

# First genetic assessment of brackish water polychaete *Tylorrhynchus heterochaetus*: mitochondrial COI sequences reveal strong genetic differentiation and population expansion in samples collected from southeast China and north Vietnam

Xing-Han Chen<sup>1#</sup>, Sen Yang<sup>2#</sup>, Wei Yang<sup>1</sup>, Yuan-Yuan Si<sup>1</sup>, Rui-Wen Xu<sup>1</sup>, Bin Fan<sup>1</sup>, Le Wang<sup>3</sup>, Zi-Ning Meng<sup>4,5,\*</sup>

<sup>1</sup> Department of Food and Environmental Engineering, Key Laboratory for Marine Estuary Fishery Resources Protection of Yangjiang City, Yangjiang Polytechnic, Yangjiang, Guangdong 529500, China

<sup>2</sup> College of Animal Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, Guangdong 520175, China

<sup>3</sup> Temasek Life Sciences Laboratory, National University of Singapore, Singapore 117604, Singapore

<sup>4</sup> State Key Laboratory of Biocontrol, Institute of Aquatic Economic Animals, and the Guangdong Province Key Laboratory for Aquatic Economic Animals, Life Science School, Sun Yet-Sen University, Guangzhou, Guangdong 510275, China

<sup>5</sup> Southern Laboratory of Ocean Science and Engineering, Zhuhai, Guangdong 519000, China

## ABSTRACT

*Tylorrhynchus heterochaetus* is a widespread benthic polychaete worm found in coastal brackish waters of the west Pacific. It has high ecological and economic value as a biomarker of water quality and as a high-quality feed in aquaculture and fisheries and is considered a delicacy in some areas of Asia. However, it has experienced a marked reduction in recent years due to overexploitation as well as changes in the environment and climate. Here, to comprehensively understand its genetic background and thus provide insights for better conservation and utilization of this species, we assessed the genetic variability and demographic history of *T. heterochaetus* individuals sampled from eight locations along the coasts of southeast China and

north Vietnam based on mitochondrial cytochrome c oxidase I (COI) sequences. We observed high haplotype diversity ( $Hd$ ), with an average of 0.926, but relatively low nucleotide diversity ( $\pi$ ), with a mean of 0.032 across all samples. A total of 94 polymorphic sites and 85 haplotypes were identified among 320 individuals. The pairwise genetic distances among haplotypes ranged from 0.001 to 0.067, with the high intraspecific divergence possibly reflecting geographic isolation and gene pool fragmentation. Significant genetic structures were revealed among the studied locations; specifically, the eight locations could be treated as six genetically

Received: 03 July 2019; Accepted: 29 October 2019; Online: 29 October 2019

Foundation items: This work was supported by the Industry-Academic Research Cooperation Program of Academician Workstation of Guangdong Province (2013B090400017), Natural Science Foundation of Guangdong Province (2016A030307036), and Yangfan Innovative & Entrepreneurial Research Team Project for Guangdong Province (201312H10)

#Authors contributed equally to this work

\*Corresponding author, E-mail: mengzn@mail.sysu.edu.cn

DOI: 10.24272/j.issn.2095-8137.2020.006

## Open Access

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2020 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

different populations based on pairwise  $\Phi_{ST}$  results (0.026–0.951,  $P < 0.01$ ). A significant pattern of isolation-by-distance was detected between the genetic and geographic distances ( $r = 0.873$ ,  $P = 0.001$ ). Three geographic lineages were defined based on phylogenetic tree and network analyses of *COI* haplotypes. AMOVA results indicated that genetic variations mainly occurred among the three lineages (89.96%). Tests of neutrality and mismatch distribution suggested that *T. heterochaetus* underwent recent population expansion. These results provide the first report on the genetic status of *T. heterochaetus* and will be valuable for the management of genetic resources and better understanding of the ecology and evolution in this species.

**Keywords:** *Tylorrhynchus heterochaetus*; Mitochondrial DNA; Genetic diversity; Population structure; Demographic history

## INTRODUCTION

*Tylorrhynchus heterochaetus*, a member of the polychaete family Nereididae (Annelida: Phyllodoceida), is a widespread benthic invertebrate found in brackish waters along the coasts of China, Japan, and Southeast Asia (Tuan, 2018). Due to its high sensitivity to water quality, the species is widely used as biomarker of marine environmental conditions (Dean, 2008). Moreover, it has great potential in both aquaculture and recreational fisheries as a high-quality feed (Costa et al., 2006) and is also a favored and relatively expensive delicacy in some areas of Asia, such as Vietnam and southern China (Glasby & Timm, 2008).

During the breeding season, *T. heterochaetus* adults tend to swim to higher salinity waters and aggregate for spawning, and then die after releasing gametes. The salinity levels suitable for reproduction range from 10 to 13 ppt, and the hatching rate can be significantly affected by different salinities (Duan et al., 2017). For further development, however, early setiger larvae (3–5 d after hatching) prefer a low salinity environment, where they settle into mud until they reach sexual maturity in the following year. Thus, *T. heterochaetus* individuals are usually confined to the muddy bottom in brackish water estuaries and are thus extremely vulnerable to the impact of sudden changes in the external environment. In recent years, the increasing demand for commercial utilization, together with changes in both climate and environment, such as water pollution, ocean salinization, and habitat fragmentation, has led to a marked reduction in the natural resources of *T. heterochaetus*, even in historically high-yield habitats (Tuan, 2018).

Based on the above threats, a comprehensive understanding of the genetic background of *T. heterochaetus* could greatly facilitate its conservation and management and thus better utilization of its genetic resources. To date,

however, studies on *T. heterochaetus* have mainly focused on reproductive physiology (Kaoawa, 1954; Okada, 1952; Sato & Osanai, 1990; Tuan, 2018) and the function of individual genes (Green et al., 1995; Suzuki & Gotoh, 1986; Suzuki et al., 1990). As such, the genetic diversity and structure of natural *T. heterochaetus* populations remain unclear.

