmarked fund for China Agriculture Research System-Mollusk (CARS-48).

**References**

Arnaud S., Monteforte M., Galtier N., Bonhomme F., and Blanc F., 2000, Population structure and genetic variability of pearl oyster Pinctada margaritifera along Pacific coasts from Mexico to Panama, Conservation Genetics, 1(4): 299-308

https://doi.org/10.1023/A:1011575722481

Boosy G., Hendriks R., Smulders M., Van Groenendael J., and Vosman B., 2000, Genetic diversity and the survival of populations, Plant biology, 2(04): 379-395

https://doi.org/10.1055/s-2000-5958

Borodina A., 2018, Carotenoids in the gonads of the Bivalved Mollusk *Anadara kagoshimensis* (Tokunaga, 1906), J. Evol. Biochem. Physiol., 54, (4): 267–272

https://doi.org/10.1134/S0022093018040026

Borodina A., and Soldatov A., 2016, The qualitative composition of carotenoids and their seasonal dynamics in tissues of the bivalve *Anadara kagoshimensis* (Tokunaga, 1906), Rus. J. Marine Biol., 42(2): 166-178

https://doi.org/10.1134/S1063074016020024

Borodina A.V., and Zadorozhny P.A., 2020, The Annual Dynamics of Tissue Carotenoids in a Bivalve Mollusk *Cerastoderma glaucum* (Bruguère, 1789), Journal of Evolutionary Biochemistry and Physiology, 56(1): 1-10

https://doi.org/10.1134/S0022093020010019

Carlsson J., Morrison C.L., and Reece K.S., 2006, Wild and aquaculture populations of the eastern oyster compared using microsatellites, Journal of Heredity, 97(6): 595-598

https://doi.org/10.1093/ijhered/es034

PMid:17071849

Colson I., and Hughes R.N., 2004, Rapid recovery of genetic diversity of dogwhelk (*Nucella lapillus*) populations after local extinction and recolonization contradicts predictions from life-history characteristics, Molecular Ecology, 13(8): 2223-2233

https://doi.org/10.1111/j.1365-294X.2004.02245.x

PMid:15245396

Cowen R.K., Paris C.B., and Srinivasan A., 2006, Scaling of connectivity in marine populations, Science, 311(5760): 522-527

https://doi.org/10.1126/science.1122059

PMid:16357224

Excoffier L., Laval G., and Schneider S., 2005, Arlequin (version 3.0): An integrated software package for population genetics data analysis, Evol Bioinform, 1: 47-50

https://doi.org/10.1177/117693430501000003
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(2): 479-491
https://doi.org/10.1093/genetics/131.2.479
PMid:1644282 PMCID:PMC1205020

Gostyukhina O.I., and Andreenko T.I., 2020, Enzymatic and Low-Molecular Weight Parts of Antioxidant Complex in Two Species of Black Sea Mollusks with Different Resistance to Oxidative Stress: Mytilus galloprovincialis Lam. and Anadara kosgohimensis (Tokunaga, 1906), Biology Bulletin Reviews. 10(1): 38-47
https://doi.org/10.1134/S2079086420010041

Goudet J., 1995, FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics, Journal of Heredity, 86(6): 485-486
https://doi.org/10.1093/oxfordjournals.jhered.a111627

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, Journal of Shellfisheries Research, 23: 379-385

Huang X., Yu F., Zhu J., Chen X., and Wang Y., 2014, Characterization of novel microsatellite markers in the black-lipped pearl oyster, Pinctada margaritifera (Linnaeus, 1758), Conservation Genetics Resources, 6(1): 63-64
https://doi.org/10.1007/s12686-013-0004-8

Li Q., Li Y., Wang Y., and Luo B., 2013, A preliminary study on annual gonadal development of Paphia schnelliana in the Guangxi sea area, J. Biol., 30: 37-39

Li Q., Luo B., Li Y., and Wang Y., 2012, Study on the larval development and growth of Paphia schnelliana. Fish. Sci. Technol. Inform., 39: 275-278

