African Journal of Biotechnology Vol. 11(16), pp. 3738-3750, 23 February, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.1643
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Population genetic structure and phylogeography of the mud-flat crabs *Helice tientsinensis* and *Helice latimera* along the coast of China seas based on mitochondrial DNA

Jing Ming Xu[1], FanYang[1] and Shi Chun Sun[2]*

[1]College of Life Science and Technology, Chongqing University of Arts and Sciences, Chongqing, 402168, People's Republic of China.
[2]The Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao, 266003, People's Republic of China.

Accepted 27 January, 2012

Genetic structure and evolutionary demography of the mud-flat crabs (*Helice tientsinensis* and *H. latimera*) were investigated using sequence data of mitochondrial cytochrome *c* oxidase I (COI) gene. Samples of 213 individuals were collected from nine localities over most of the species' range along the coast of China seas. There were 38 variable sites defining 42 haplotypes. Two distinct lineages were detected, which might be divergent in different marginal seas of the Northwestern Pacific during periods of the low sea level. There was a significant difference in geographical distribution frequency between the two lineages. Analysis of molecular variance (AMOVA) and conventional population statistics (F_{ST}) revealed significant genetic structure in the 10 populations of *H. tientsinensis* and *H. latimera*. For all populations, the AMOVA showed three different groups in the marginal seas of East Asia. Within-lineage variation was also highly structured in the two lineages. Both mismatch distribution and neutrality tests indicate population expansion in the two lineages. The two co-existing lineages in the East China Sea might be explained by population reestablishment after the last glacial maximum.

Key words: *Helice tientsinensis*, *Helice latimera*, mtDNA, cytochrome *c* oxidase subunit I gene, population genetic structure, population expansion.

INTRODUCTION

A unique topographic feature of the northwestern Pacific is the series of linked marginal seas, including the Japan Sea, Yellow Sea, East China Sea, and South China Sea. The lowering and rising of sea levels during Pliocene and Pleistocene glacial-interglacial cycles greatly affected the areas and configurations of these marginal seas (Wang, 1999). A few studies in the northwestern Pacific have reported that the isolation of populations in marginal seas during lower sea levels in the Pliocene and Pleistocene, followed by range expansions and secondary contacts, were important in speciation and in shaping the spatial population genetic structure of near-shore fishes (Liu et al., 2006, 2007) and a crab with a catadromous lifecycle (Wang et al., 2008). *Helice latimera* and *Helice tientsinensis*, belonging to the genus *Helice*, are common crabs which burrow in intertidal mudflats and estuaries, especially in the high intertidal and supralittoral zones along the coast of China seas (Dai and Yang, 1991). The identification of the two species mostly relied on the size,

*Corresponding author. E-mail: sunsc@ouc.edu.cn. Tel/Fax: +86 532 8203 2216.

Abbreviations: COI, Cytochrome *c* oxidase I; F_{ST}, conventional population statistics; **AMOVA,** Analysis of molecular variance.

shape and number of infraorbital crest. However, the key morphological differences are variable within species and greatly differ between sexes, and the females of both species could not be distinguished by the shape and number of infraorbital crest.

The molecular identification of species is clear-cut and efficient irrespective of phenotypic variation, except in cases of incomplete lineage sorting among recently separated species (Nielsen and Matz, 2006). This method has another strong point in that it could reveal cryptic species that are hard to discriminate by traditional approaches (Hebert et al., 2004). If the molecular data reveal adequate intra-specific variation, population genetic analyses can be used to infer the evolutionary pathways of the species concerned, such as their population structure, demographic history and phylogeography. Shih and Suzuki (2008) used the large subunit (16S) ribosomal (r) RNA and cytochrome oxidase subunit I (COI) to study the taxonomy, phylogeny, and biogeography of the endemic genera *Helice*, *Helicana* and *Chasmagnathus* from East Asia.

Their results showed that three species of the genus *Helice*, *H. latimera*, *H. formosensis* and *H. tientsinensis*, formed an unresolved "*H. latimera* clade", implying that they may represent the same species with an intraspecific variable number of suborbital tubercles. However, Sun et al. (2009) used three different genes (16S rDNA-trnL1-IGS-nad1) to study the taxonomy of the *Helice* sensu lato crabs.

The results provide independent molecular evidence supporting the splitting of the Asian species into *Helice* and *Helicana*. In their results, the two *Helice* species, *H. latimera* and *H. tientsinensis*, are valid species. Yin et al. (2009) also believed that these three species are valid species based on the data from three mitochondrial and two nuclear genes, and that some straits (Tokara Strait, Tsushima/Korea Strait, and Taiwan Strait), the Okinawa Trough and several currents (Kuroshio Current, Tsushima Current, and Taiwan Strait Warm Current) have acted as geographic barriers resulting in allopatric speciation among onshore marine animals in the northwestern Pacific.

In this study, we analyzed 10 populations of *H. latimera* and *H. tientsinensis* along the coast of China seas using the sequence data of the partial COI gene so as: (1) to examine sequence variability and geographic structure of two species, and particularly to examine whether the two species distributed along the coast of China seas have diverged genetically or represents the same species with an intraspecific variable number of suborbital tubercles; (2) to infer historical population processes (e.g., population fragmentation, range expansion or long distance colonization), or present-day processes (e.g., restricted gene flow) that might have affected the current distribution of the two species, in order to improve the knowledge of the origin and evolution of these two species.