Mitochondrial DNA (mtDNA) is an effective tool for molecular phylogenesis and population genetics analysis of polychaetes, such as *Marenzelleria* (Blank & Bastrop, 2009), *Phascolosoma esculenta* (Gao et al., 2018), *Perinereis aibuhitensis* (Liu et al., 2012), *Pygospio elegans* (Kesäniemi et al., 2012), *Aglaophamus australiensis*, and *Nephtys longipes* (Smith et al., 2015), due to the advantages of maternal inheritance, relatively fast mutation rate, and non-recombination (Birky et al., 1989). In the present study, we used the cytochrome c oxidase subunit I (*COI*) gene from mtDNA to investigate the genetic diversity, population structure, and demographic history of *T. heterochaetus* along the coasts of southeast China and north Vietnam. This genetic survey will provide useful information for the development of effective conservation and utilization strategies of this species.

## MATERIALS AND METHODS

### Sample collection and DNA extraction

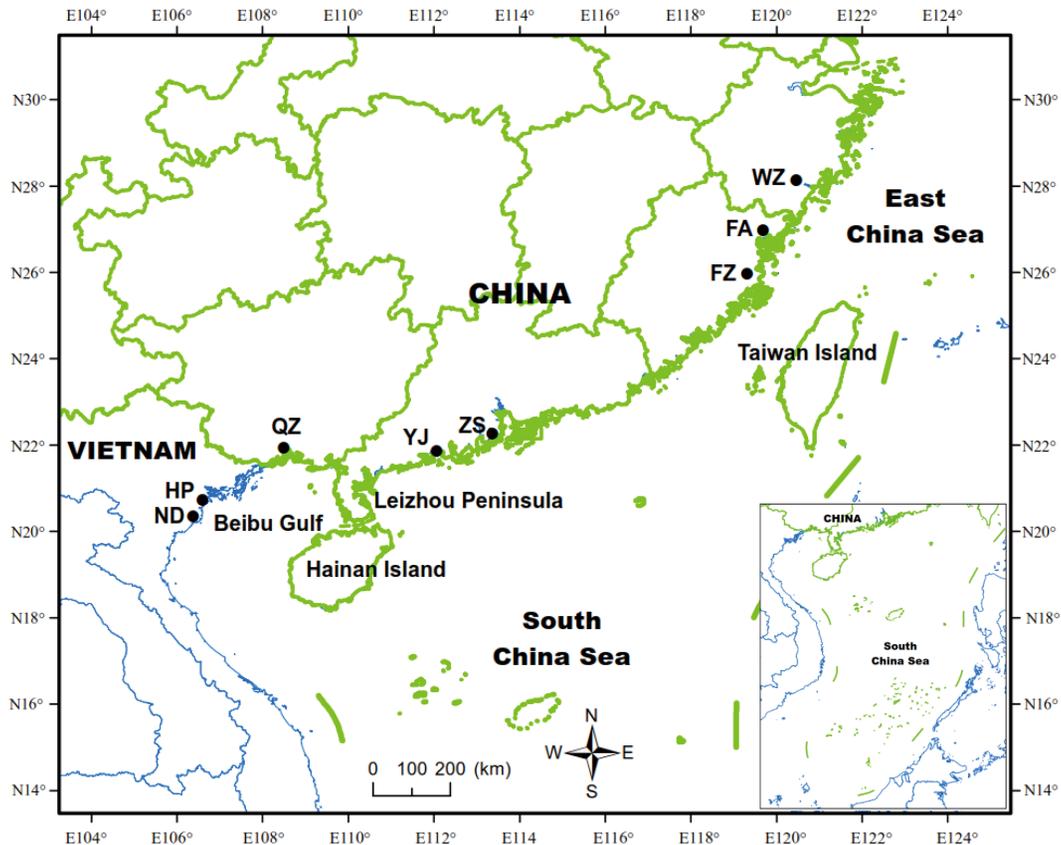
A total of 320 individuals from eight locations were collected during 2016–2018 from the coastal waters of southeast China and north Vietnam. Two locations were along the coast of Vietnam (Hai Phong (HP), Nam Dinh (ND)) and six locations were along the coast of China (Wenzhou (WZ), Fu' an (FA), Fuzhou (FZ), Yangjiang (YJ), Zhongshan (ZS), Qinzhou (QZ)). Detailed geographic locations and sampling information are shown in Figure 1 and Table 1. The samples were stored in 95% ethyl alcohol at  $-20^{\circ}\text{C}$  until DNA extraction. Muscle tissues ( $\sim 10$  mg for each individual) were dissected and genomic DNA was extracted using a Genomic DNA Extraction Kit (Tiangen Biotech, DP304, Beijing, China) according to the manufacturer's protocols. The extracted DNA was stored at  $-80^{\circ}\text{C}$  before use.

### Gene amplification and sequencing

Partial sequences of mtDNA *COI* were amplified and sequenced using universal DNA primers *COI*-LCO1490 and HCO2198 (Folmer et al., 1994). Polymerase chain reaction (PCR) amplification was conducted in a 50  $\mu\text{L}$  volume containing 0.5  $\mu\text{mol/L}$  of each primer, 0.2 mmol/L of each dNTP, 30 ng of template DNA, 1.5 mmol/L  $\text{MgCl}_2$ , 1 $\times$ PCR buffer, and 1 unit of Taq DNA polymerase (Fermentas, Thermo Scientific, USA). PCR was conducted with an initial denaturation at  $94^{\circ}\text{C}$  for 5 min, followed by 35 cycles of  $94^{\circ}\text{C}$  for 30 s,  $55^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 45 s, with a final extension at  $72^{\circ}\text{C}$  for 3 min. The amplified products were then purified and sequenced on an ABI Prism 3730 DNA sequencer (Applied Biosystems, USA) using both forward and reverse primers individually.

### Data analysis

Using the *COI* gene from the complete mitochondrial genome



**Figure 1** Sampling locations of *T. heterochaetus* used in this study (indicated by solid cycles)

Corresponding location abbreviations are as in Table 1.