Li Q., Yu F., Luo B., Li Y., Zhu J., and Wang Y., 2014, Microsatellite markers development in the clam, Paphia amabilis (Philippi, 1847), Conservation genetics resources, 6(1): 245-246
https://doi.org/10.1007/s12686-013-0072-9

Li Z., Li J., Feng X., Xie N., Feng J., and Yue G., 2010, Sixteen polymorphic microsatellites for breeding of Chinese soft-shelled turtles (Pelodiscus sinensis), Anim Genet, 41: 446-447
https://doi.org/10.1111/j.1365-2052.2010.02029.x
PMid:20331610

Liu H., Tan K.S., Zhang X., Zhang H., Cheng D., Ting Y., Li S., Ma H., and Zheng H., 2020, Comparison of Gut Microbiota Between Golden and Brown Noble Scallop Chlamys nobilis and Its Association With Carotenoids, Frontiers in Microbiology, 11: 36
https://doi.org/10.3389/fmicb.2020.00336
PMid:32117095 PMCID:PMC7018768

Mahidol C., Na-Nakorn U., Sukmanomon S., Taniguchi N., and Nguyen T.T., 2007, Mitochondrial DNA diversity of the Asian moon scallop, Amusium pleuronectes (Pectinidae), in Thailand, Mar Biotechnol (NY), 9(3): 352-359
https://doi.org/10.1007/s10126-006-6137-y
PMid:17549563

Maoka T., 2011, Carotenoids in marine animals, Mar Drugs, 9(2): 278-293
https://doi.org/10.3390/md9020278
PMid:2156799 PMCID:PMC3093257

Maoka T., Akimoto N., Yim M.-J., Hosokawa M., and Miyashita K., 2008, New C37 skeletal carotenoid from the clam, Paphia amabilis, Journal of Agricultural and Food Chemistry, 56(24): 12069-12072
https://doi.org/10.1021/jf080271b
PMid:19053385

Nei M., Tajima F., and Tateno Y., 1983, Accuracy of estimated phylogenetic trees from molecular data, Journal of molecular evolution, 19(2): 153-170
https://doi.org/10.1007/BF02300753
PMid:6571220

Peakall R., and Smouse P.E., 2006, genalex 6: genetic analysis in Excel. Population genetic software for teaching and research, Molecular Ecology Notes, 6(1): 288-295
https://doi.org/10.1111/j.1471-8286.2005.01155.x

Qiu Y., Huang X., Lu H., Shi Y., Wang A., and Wang Y., 2013, Development of EST-derived microsatellite markers in the pearl oyster Pinctada martensi (Dunker) for genetic resource monitoring, Conservation Genetics Resources, 5(2): 401-403
https://doi.org/10.1007/s12686-012-9813-4

Raymond M., and Rousset F., 1995, GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism, Journal of Heredity. 68: 248-249
https://doi.org/10.1093/oxfordjournals.jhered.a111573

Rice W.R., 1989, Analyzing tables of statistical tests, Evolution, 43: 223-225
https://doi.org/10.1111/j.1558-5646.1989.tb04220.x
PMid:28568501