MATERIALS AND METHODS

Sample collection

A total of 213 individuals of *H. latimera* and *H. tientsinensis* from nine localities were collected over its range during 2004 to 2006 (Table 1 and Figure 1). All specimens were either preserved in 95% ethanol or packed in ice after collection and deposited in the Key Laboratory of Mariculture, Ministry of Education, Ocean University of China.

DNA extraction, amplification and sequencing

Genomic DNA was extracted from the muscle tissue of walking legs or claws using a phenol-chloroform method. The COI gene fragments were amplified using primers COIL1490: 5′-GGTCAAC-AAATCATAAAGATATTGG-3′ and COIH2198: 5′-TAAACTTCAG-GGTGACCAAAAAATCA-3′ (Folmer et al., 1994). Polymerase chain reaction (PCR) amplification was carried out in an Eppendorf authorized thermal cycler. Reactions were conducted in 25 µL volumes containing 1 U *Taq* DNA polymerase (Takara, China), 0.1 mM primers, 2.0 mM $MgCl_2$, 0.1 mM dNTPs, 2.5 µL 10X PCR buffer and approximately 30 ng of DNA as template. The cycling conditions were as follows: initial 2 min at 94°C, followed by 35 cycles of 45 s at 94°C, 50 s at 47°C, and 1 min at 72°C, with a final extension for 7 min at 72°C. The purified products were used as the template DNA for cycle sequencing reactions performed using BigDye Terminator Cycle Sequencing Kit (ver. 2.0, PE Biosystems, Foster City, California), and sequencing was conducted on an ABI Prism 3730 (Applied Biosystems) automatic sequencer with both forward and reverse primers. The primers used for sequencing were the same as those for PCR amplification. COI sequences have been deposited in the GenBank database under Accession numbers HQ121442-HQ121487.

Data analysis

Sequences were edited and aligned using DNASTAR software (DNASTAR Inc., Madison, USA) and individual consensus sequences were retrieved with both alignment and manual check. The accuracy of the COI sequencing was confirmed by translating the nucleotide data to amino acid sequences. Molecular diversity indices such as number of haplotypes, polymorphic sites, transitions, transversions, and indels, were obtained using Arlequin (Ver. 2.0, Schneider et al., 2000). Haplotype diversity (h), nucleotide diversity (π), and their corresponding variances were calculated following Nei's (1987) as implemented in Arlequin. Implemented with Modeltest 3.06 (Posada and Crandall, 1998), hierarchical series of likelihood ratio tests (Huelsenbeck and Rannala, 1997) were used to identify the appropriate nucleotide substitution models. The net average genetic distance between lineages given by $dA = dXY – (dX + dY)/2$ where, dXY is the average distance between lineages X and Y, and dX and dY are the mean within-lineages distances, was calculated with MEGA 4.0 (Tamura et al., 2007).

Genetic distances were generated for phylogenetic reconstruction with MEGA4.0 using the model of HKY85 (Hasegawa et al., 1985) with no invariable sites and equal mutation rate given by the Modeltest. The neighbor-joining (NJ) tree of the haplotype was constructed using MEGA (Saitou and Nei, 1987) and evaluated with 1000 bootstrap replicates. A bootstrap analysis with 1,000 replicates was used to evaluate the reliability of the phylogenetic analysis (Felsenstein, 1985). In addition, genealogical relationships were examined by constructing haplotype networks using a median

Table 1. Distribution of haplotypes of COI among the populations of *H. tientsinensis* and *H. latimera*.

Population	Haplotype																					Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
HKL	19							1														20
BHL	18	1	1																			20
QZL				15					1		1							1	1	1	1	21
NBL	2			6					11	1	1	1										22
NB T	3			2					11				1	1								18
RZ T	3			3	1				4		1											12
WH T	2			9	1	1			1													13
TG T	5			12																		17
PJ T	1			14	3	1	1								2	1	1					24
DD T	3			15	2	1	1															22
Total	56	1	1	76	7	3	2	1	28	1	3	1	1	1	2	1	1	1	1	1	1	

Population	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	
HKL																						0
BHL																						0
QZL	1	1																				2
NBL																						0
NB T																						0
RZT			1	1	1	1	1	1	1	1												8
WH T														3	1	1	1	1	1	1	1	11
TG T											1	1	1									3
PJ T																						0
DD T																						0
Total	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1	1	1	1	1	1	213

"L" means four *H. latimera* populations, and "T" means six *H. tientsinensis* populations.

-network approach (Bandelt et al., 1995, 2000). Furthermore,an analysis of molecular variation (AMOVA, Excoffier et al., 1992) was used to statistically test the population structure. AMOVA and a bootstrap analysis with 5,000 replicates were performed in Arlequin. For the COI data, the appropriate model of nucleotide substitution was the HKY85 model (Hasegawa et al., 1985) with no invariable sites and equal mutation rate. Due to the fact that the HKY model was not available in Arlequin, the more inclusive Tamura-Nei (TrN) (Tamura and Nei, 1993) model was used to calculate the genetic pair-wise distances between haplotypes. The conventional population (sequence) statistics (F_{ST}) were calculated between pairs of populations. The significance (5% level) of the F_{ST} was tested by 1,000 permutations for each pair-wise comparison.