**Table 1** Sampling information on *T. heterochaetus* in this study

| Location code | Locality             | Country | Sample size ( <i>n</i> ) | Longitude (E) | Latitude (N) |
|---------------|----------------------|---------|--------------------------|---------------|--------------|
| WZ            | Wenzhou, Zhejiang    | China   | 40                       | 121°18'       | 28°38'       |
| FA            | Fu'an, Fujian        | China   | 40                       | 119°40'       | 26°59'       |
| FZ            | Fuzhou, Fujian       | China   | 40                       | 119°18'       | 25°58'       |
| YJ            | Yangjiang, Guangdong | China   | 40                       | 112°02'       | 21°51'       |
| ZS            | Zhongshan, Guangdong | China   | 40                       | 113°21'       | 22°15'       |
| QZ            | Qinzhou, Guangxi     | China   | 40                       | 108°29'       | 21°56'       |
| HP            | Hai Phong            | Vietnam | 40                       | 106°35'       | 20°43'       |
| ND            | Nam Dinh             | Vietnam | 40                       | 106°22'       | 20°20'       |

sequence of *T. heterochaetus* as a reference sequence (Chen et al., 2016), sequence data were aligned using Cluster X 2.0 (Thompson et al., 1997). DNA sequence polymorphisms, including number of polymorphic sites (*S*), number of haplotypes (*H*), haplotype diversity (*Hd*), nucleotide diversity ( $\pi$ ), and average number of nucleotide differences (*K*), were estimated using DnaSP 5.10 (Librado & Rozas, 2009). To examine the genealogical relationships among mtDNA haplotypes, a haplotype network was constructed based on the median-joining algorithm in Network 5.0 (Bandelt et al., 1999). A neighbor-joining phylogenetic tree of *COI* haplotypes was constructed in MEGA 7.0 (Kumar et al., 2016) with the Kimura-2-parameter model (Saitou & Nei, 1987) and 1 000

bootstrap replicates. *Perinereis aibuhitensis*, a sea worm from the family Nereididae, was used as the outgroup (GenBank accession No.: NC023943.1). Pairwise genetic distances among haplotypes or different locations of *T. heterochaetus* were also calculated using MEGA 7.0 based on the Kimura-2-parameter model.

Pairwise  $\Phi_{ST}$ , analysis of molecular variance (AMOVA), as well as correlation between genetic and geographic distance (coastline distance between sampling sites measured by Google Earth) estimated with the Mantel test, were all calculated in Arlequin v3.5 with 10 000 permutations (Excoffier et al., 2005). For examining demographic history of *T. heterochaetus*, Tajima's *D* (Tajima, 1989) and Fu's *F<sub>s</sub>* (Fu,

1997) tests were used to examine the neutrality of coding sequences with Arlequin v3.5. Additionally, mismatch distribution analysis was also performed to investigate the hypothesis of population expansion. The fitness between the observed and simulated distributions was tested using the sum of squared deviations (SSD) as well as Harpending's raggedness index (HRI) (Harpending, 1994). Time of population expansion was calculated with the formulas  $\tau = 2ut$  and  $u = 2\mu k$ , where  $t$  is the time since expansion,  $u$  is the mutation rate for the COI gene (Rogers & Harpending, 1992),  $k$  is the number of nucleotides, and  $\mu$  is the mutation rate. In this study, a mutation rate of 2% per million years was used as suggested by Olson et al. (2009) for another polychaete *Hobsonia florida*.

## RESULTS

### Genetic diversity

A 709 bp fragment of the COI gene was obtained after alignment. The average nucleotide frequency was: T=29.7%, C=24.4%, A=29.1%, and G=16.9%. Of the 320 sequences from the eight locations, a total of 94 variable sites were observed, including 32 singleton variable sites and 62 parsimony informative sites. Genetic diversity indices are presented in Table 2. A total of 85 haplotypes (GenBank accession No.: MK614603–MK614686) were identified, most of which (68 out of 85) were represented by a single individual, and pairwise genetic distances among haplotypes varied from 0.001 to 0.067. The number of haplotypes at a location ranged from 8 to 23. Overall, most locations showed moderate to high haplotype diversity (0.237–0.949) due to the large number of rare haplotypes. However, nucleotide diversity was relatively low, ranging from 0.0004 to 0.00912, across all studied locations. The WZ location exhibited the highest haplotype diversity ( $Hd=0.949$ ) and nucleotide diversity ( $\pi=0.009$ ).

### Population structure and phylogenetic analysis

Pairwise  $\Phi_{ST}$  values ranged from  $-0.013$  to  $0.951$  for all locations and were highly significant ( $P<0.01$ ), except for two

**Table 2 Genetic diversity of eight *T. heterochaetus* locations based on COI gene sequences**

| Location (lineage) | <i>n</i> | <i>S</i> | <i>H</i> | <i>Hd</i> | $\pi$   | <i>k</i> |
|--------------------|----------|----------|----------|-----------|---------|----------|
| WZ                 | 40       | 32       | 23       | 0.949     | 0.00912 | 6.465    |
| FA                 | 40       | 13       | 10       | 0.776     | 0.00474 | 3.362    |
| FZ                 | 40       | 16       | 13       | 0.783     | 0.00470 | 3.329    |
| ZS                 | 40       | 18       | 6        | 0.237     | 0.00127 | 0.900    |
| YJ                 | 40       | 12       | 13       | 0.654     | 0.00140 | 0.994    |
| QZ                 | 40       | 11       | 10       | 0.441     | 0.00104 | 0.735    |
| HP                 | 40       | 17       | 16       | 0.645     | 0.00159 | 1.124    |
| ND                 | 40       | 7        | 8        | 0.618     | 0.00105 | 0.744    |
| Lineage A          | 120      | 42       | 39       | 0.861     | 0.00711 | 5.043    |
| Lineage B          | 80       | 29       | 18       | 0.660     | 0.00175 | 1.239    |
| Lineage C          | 120      | 29       | 29       | 0.767     | 0.00274 | 1.942    |
| Overall            | 320      | 94       | 85       | 0.926     | 0.03215 | 22.794   |