Schuelke M., 2000, An economic method for the fluorescent labeling of PCR fragments, Nature biotechnology, 18: 233-234
https://doi.org/10.1038/72708
PMid:10657137
Zheng G., Ed., (2014) Mollusks of the Black Sea: Elements of Comparative and Ecological Biochemistry, Sevastopol.
Soldatov A., Gostyukhina O., Borodina A., and Golovina I., 2017, Glutathione antioxidant complex and carotenoid composition in tissues of the bivalve mollusk *Anadara kagoshimensis* (Tokunaga, 1906), J. Evol. Biochem. Physiol., 53(4): 289-297
https://doi.org/10.1134/S0022993017040056
Su D., Cheng L., Li X., Chen X., Zhu J., and Wang Y., 2015, Thirty-three microsatellite markers for the Asian moon scallop Amusium pleuronectes (Linnaeus, 1758), Conservation Genetics Resources, 7(2): 547-549
https://doi.org/10.1007/s12686-015-0421-y
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, Molecular biology and evolution, 27(4): 747-752
https://doi.org/10.1093/molbev/msq312
PMid:20022889 PMCid:PMC2877541
Van Oosterhout C., Hutchinson W., Wills D., and Shipley P., 2004, MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data, Molecular Ecology Notes, 4: 535-538
https://doi.org/10.1111/j.1471-8286.2004.00684.x
Wang S., Lv J., Zhang L., Dou J., Sun Y., Li X., Fu X., Dou H., Mao J., Hu X., and Bao Z., 2015, MethylRAD: a simple and scalable method for genome-wide DNA methylation profiling using methylation-dependent restriction enzymes, Open Biology, 5(11): 150130
https://doi.org/10.1098/rsob.150130
PMid:26581575 PMCid:PMC4680569
Wang Y., Huang X., Wu X., Li Q., Luo B., and Zhu J., 2013, *Paphia amabilis* (Philippi, 1847), a Promising Bright Red Clam in Aquaculture, International Journal of Aquaculture, 3: 1-3
Wang Y., Wang A., and Guo X., 2010, Development and Characterization of Polymorphic Microsatellite Markers for the Northern Quahog (*Mercenaria mercenaria*; Linnaeus, 1758), Journal of Shellfish Research, 29(1): 77-82, 76
https://doi.org/10.2983/035.029.0130
Waples R.S., 2014, Testing for Hardy–Weinberg proportions: have we lost the plot?, Journal of Heredity, 106(1): 1-19
https://doi.org/10.1093/hered/ess062
PMid:25425676
Wu X., Yu F., Zhu J., and Wang Y., 2014, Changes of population genetic structure of silver-lipped pearl oyster *Pinctada maxima* (Jameson, 1901) in Sanya, China under pressure from poaching and environment changes, Journal of Shellfish Research, 33(2): 666
Xu Z., 1987, The Paphia resources along coast of Fujian Province, China J. Fujian Fish., 2: 44-46
Yang S., Bao X., Chen C., and Chen F., 2003, Analysis on characteristics and mechanism of current system in west coast of Guangdong Province in the summer, Acta Oceanologica Sinica, 25: 1-8
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, Frontiers in Marine Science, 7: 516
https://doi.org/10.3389/fmars.2020.00516
Yu F., Su D., Cheng L., Chen X., Wang J., and Wang Y., 2015, Microsatellite genetic variation between and within cultured pearl oyster *Pteria penguia* populations, International Journal of Marine Science, 5: 1-8
Zhang X., and Qi Z. (1960) *The bivalve mollusk in Southern Sea*, Beijing, Science Press
Zhang Y., Zhang L., Sun J., Qiu J., Hu X., Hu J., and Bao Z., 2014, Proteomic analysis identifies proteins related to carotenoid accumulation in Yesso scallop (*Patinopecten yessoensis*), Food Chem, 147: 111-116
https://doi.org/10.1016/j.foodchem.2013.09.078
PMid:24206693
Zheng H., Liu H., Liu W., Sun Z., and Zhang Q., 2012a, Changes of total carotenoid and lipid content in scallop tissues of Chlamys nobilis (Bivalve: Pectinidae) during gonad maturation, Aquaculture, 342-343: 7-12
https://doi.org/10.1016/j.aquaculture.2012.01.037
Zheng H., Liu H., Zhang T., Wang S., Sun Z., Liu W., and Li Y., 2010, Total carotenoid differences in scallop tissues of *Chlamys nobilis* (Bivalve: Pectinidae) with regard to gender and shell colour, Food Chemistry, 122(4): 1164-1167
https://doi.org/10.1016/j.foodchem.2010.03.109
Zheng H., Zhang Q., Liu H., Liu W., Sun Z., Li S., and Zhang T., 2012b, Cloning and expression of vitelligenin (Vg) gene and its correlations with total carotenoids content and total antioxidant capacity in noble scallop *Chlamys nobilis* (Bivalve: Pectinidae), Aquaculture, 366-367: 46-53
https://doi.org/10.1016/j.aquaculture.2012.08.031
Zouros E., and Foltz D., 1984, Possible explanations of heterozygote deficiency in bivalve mussels, Malacologia, 25(2): 583-591