Both mismatch analysis and neutrality tests were performed in Arlequin. The historical demographic expansions were examined by two different approaches. First, the Tajima's *D* test (Tajima, 1989) and the Fu's *Fs* test (Fu, 1997) were used to test whether neutrality holds. Significant negative *D* and *Fs* statistics can be interpreted as signatures of population expansion. Historic demographic expansions were also investigated by examination of frequency distributions of pair-wise differences between sequences (mismatch distribution), which was based on three parameters: θ_0, θ_1 (before and after the population growth) and τ (time since expansion expressed in units of mutational time) (Rogers and Harpending, 1992; Rogers, 1995). The mismatch distribution is usually multimodal in samples drawn from populations at demographic equilibrium, but it is usually unimodal in populations following a recent population demographic expansion and range expansion (Rogers and Harpending, 1992; Slatkin and Hudson, 1991; Ray et al., 2003; Excoffier, 2004). The parameters of the demographic expansion τ, θ_0 and θ_1 are estimated by a generalized non-linear least-square approach, and confidence intervals of the parameters are computed using a parametric bootstrap approach (Schneider and Excoffier, 1999). The values of τ were transformed to estimate the real time since expansion with the Equation: $\tau = 2ut$, where u is the mutation rate per generation for the whole sequence under study and t is the time measured in years since expansion.

Molecular clock calibrations based on geminate species separated by the Isthmus of Panama, estimate the average pair-wise sequence divergence rate at about 2.3% per million years for the COI gene in crustaceans and other arthropods (Knowlton et al., 1993; Brower, 1994; Schubart et al., 1998; Daniels et al., 2002). Unless stated otherwise, we used this rate for our analysis (that is-the estimated single-lineage value for μ was 1.15×10^{-8}).

RESULTS

Genetic variation

A 685 bp fragment of mitochondrial COI gene was

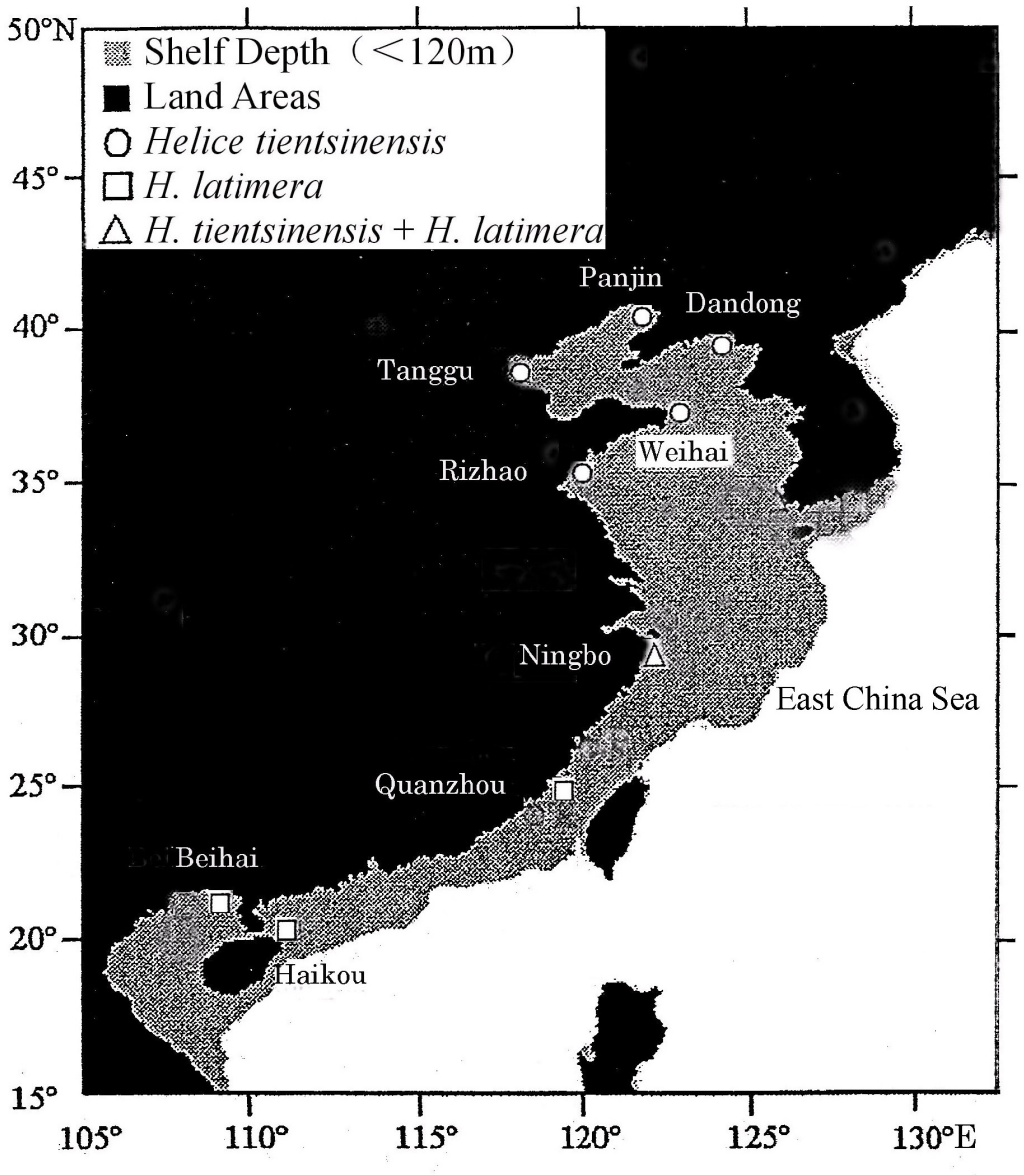

Figure 1. Map showing sample locations of *H. tientsinensis* and *H. latimera*. Shaded sea areas are continental shelves that would have been exposed to the air during periods of low sea-level.