*n*: Sample size; *H*: Number of haplotypes; *S*: Number of polymorphic sites; *Hd*: Haplotype diversity;  $\pi$ : Nucleotide diversity; *k*: Mean number of pairwise differences. WZ: Wenzhou; FA: Fu'an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dinh. Lineage A=WZ+FA+FZ, Lineage B=ZS+YJ, Lineage C=QZ+HP+ND.

close location pairs (FA-FZ and HP-ND), whereas pairwise genetic distance varied from 0.001 to 0.063 (Table 4). Thus, all eight locations sampled could be treated as six genetically different populations (WZ, FA+FZ, ZS, YJ, QZ, HP+ND, Table 4). The topologies produced from both the haplotype phylogenetic tree (Figure 2) and haplotype median-joining network (Figure 3) showed a consistent structure. All eight locations could be characterized into three geographically distinguishable lineages (i.e., lineage A: WZ, FA+FZ; lineage B: ZS, YJ; lineage C: QZ, HP+ND, Figures 1–3). The AMOVA results based on the three lineages revealed that genetic variation mainly occurred among lineages (89.86%). Only 6.67% and 3.47% of variation occurred within populations and among populations within lineages, respectively (Table 3), suggesting strong genetic divergence among the different regions. Additionally, a significant pattern of isolation-by-distance was detected across all studied locations using the Mantel test ( $r=0.8731$ ,  $P=0.001$ ).

**Table 3 AMOVA results of *T. heterochaetus* based on mtDNA**

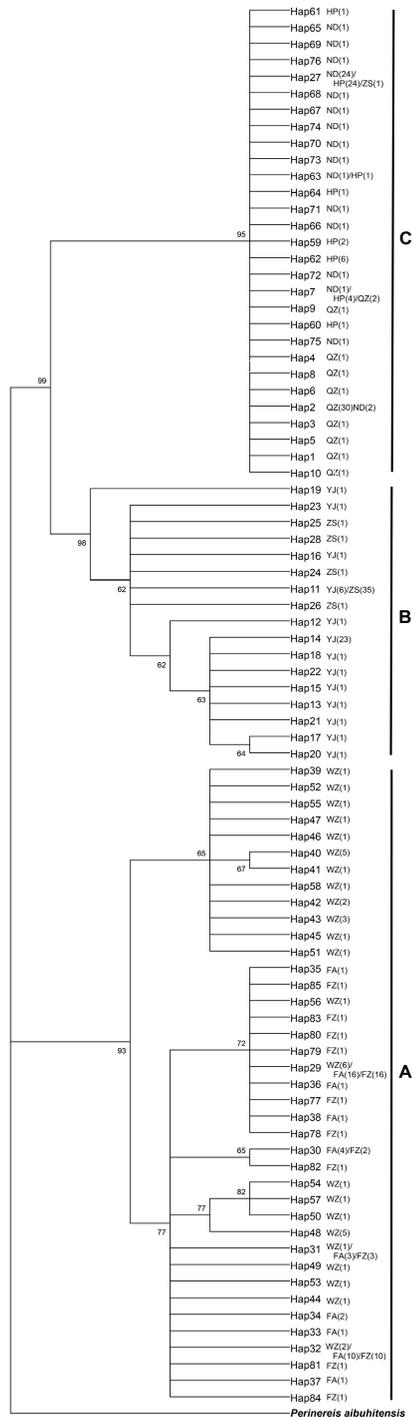
| Source of variation               | <i>d. f.</i> | Sum of squares | Variance component | Percentage of variation (%) | Fixation indices |
|-----------------------------------|--------------|----------------|--------------------|-----------------------------|------------------|
| Among lineages                    | 2            | 3171.121       | 14.871 Va          | 89.86                       | $F_{SC}=0.342^*$ |
| Among populations within lineages | 5            | 120.279        | 0.573 Vb           | 3.47                        | $F_{ST}=0.933^*$ |
| Within populations                | 312          | 344.225        | 1.103 Vc           | 6.67                        | $F_{CT}=0.818^*$ |
| Total                             | 319          | 3635.625       | 16.548             |                             |                  |

\*:  $P<0.05$ ; *d. f.*: Degree of freedom; Va: Variance component due to differences among lineages, Vb: Variance component due to differences among populations within lineages, Vc: Variance component due to differences among individuals within populations;  $F_{SC}=Vb/(Vb+Vc)$ ,  $F_{ST}=(Va+Vb)/(Va+Vb+Vc)$ ,  $F_{CT}=Va/(Va+Vb+Vc)$ .

### Historical demography

As genetic differences were not significant in the two geographically close location pairs FA-FZ and HP-ND, we treated them as a whole when analyzing demographic history.

Both Tajima's *D* and Fu's *F* tests showed negative values for all genetically differentiated populations (Table 5), indicating departure from mutation-drift equilibrium and possible population demographic expansion. Furthermore, mismatch



**Figure 2 Neighbor-joining phylogenetic tree of *T. heterochaetus* based on *COI* haplotypes**

Bootstrap values >60% are shown near nodes (1 000 replicates), number in parentheses after location code is frequency of haplotypes. *Perinereis aibuhitensis* (GenBank accession No.: NC023943.1) was used as an outgroup. WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dinh.

**Table 4 Pairwise  $\Phi_{ST}$  (below diagonal) and genetic distance (above diagonal) among different *T. heterochaetus* locations**

| Location code | WZ           | FA           | FZ           | ZS           | YJ           | QZ           | HP    | ND    |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|-------|
| WZ            | –            | 0.009        | 0.009        | 0.058        | 0.057        | 0.061        | 0.061 | 0.060 |
| FA            | <b>0.226</b> | –            | 0.005        | 0.059        | 0.058        | 0.062        | 0.063 | 0.063 |
| FZ            | <b>0.238</b> | –0.013       | –            | 0.059        | 0.058        | 0.062        | 0.063 | 0.063 |
| ZS            | <b>0.906</b> | <b>0.946</b> | <b>0.946</b> | –            | 0.002        | 0.020        | 0.019 | 0.019 |
| YJ            | <b>0.903</b> | <b>0.944</b> | <b>0.944</b> | <b>0.378</b> | –            | 0.022        | 0.021 | 0.021 |
| QZ            | <b>0.912</b> | <b>0.950</b> | <b>0.951</b> | <b>0.941</b> | <b>0.942</b> | –            | 0.005 | 0.005 |
| HP            | <b>0.907</b> | <b>0.944</b> | <b>0.947</b> | <b>0.922</b> | <b>0.926</b> | <b>0.711</b> | –     | 0.001 |
| ND            | <b>0.911</b> | <b>0.951</b> | <b>0.951</b> | <b>0.936</b> | <b>0.938</b> | <b>0.771</b> | 0.017 | –     |

$\Phi_{ST}$  values in bold type indicate statistical significance ( $P < 0.01$ ). WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dinh.

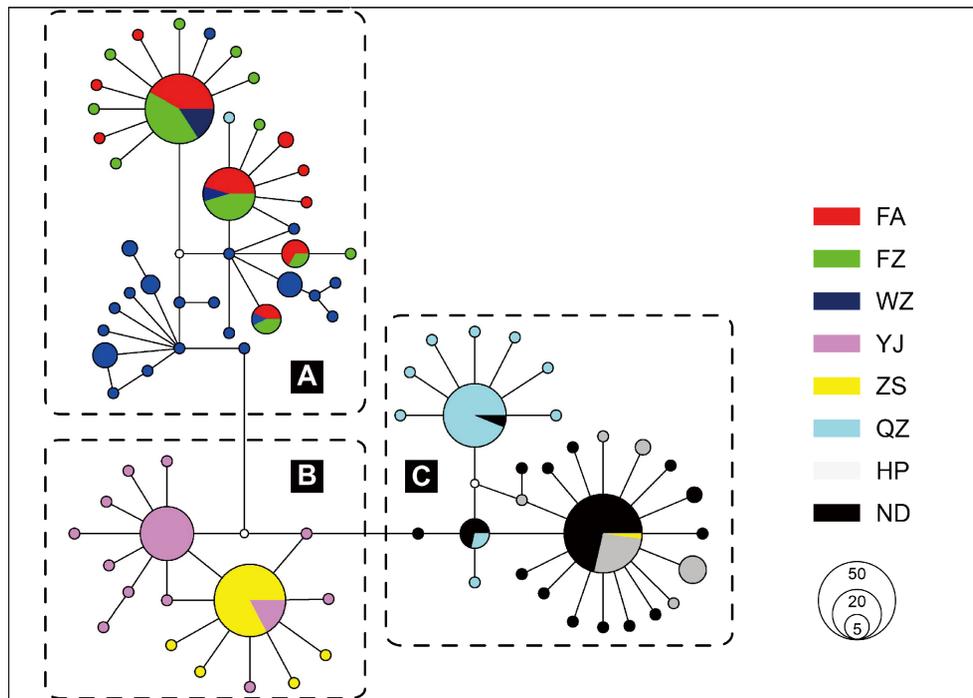
distribution did not differ significantly from the model of sudden expansion when using either *SSD* or *HRI* for goodness-of-fit (Table 5), further supporting the hypothesis of population expansion in *T. heterochaetus*. The  $t$  value across populations varied from 0.969 to 13.936. Using the mutation rate of 2% per million years, it was estimated that the *T. heterochaetus* population expansion occurred about 17 000–246 000 years ago in the middle to late Pleistocene.

## DISCUSSION

### Genetic diversity

This study examined the population genetic variability of *T. heterochaetus* using mtDNA for the first time. Compared with other polychaetes, the northern location (FA, FZ, WZ) in lineage A showed high haplotype and nucleotide diversities ( $Hd=0.861$ ,  $\pi=0.007$  11; Figure 1, Table 1), differing from several other species, such as *Branchipolynoe symmytilida* ( $Hd=0.970$ ,  $\pi=0.007$ ; Plouviez et al., 2009), *Owenia fusiformis* clade 1 and 2 ( $Hd=0.924$ – $0.978$ ,  $\pi=0.007$  2– $0.007$  7; Jolly et al., 2006), and *Aglaophamus australiensis* ( $Hd=0.78$ – $0.99$ ; Smith et al., 2015). However, remarkably lower nucleotide diversity was detected in the southern locations (YJ, ZS, QZ, HP, and ND; Figure 1, Table 1), which may be due to the commercial overexploitation of wild stocks as *T. heterochaetus* is a very popular specialty food in this region. Although small-scale artificial breeding of *T. heterochaetus* has been conducted in southern China and Vietnam, production still comes primarily from natural exploitation.

Previous meta-analysis revealed that overharvesting can drive the decay of genetic diversity in many highly abundant marine species (Pinsky & Palumbi, 2014). Genetic diversity is closely related to the long-term adaptability and survivability of populations, especially in suddenly and drastically changing marine environments (Barrett & Schluter, 2008). Thus, the strikingly varied genetic diversity among different *T. heterochaetus* locations, as shown in this study, should be an important consideration in relation to conservation strategies or artificial breeding programs for future aquaculture enhancement.



**Figure 3** Median-joining network of *T. heterochaetus* based on *COI* haplotypes

Each circle represents a haplotype and area is proportional to its frequency. Colors of circles denote geographic origin, white dots represent hypothetical intermediate haplotypes. WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dinh.