polymorphic sites (15 were parsimony informative) with 36 transitions and 3 transversions (Table 7). 42 haplotypes were defined from all individuals sequenced (designated as Hap1-42). The number of detected haplotypes within samples ranged from two (HKL) to 13 (RZT, WHT). The majority of haplotypes (33/42, 78.6%) were singletons (haplotypes represented by a single sequence in the sample). Of the remaining nine haplotypes, seven were shared among populations and two were found in more than one individual, but only in one population. Three haplotypes (7.1%) were common and found in 75% (160/213) of the individuals studied. Four haplotypes were shared among species. Haplotype frequencies of the COI fragment and their distributions in the 10 populations are shown in Table 1. Sequence differences of the 42 haplotypes ranged from 0.15 to 1.23%.

Genetic relationships among haplotypes

Phylogenetic analysis of haplotypes revealed two distinct lineages (lineage A and B) in both the NJ tree and the haplotype network (Figures 2 and 3). Net average genetic distance between the two lineages was 0.46%. Applying the sequence divergence rate (2.3% per million years) in

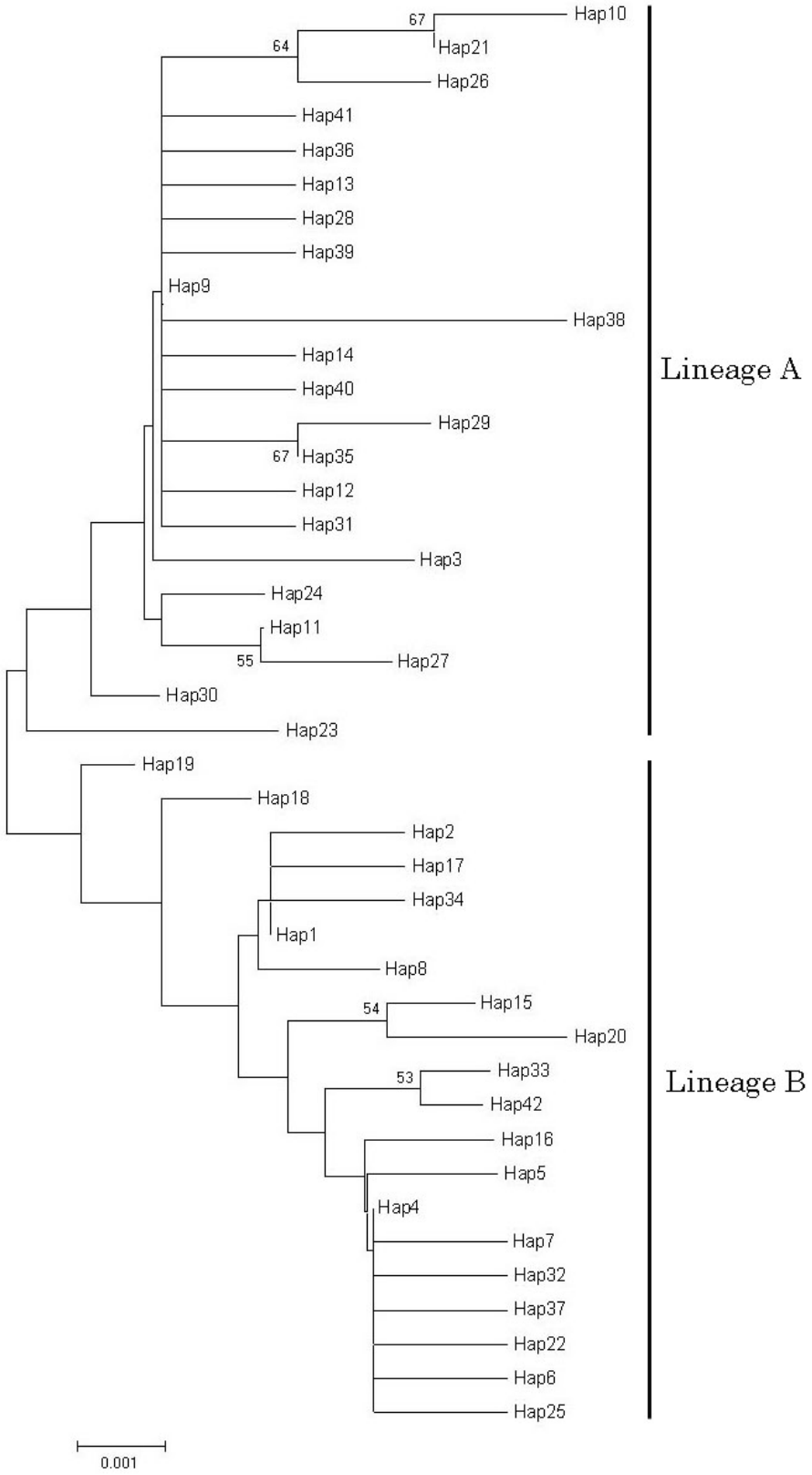

Figure 2. Unrooted neighbor-joining tree for 42 haplotypes of *H. tientsinensis* and *H. latimera*. Bootstrap supports of >50% in 1000 replicates are shown.