**Table 5** Parameters of neutrality test and mismatch distribution analysis for six *T. heterochaetus* populations

| Population | Neutrality test   |                           | Mismatch distribution analysis |            |                  | Expansion time |
|------------|-------------------|---------------------------|--------------------------------|------------|------------------|----------------|
|            | Tajima's <i>D</i> | Fu's <i>F<sub>s</sub></i> | <i>SSD</i>                     | <i>HRI</i> | Tau ( <i>t</i> ) | <i>t</i> (Ma)  |
| WZ         | -0.486            | -7.786**                  | 0.023                          | 0.024      | 6.425            | 0.113          |
| FA+FZ      | -0.395            | -5.076*                   | 0.052                          | 0.123      | 4.834            | 0.085          |
| ZS         | -2.571**          | -1.607                    | 0.002                          | 0.408      | 13.936           | 0.246          |
| YJ         | -1.995**          | -11.090**                 | 0.001                          | 0.06       | 1.062            | 0.019          |
| QZ         | -2.172**          | -7.960**                  | 0.000 1                        | 0.124      | 1.103            | 0.019          |
| HP+ND      | -2.333**          | -24.297**                 | 0.003                          | 0.081      | 0.969            | 0.017          |

\*:  $P < 0.05$ , \*\*:  $P < 0.01$ . WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dinh.

### Population structure and phylogenetic analysis

According to the pairwise  $\Phi_{ST}$  values (Table 4), we identified six genetically different populations (WZ, FA+FZ, ZS, YJ, QZ, HP+ND) among the eight locations in this study. Given the geographic positions and significant divergences among the populations, Leizhou Peninsula and Hainan Island appear to be remarkable oceanographic barriers that block gene flow among the different populations inside the Beibu Gulf and others (Figure 1), as also demonstrated in *Perinereis aibuhitensis* populations from the southern coastal zone of China (Liu et al., 2014). Over broader geographic scales, long distance/isolation between populations may play an important role in the high level of genetic structuring. In fact, an evident isolation by distance pattern was detected using the Mantel test ( $r=0.873$  1,  $P=0.001$ ). As for the genetic homogeneity in

the two location pairs FA-FZ and HP-ND, the short distance along the coastline (<60 km, measured by Google Earth) may not be strong enough to prevent effective gene flow between the two adjacent estuaries.

Geographically, the southern populations (YJ, ZS, QZ, HP+ND) are isolated from the northern populations (FA+FZ, WZ) by several main water systems, such as the Pearl River, Hanjiang River, and Minjiang River. Such isolation could be a strong barrier to gene flow. In addition, according to our resource investigations over the last five years, many unsuitable habitat patches exist along the coastline among the sampling sites. This habitat discontinuity could also lead to the significant genetic divergences observed in different populations.

Apart from geographic factors, its narrow habitat niche is

also suspected to be partly responsible for the high structuring among *T. heterochaetus* populations. The pairwise  $\Phi_{ST}$  results among sampling sites suggest striking genetic divergence, except for the FA-FZ and HP-ND pairs, which exhibit close geographical positions (Figure 1, Table 4), implying that effective dispersal of *T. heterochaetus* over large distances is limited. In marine invertebrates with sedentary adults, a longer planktonic larval stage is correlated with increased dispersal capacity and higher connectivity among populations (Cowen & Sponaugle, 2009; Kyle & Boulding, 2000). However, pelagic larva duration may not always be a good predictor of gene flow and population structure (Weersing & Toonen, 2009). Although it has a relatively long planktonic larval stage (~20 d), wild *T. heterochaetus* individuals favor the muddy bottoms of brackish environments with low salinity, such as estuaries (Tuan, 2018). This specialized habitat and narrow salinity requirement may limit colonization potential of the dispersed larvae. A high level of genetic divergence among invertebrates inhabiting estuarine systems is not uncommon (e. g., Darling et al., 2004; Olson et al., 2009; Virgilio et al., 2006), as estuaries represent spatially discrete habitats that tend to restrict gene flow and lead to different levels of isolation (Bilton et al., 2002).

In the present study, the phylogenetic tree, haplotype network, and AMOVA results indicated that the eight geographical locations of *T. heterochaetus* along the coasts of southeast China and north Vietnam could be characterized into three distinguishable lineages (i.e., A, B, C; Figures 2–3), corresponding to different geographic regions (A: East China Sea; B: Eastern Leizhou Peninsula; C: Western Leizhou Peninsula). Pairwise genetic distances among the 85 haplotypes ranged from 0.001 to 0.067, some of which were higher than 0.02, which is a commonly used standard for species identification (Avice, 2000). Hebert et al. (2003) suggested that higher intraspecific divergences ordinarily occur as geographic isolates, reflecting gene pool fragmentation in the origin of species in past episodes. In the last glacial maximum of the late-Pleistocene, the mean sea level was about 120 m lower than the current level (Fairbanks, 1989). During this period, the East China Sea shrank to a long, narrow ocean trough (Okinawa trough) and was isolated with the South China Sea (Koizumi et al., 2006). In addition, Hainan Island and the Chinese mainland were connected by the Qiongzhou Strait land-bridge (Voris, 2000). These historical geographic barriers may have blocked gene flow among different coastal regions and led to the deep lineages observed in the current *T. heterochaetus* populations.

Population genetic analysis can provide guidelines for strategies in species conservation and germplasm resource management (Loeschcke et al., 2013; Whiteley et al., 2006). Different lineages can show varying degrees of physiological adaptations (Méndez et al., 2001). Thus, we suggest that the genetically divergent lineages confirmed in this study should be treated as independent management units regarding conservation issues. Based on the relatively low level of genetic diversity in samples from ZS and QZ, special attention

should be paid to these two populations. As a typical *r*-selected species, *T. heterochaetus* is characterized by its short life span (one year) and hypersensitivity to environmental deterioration, habitat fragmentation, and overfishing (Tuan, 2018). To effectively protect the local populations of *T. heterochaetus*, water pollution control, overfishing reduction, and habitat preservation are all high priorities for future conservation.

Previous studies have suggested that the genetic structure of marine invertebrates can be shaped by various factors, such as currents, geographic segregation, life history characteristics, human-mediated transfer, and local selection (Selkoe & Toonen, 2011; Simon & Sato-Okoshi, 2015; Zakas & Wares, 2012; Zardi et al., 2007). Therefore, further research is still needed to identify the underlying mechanisms that may contribute to the substantial genetic heterogeneity in *T. heterochaetus*.