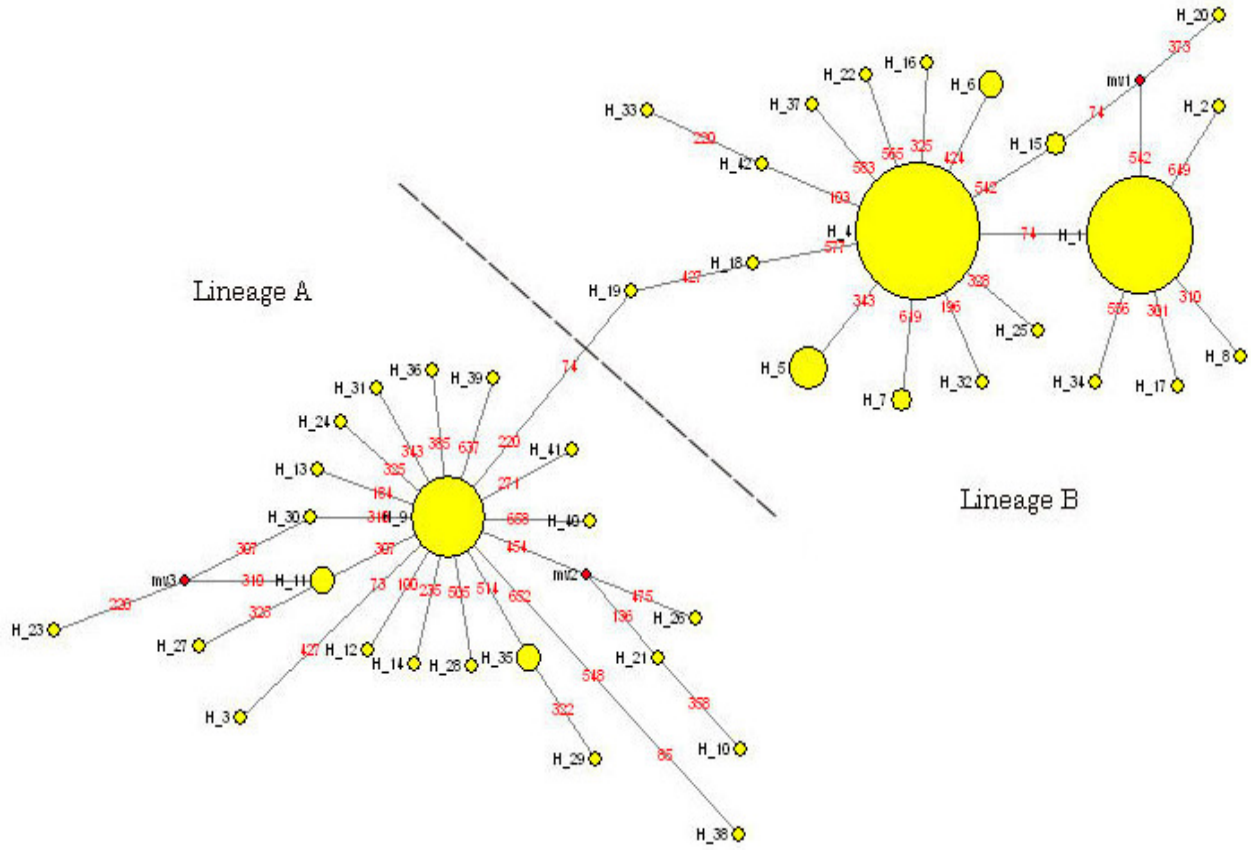

Figure 3. Median networks showing genetic relationship among COI gene haplotypes in lineages A and B. The sizes of circles are proportional to haplotype frequency. Haplotypes are marked by names that correspond to Table 1. Number marks on lines joining haplotypes represent the position and number of nucleotide substitutions. COI, cytochrome c oxidase

the COI gene (Schubart et al., 1998), the divergence of lineage A and B occurred about 200,000 years ago. Lineage A contained 53 individuals, with a haplotype diversity of 0.721 ± 0.069. This lineage had 24 polymorphic sites and the nucleotide diversity was 0.0020 ± 0.0014. Moreover, lineage B included 160 individuals, with a haplotype diversity of 0.653 ± 0.026 and a nucleotide diversity of 0.0013 ± 0.0010. The average pair-wise distances between individuals, within the lineages (±SE) were 0.43 ± 0.09% and 0.38 ± 0.1%, respectively.

There were obviously geographical differences in haplotype frequencies of the two lineages (Tables 1 and 2). Lineage A dominates the East China Sea populations, but the frequency declined steadily along the coast from Ningbo to north and south, respectively. More also, lineage A was not found in the South China Sea population Haikou and the north Yellow and Bohai Sea populations (Dandong, Tanggu, Panjin). On the other hand, lineage B has a very wide geographical distribution and is present in all ten populations from Dandong to Haikou. The frequency of lineage B showed the opposite trend, increased gradually along the coast from Ningbo to north and south, respectively. Furthermore, the network of lineage A was star-shaped with a dominant haplotype (53%) shared by six populations (BHL, QZL, NBL, NBT, RZT, WHT). A larger number of haplotypes and private haplotypes were found in northern populations than in southern ones of the East China Sea group (Tables 1 and 2). There was only one haplotype of lineage A that was found in the BHL population. Meanwhile, in lineage B, the network represented a “double star” shape, where most haplotypes were very closely related to the two common haplotypes (Table 2; Figure 3). The two common haplotypes were separated by a single substitution. Haplotype 1 was the dominant haplotype of the South China Sea shared by nine populations except for QZL. The proportion of haplotype 1 in these nine samples ranged from 4 to 95%. Haplotype 4 was the dominant haplotype of the Yellow and Bohai Sea populations shared by eight populations except for BHL and HKL. The proportion of haplotype 4 in these eight samples ranged from 11 to 68%.This star-like network is a signature which suggests population expansion.