#### Historical demography

Both mismatch distribution analysis and significant negative values of neutrality tests (Tajima's *D* and Fu' s *F*s) indicate a pattern of recent population expansion in *T. heterochaetus*. This hypothesis was also supported by the star-shaped haplotype network (Figure 3), a characteristic of exponential population growth (Slatkin & Hudson, 1991). The demographic history was reflected in the genetic indices of this species, which showed low nucleotide diversity ( $\pi=0.032$  15) but high haplotype diversity ( $Hd=0.926$ ) (Table 2). As described by Grant & Bowen (1998), high *Hd* and low  $\pi$  can be attributed to rapid population expansion that enhances the retention of new mutations, which is consistent with the large number of unique and low-frequency haplotypes found within lineages in the present study (Figure 3).

Due to the lack of fossil and geological records, which are major obstacles for phylogeographic analysis of marine invertebrates (Provan & Bennett, 2008), the precise population expansion time based on a species-specific molecular clock for *T. heterochaetus* is not available. Based on a mutation rate of 2% per million years, our rough estimation suggests population expansion for *T. heterochaetus* in the middle to late Pleistocene, a period dominated by glaciation cycles (Imbrie et al., 1992) and periodic climatic oscillations, which may have impacted the distribution of *T. heterochaetus*. A similar demographic history pattern has also been found in other marine invertebrates along the Chinese coast (Gao et al., 2018; Liu et al., 2012).

#### CONCLUSIONS

Based on partial sequences of the *COI* gene, this study provides preliminary information on the genetic status of *T. heterochaetus* collected from the coasts of southeast China and north Vietnam. The genetic diversity in this species was highly variable and we identified six populations with significant genetic differences. These six populations could be divided into three genetically divergent lineages, corresponding to three geographic regions. We suggest that

the lineages confirmed in this study should be managed separately from a conservation point of view. For future studies, the application of multiple markers with higher resolution (e. g., microsatellites and single-nucleotide polymorphisms) and greater spatial sampling, as well as a broader understanding of biological and ecological factors, should provide a more detailed assessment of the *T. heterochaetus* population structure.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

Z.N.M. and S.Y. conceived and designed the experiment; X.H.C. and S.Y. performed the experiments; X.H.C. and Z.N.M. wrote the manuscript; W.Y., Y.Y.S., and R.W.X. contributed reagents/materials/analysis tools; B.F. and L. W. collected the samples and made English corrections. All authors read and approved the final version of the manuscript.

## ACKNOWLEDGEMENTS

We thank Bo Li from Yangjiang Polytechnic for his assistance in field investigation.

## REFERENCES

- Avise JC. 2000. Phylogeography: the History and Formation of Species. Cambridge, MA: Harvard University Press.
- Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**(1): 37–48.
- Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**(1): 38–44.
- Bilton DT, Paula J, Bishop JDD. 2002. Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal and Shelf Science*, **55**(6): 937–952.
- Birky CW, Fuerst P, Maruyama T. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**(3): 613–627.
- Blank M, Bastrop R. 2009. Phylogeny of the mud worm genus *Marenzelleria* (Polychaeta, Spionidae) inferred from mitochondrial DNA sequences. *Zoologica Scripta*, **38**(3): 313–321.
- Chen X, Li M, Liu H, Li B, Guo L, Meng Z, Lin H. 2016. Mitochondrial genome of the polychaete *Tylorrhynchus heterochaetus* (Phyllodocida, Nereididae). *Mitochondrial DNA Part A*, **27**(5): 3372–3373.
- Costa PFE, Gil J, Passos A M, Pereira P, Melo P, Batista F, Da Fonseca LC. 2006. The market features of imported non-indigenous polychaetes in Portugal and consequent ecological concerns. *Scientia Marina*, **70**(S3): 287–292.
- Cowen RK, Sponaugle S. 2009. Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, **1**: 443–466.
- Darling JA, Reitzel AM, Finnerty JR. 2004. Regional population structure of a widely introduced estuarine invertebrate: *Nematostella vectensis* Stephenson in New England. *Molecular Ecology*, **13**(10): 2969–2981.
- Dean HK. 2008. The use of polychaetes (Annelida) as indicator species of

marine pollution: a review. *Revista de Biología Tropical*, **56**(4): 11–38.

- Duan XH, Liu WS, Xu AY, Xu CH, He QT, Xiao RY, Pan G. 2017. Observation on embryonic development of *Tylorrhynchus heterochaeta* and effect of salinity on their hatching. *South China Fisheries Science*, **13**(4): 115–121.
- Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**: 47–50.
- Fairbanks RG. 1989. A 17, 000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deepocean circulation. *Nature*, **342**(6250): 637–642.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**(5): 294–299.
- Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**(2): 915–925.
- Gao X, Mu D, Lou J, Hou C, Zhu J, Wang J, Jin S. 2018. Genetic diversity and structure analysis of *Phascolosoma esculenta* in the coastal zone of south-eastern China based on mitochondrial Cyt b gene. *Turkish Journal of Fisheries and Aquatic Sciences*, **18**(4): 519–528.
- Glasby CJ, Timm T. 2008. Global diversity of polychaetes (Polychaeta; Annelida) in freshwater. *Hydrobiologia*, **595**: 107–115.
- Grant WAS, Bowen BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**(5): 415–426.
- Green BN, Suzuki T, Gotoh T, Kuchumov AR, Vinogradov SN. 1995. Electrospray ionization mass spectrometric determination of the complete polypeptide chain composition of *Tylorrhynchus heterochaetus* hemoglobin. *Journal of Biological Chemistry*, **270**(31): 18209–18211.
- Harpending H. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**(4): 591–600.
- Hebert PDN, Ratnasingham S, de Waard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**(S1): S96–S99.
- Imbrie J, Boyle EA, Clemens SC, Duffy A, Howard WR, Kukla G, Kutzbach J, Martinson DG, McIntyre A, Mix AC, Molino B, Morley JJ, Peterson LC, Pisias NG, Prell WL, Raymo ME, Shackleton NJ, Toggweiler JR. 1992. On the structure and origin of major glaciation cycles 1. Linear responses to Milankovitch forcing. *Paleoceanography and Paleoceanography*, **7**(6): 701–738.
- Jolly MT, Viard F, Gentil F, Thiébaud É, Jollivet D. 2006. Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology*, **15**(7): 1841–1855.
- Kaoawa Y. 1954. Histological Observations on the transformation to mature worms of *Tylorrhynchus heterochaetus*. *Journal of Gakugei, Tokushima University*, **4**: 65–71.
- Kesäniemi JE, Rawson PD, Lindsay SM, Knott KE. 2012. Phylogenetic analysis of cryptic speciation in the polychaete *Pygospio elegans*. *Ecology and Evolution*, **2**(5): 994–1007.

- Koizumi I, Tada R, Narita H, Irino T, Aramaki T, Oba T, Yamamoto H. 2006. Paleocceanographic history around the Tsugaru Strait between the Japan Sea and the Northwest Pacific Ocean since 30 cal kyr BP. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **232**(1): 36–52.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870–1874.
- Kyle CJ, Boulding EG. 2000. Comparative population genetic structure of marine gastropods (*Littorina spp.*) with and without pelagic larval dispersal. *Marine Biology*, **137**(5-6): 835–845.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**(11): 1451–1452.
- Liu F, Guo QS, Shi HZ, Lv F, Yu YB, Lv LL, Huang JT, Wang AM, Liang HX. 2014. Genetic diversity analysis of *Perinereis aibuhitensis* based on ISSR and SRAP markers of Chinese coast populations. *Biochemical Systematics and Ecology*, **57**: 262–269.
- Liu H, Liu M, Ge S, Wang Q, Yu D, Guan S. 2012. Population structuring and historical demography of a common clam worm *Perinereis aibuhitensis*, near the coasts of Shandong Peninsula. *Biochemical Systematics and Ecology*, **44**(4): 70–78.
- Loeschcke V, Tomiuk J, Jain SK. 2013. Conservation genetics (Vol. 68). Basel: Birkhäuser, 37–53.
- Méndez N, Linke-Gamenick I, Forbes VE, Baird DJ. 2001. Sediment processing in *Capitella spp.* (Polychaeta: Capitellidae): strain-specific differences and effects of the organic toxicant fluoranthene. *Marine Biology*, **138**(2): 311–319.
- Okada K. 1952. Experimental studies on the Japanese palolo, *Tylorrhynchus heterochaetus*. I. activation of egg-cells. *Journal of Gakugei, Tokushima University*, **2**: 43–52.
- Olson MA, Zajac RN, Russello MA. 2009. Estuarine-scale genetic variation in the polychaete *Hobsonia florida* (Ampharetidae; Annelida) in Long Island Sound and relationships to Pleistocene glaciations. *The Biological Bulletin*, **217**(1): 86–94.
- Pinsky ML, Palumbi SR. 2014. Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology*, **23**(1): 29–39.
- Plouviez S, Shank TM, Faure B, Daguin-Thiebaut C, Viard F, Lallier FH, Jollivet D. 2009. Comparative phylogeography among hydrothermal vent species along the East Pacific Rise reveals vicariant processes and population expansion in the South. *Molecular Ecology*, **18**(18): 3903–3917.
- Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**(10): 564–571.
- Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**(3): 552–569.
- Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**(4): 406–425.
- Sato M, Osanai K. 1990. Sperm attachment and acrosome reaction on the egg surface of the polychaete, *Tylorrhynchus heterochaetus*. *The Biological Bulletin*, **178**(2): 101–110.
- Selkoe KS, Toonen RJ. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, **436**: 291–305.
- Simon CA, Sato-Okoshi W. 2015. Polydroid polychaetes on farmed molluscs: distribution, spread and factors contributing to their success. *Aquaculture Environment Interactions*, **7**(2): 147–166.
- Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**(2): 555–562.
- Smith LM, Hutchings P, Fraser CI. 2015. Molecular evidence supports coastal dispersal among estuaries for two benthic marine worm (Nephtyidae) species in southeastern Australia. *Marine Biology*, **162**(6): 1319–1327.
- Suzuki T, Gotoh T. 1986. The complete amino acid sequence of giant multisubunit hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. *Journal of Biological Chemistry*, **261**(20): 9257–9267.
- Suzuki T, Takagi T, Gotoh T. 1990. Primary structure of two linker chains of the extracellular hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. *Journal of Biological Chemistry*, **265**(21): 12168–12177.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**(3): 585–595.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**(24): 4876–4882.
- Tuan NN. 2018. Biological characteristics and effects of salinity on reproductive activities of marine worm (*Tylorrhynchus heterochaetus*, Quatefages, 1865) in summer season in Hai Phong - Viet Nam. *SNRU Journal of Science and Technology*, **10**(1): 25–31.
- Virgilio M, Backeljau T, Abbiati M. 2006. Mitochondrial DNA and allozyme patterns of *Hediste diversicolor* (Polychaeta: Nereididae): the importance of small scale genetic structuring. *Marine Ecology Progress Series*, **326**: 157–165.
- Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**(5): 1153–1167.
- Weersing K, Toonen RJ. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**: 1–12.
- Whiteley AR, Spruell P, Allendorf FW. 2006. Can common species provide valuable information for conservation?. *Molecular Ecology*, **15**(10): 2767–2786.
- Zakas C, Wares JP. 2012. Consequences of a poecilogonous life history for genetic structure in coastal populations of the polychaete *Streblospio benedicti*. *Molecular Ecology*, **21**(22): 5447–5460.
- Zardi GI, McQuaid CD, Teske PR, Teske PR, Barker NP. 2007. Unexpected genetic structure of mussel populations in South Africa: indigenous *Perna perna* and invasive *Mytilus galloprovincialis*. *Marine Ecology Progress Series*, **337**: 135–144.