Table 2. Number and proportion of individuals and number of haplotypes for phylogenetic lineages A and B in different populations.

ID of population	Species	Sample size	Lineage A		Lineage B	
			Number of individuals (proportion, %)	Number of haplotype (private)	Number of individual (proportion,%)	Number of haplotype (private)
HKL	*H. latimera*	20	0	0	20 (100.0)	2 (1)
BHL		20	1 (5.0)	1 (1)	19 (95.0)	2 (1)
QZL		23	4 (17.4)	4 (2)	19 (82.6)	5 (4)
NBL		22	14 (63.6)	4 (2)	8 (36.4)	2 (0)
NBT	*H. tientsinensis*	18	13 (72.2)	3 (2)	5 (27.8)	2 (0)
RZT		20	12 (60.0)	9 (7)	8 (40.00)	4 (1)
WHT		24	9 (37.5)	7 (6)	15 (62.5)	6 (2)
TGT		20	0	0	20 (100.0)	5 (3)
PJT		24	0	0	24 (100.0)	8 (3)
DDT		22	0	0	22 (100.0)	5 (0)
		213	53 (24.9%)	22	160 (75.1%)	20

Population genetic structure

Significant separation of the two lineages was also supported by AMOVA, with 77.09% of all variance being partitioned between the two lineages (F_{ST} = 0.771, P = 0.00) (Table 6). The AMOVA analysis showed that 30.31% of the genetic diversity was found among groups (P = 0.00). A smaller (8.23%) but significant (P = 0.00) amount of genetic diversity was found among populations within groups, while 61.46% of the divergence was found within populations. The genetic structures of the populations within the two lineages were also investigated by AMOVA (Table 6). In lineage A, 10.40% (P=0.006) of the genetic variation was found within populations, whereas 89.60% of the variation was between populations, indicating existence of little genetic differentiation in lineage A. Moreover, AMOVA analyses in lineage B showed that the genetic variation among groups was 33.09% (P=0.00), while among populations within groups it was not significant (3.8%, P = 0.063), indicating a significant population structure in lineage B. Most of the pair-wise F_{ST} values were high and significant after sequential Bonferroni correction, with the exception of nine comparisons between populations mostly from the East China Sea (Table 3).

Furthermore, pair-wise population F_{ST} values were calculated for each lineage. For lineage A, the genetic differences between the five East China Sea populations and the South China Sea population (BHL) were strong (F_{ST} = 0.239 to 0.857), but not statistically significant after the sequential Bonferroni correction (P = 0.99). The genetic differences between Weihai and the other four East China Sea populations were small (F_{ST} = 0.061 to 0.118) but statistically significant (P < 0.05) (Table 4). There were also small significant values of F_{ST} between NBT and other three populations (QZL\RZT\WHT, F_{ST}=0.039 to 0.330) (Table 4). For lineage B, the genetic differences between the two South China Sea populations (HKL and BHL) and the other populations were strong (F_{ST} = 0.414 to 0.744) and statistically significant after the sequential Bonferroni correction (P < 0.05). There were also small significant values of F_{ST} between NBT and the other two populations (QZL\PJT, F_{ST} = 0.189 to 0.249) (Table 5).

Historical demography

The mismatch distributions for lineages A and B were both unimodal, supporting a model of sudden expansion (Figure 4a and b). The P_{SSD} and raggedness tests could not reject the expansion hypothesis. The tau value (τ), which reflects the location of the mismatch distribution crest, provided a rough estimation of the time when rapid population expansion started. For lineage A, it was estimated that population expansion occurred at

Table 3. Pair-wise F_{ST} (below diagonal) and associated *P* values (above diagonal) among *H. tientsinensis* and *H. latimera* populations.

Population	BHL	HKL	NBL	QZL	DDT	NBT	PJT	RZT	TGT	WHT
BHL		0.76616	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00040
HKL	0.00003		0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00010
NBL	0.44109	0.49431		0.00158	0.00000	0.44213	0.00000	0.86378	0.00010	0.05772
QZL	0.34402	0.38177	0.25602		0.02802	0.00010	0.01020	0.00208	0.06257	0.11692
DDT	0.59535	0.67178	0.51027	0.07673		0.00000	0.91060	0.00000	0.16771	0.00198
NBT	0.57843	0.64215	-0.01933	0.38797	0.65724		0.00000	0.33492	0.00000	0.01267
PJT	0.56168	0.61664	0.50588	0.08288	-0.02491	0.64098		0.00000	0.03960	0.00099
RZT	0.35502	0.40140	-0.03193	0.21034	0.43592	-0.00677	0.43594		0.00000	0.08633
TGT	0.43278	0.50407	0.44131	0.05453	0.02058	0.58945	0.05161	0.36961		0.00644
WHT	0.23866	0.27498	0.06854	0.03665	0.19928	0.15517	0.20956	0.04574	0.14472	

approximately 91,000 years ago, while for the lineage B this time was 65,000 years ago based on the mutation rate of 1.15%/MY for COI gene and the equation $\tau = 2ut$. The Fu's F_S and Tajima's *D* tests for lineages A and B were negative and significant (Table 7), which also indicates population expansion. These were consistent with the star shape networks.

DISCUSSION

Relationship between the two species

Since the females of *H. latimera* and *H. tientsinensis* are not distinguishable from each other, all individuals of *H. tientsinensis* and *H. latimera* collected in this study were regarded as males. Of the 213 individuals, 85 belonged to *H. latimera* and 128 belonged to *H. tientsinensis*. Four haplotypes were shared between the two species (Table 1). The average pair-wise divergences between individuals within species (SE) were 0.28% (0.10%) and 0.36% (0.12%) for *H. latimera* and *H. tientsinensis*, respectively. However, the average distance and net average genetic distances between the two species (SE) were 0.34% (0.12%) and 0.01%, respectively, less than the distance among individuals. In addition, Φ-statistics was small and not significant (0.0307, $P = 0.209$) when samples were divided based on the two species (Table 6). The taxonomy of the genus *Helice* has not been well settled for a long time (Sakai and Yatsuzuka, 1980; Sakai et al., 2006). Dai and Song (1977) and Sakai and Yatsuzuka (1980) proposed that *H. tientsinensis* and *H. latimera* should be ranked at the species level within the genus *Helice*. Several molecular studies on this group supported this conclusion (Sun et al., 2009; Yin et al., 2009). However, a study by Shih and Suzuki (2008) using 16S rRNA and COI implied that *H. latimera*, *H. formosensis* and *H. tientsinensis* could be considered the same species with an intra-specific variation in the number of suborbital tubercles. Our data as presented here are consistent with this result.

Phylogeographic patterns

Many authors have suggested that Pleistocene glaciations have been the most significant events to shape the phylogeographic mitochondrial DNA patterns and population structure in marine species (Planes et al., 2001; Liu et al., 2006, 2007; Wang et al., 2008). In this study, two lineages were found, likely reflecting isolation of marginal seas of the East Asia during low sea level periods in the Pleistocene. The frequency distribution of lineage A indicated an origin in the East China Sea. In lineage B, there were two common haplotypes. Haplotype 1 dominates the two South China Sea locations, indicating that this haplotype was isolated and diverged in the South China Sea. Haplotype 4 was the dominant haplotype of the Yellow and Bohai Sea, implying this haplotype was isolated in the Yellow Sea.

The COI data suggests a divergence time of 200,000 years between the two lineages, in the late Pleistocene. During this period, a series of glacial events took place, causing considerable changes in water temperature and sea level. In the late Pleistocene's low sea level stage, the sea level was 130 to 150 m lower than the present level in the East China Sea. Consequently, the entire Bohai Sea and most of the Yellow Sea were exposed, and the Yellow Sea was only a narrow waterway close to the Korean side of the modern

Table 4. Pair-wise F_{ST} (below diagonal) and associated *P* values (above diagonal) of lineage A among *H. tientsinensis* and *H. latimera* populations.

Population	NBL	NBT	QZL	RZT	WHT	BHL
NBL		0.74636	0.06722	0.45322	0.00406	0.99990
NBT	-0.00216		0.00614	0.00594	0.00010	0.99990
QZL	0.16159	0.32980		0.39075	0.03069	0.99990
RZT	-0.00278	0.03929	0.00767		0.00624	0.99990
WHT	0.08265	0.10123	0.11786	0.06055		0.99990
BHL	0.69709	0.85739	0.23892	0.43953	0.33987	

Table 5. Pair-wise F_{ST} (below diagonal) and associated *P* values (above diagonal) of lineage B among *H. tientsinensis* and *H. latimera* populations.

Population	BHL	HKL	NBL	QZL	NBT	RZT	WHT	DDT	TGT	PJT
BHL		0.74399	0.00010	0.00000	0.03386	0.00059	0.00000	0.00000	0.00000	0.00000
HKL	0.00015		0.00020	0.00000	0.03386	0.00069	0.00000	0.00000	0.00000	0.00000
NBL	0.73677	0.74432		0.31670	0.29651	0.99990	0.94832	0.74151	0.99990	0.58469
QZL	0.68414	0.69015	0.00615		0.02129	0.05158	0.40758	0.17820	0.03524	0.34007
NBT	0.41446	0.42677	0.08007	0.24876		0.64825	0.08306	0.07197	0.41738	0.02831
RZT	0.51760	0.52814	-0.06519	0.09468	-0.07421		0.34967	0.29393	0.70934	0.18018
WHT	0.65113	0.65828	-0.05100	0.00232	0.15607	0.00892		0.94624	0.37313	0.78913
DDT	0.66606	0.67178	-0.04151	0.01387	0.20353	0.02372	-0.03977		0.17018	0.91535
TGT	0.49696	0.50407	-0.07582	0.06324	-0.00004	-0.03636	0.00164	0.02058		0.03564
PJT	0.61070	0.61664	-0.00956	0.00677	0.18874	0.03996	-0.02192	-0.02491	0.05161	

Yellow Sea depression (Chen, 1991). A large land bridge, which was formed in the late Pleistocene, extended from Eastern China to Taiwan, Ryukyu and probably to the main islands of Japan. This is likely to have isolated the East China Sea from Pacific Ocean and the South China Sea. *H. tientsinensis* and *H. latimera* like to cave on the coast, especially near mouths of estuaries and mudflats or muddy shores, and they are able to survive over a wide range of salinities and temperature. For this reason, it is likely that lineage A became isolated in the East China Sea, whereas lineage B became isolated in the Yellow Sea and South China Sea. The reduced diversity and distribution of haplotypes on the network of lineage B indicate that the two South China Sea populations BHL and HKL were isolated in the South China Sea and have experienced recent population expansions. These two lineages were found in the East China Sea populations, indicating secondary contact after extensive isolation in this area.

Population genetic structure and historical demography

The results of F_{ST} and AMOVA indicate significant genetic division between these 10 populations. For all populations, the AMOVA showed three different groups in marginalseas of East Asia. Within-lineage variation was also highly structured for the two lineages (Table 6). F_{ST} values between the South China Sea populations and the others, between the East China Sea and the Bohai and the North Yellow Sea populations showed significant genetic differences (Table 3), indicating that at least three distinct populations exist in these marginal seas. The nine insignificant comparisons between populations were within-group populations, mostly from the East China Sea populations. For lineage A, the genetic differences between WHT and other four East China Sea populations (QZL\RZT\NBL\NBT), between NBT and other two populations (QZL\RZT) were small but statistically significant ($P < 0.05$) (Table 4). For lineage B, the genetic differences between the three marginal sea populations were large than within marginal sea populations. These results therefore indicate a low level of dispersal in *H. tientsinensis* and *H. latimera*.

H. tientsinensis and *H. latimera* burrow in intertidal muddy flats and estuaries, especially in the high intertidal and supralittoral zones along the coast of China (Dai and Yang, 1991). Their larvae are typically found in near-shore waters and few reach waters off the coast. This biological characteristic may be responsible for the genetic differentiation in th marginal seas of the East

Table 6. The results of AMOVA for all ten populations and for phylogenetic lineages A and B in different populations.

Grouping	Variance component	Total variance (%)	Φ-Statistics	*P*
For all populations				
PJT, TGT, DDT, WHT, RZT, NBT, NBL, QZL, HKL, BHL	AG	32.06	Φ_{ST} = 0.3206	<0.0001
Based on species				
H. tientsinensis {PJT, TGT, DDT, WHT, RZT, NBT}	AG	3.07	Φ_{CT} = 0.0307	=0.2092
H. latimera {NBL, QZL, HKL, BHL}				
Based on lineages				
Lineage A	AG	77.09	Φ_{CT} = 0.7709	<0.0001
ineage B				
Based ocean division				
Group 1 {PJT, TGT, DDT}	AG	30.31	Φ_{CT} = 0.3031	<0.0001
Group 2 {WHT, RZT, NBT, NBL, QZL}			Φ_{SC} = 0.1181	<0.0001
Group 3 (HKL, BHL}				
For lineage A				
Group 1 {WHT, RZT, NBT, NBL, QZL, BHL}	AG	10.40	Φ_{ST} = 0.1040	=0.0059
For lineage B				
1 Group 1 {PJT, TGT, DDT, WHT, RZT, NBT, NBL, QZL, HKL, BHL}	AG	33.09	Φ_{ST} = 0.3309	<0.0001
2 Group 1 {PJT, TGT, DDT}	AG	37.26	Φ_{CT} = 0.3726	=0.0362
Group 2 {WHT, RZT, NBT, NBL, QZL}			Φ_{SC} = 0.0379	=0.0626
Group 3 {HKL,BHL}				

AG, among-groups component of variance; Φ_{CT}, source of variation among-groups; Φ_{SC}, source of variation among population within-groups; Φ_{ST}, source of variation within-populations.

Asia. Similar genetic breaks have also been described between East China Sea and South China Sea populations of coastal marine invertebrates and vertebrates (Li et al., 2003; Pan et al., 2005; Liu et al., 2007). These concordant patterns strongly suggest that historical geographical factors greatly influenced the evolutionary genetic structure of marine organisms in the marginal seas of the East Asia. Besides, both mismatch distribution and neutrality tests indicate population expansion for lineage A and B. Population range expansion and demographic expansion may both influence the pattern of genetic diversity for lineage A and lineage B. In this case, the occurrence of the two lineages in the East China Sea may reflect founding of new populations in previously unoccupied habitats. As sea levels began to rise, large amounts of shallow coastal habitat would have been created due to the shallow gradient of the East China Sea shelf. Although lineage A haplotypes are common in the

Table 7. Summary of molecular diversity for *H. tientsinensis* and *H. latimera*. Number of individuals (n), number of haplotype, mean number of pair-wise differences (*k*), number of polymorphic sites (*S*), haplotype diversity (*h* ± SD), nucleotide diversity (*Л*±SD) for each grouping of samples. Tajima's *D* and Fu's *Fs*, corresponding *P*-value, and mismatch distribution parameter estimates for each lineage were also indicated.

Group	n	Number of haplotype	*k*	*S*	*h*	*Л*	Tajima's *D*		Fu's *Fs*		Mismatch distribution		
							D	*P*	*Fs*	*P*	τ (95% CI)	$\theta 0$	$\theta 1$
All	213	42	2.216±1.228	38	0.787 ± 0.019	0.0034 ± 0.0021	-	-	-	-	-	-	-
Lineage A	53	22	1.332 ± 0.841	24	0.721 ± 0.069	0.0020 ± 0.0014	-2.418	0.000	-24.589	0.000	1.377 (0.271,2.494)	0.000	12.630
Lineage B	160	20	0.860 ± 0.610	19	0.653 ± 0.026	0.0013 ± 0.0010	-2.026	0.007	-19.499	0.000	0.982 (0.730,1.326)	0.000	99999.000
HKL	20	2	0.100 ± 0.178	1	0.100 ± 0.088	0.0002 ± 0.0003	-	-	-	-	-	-	-
BHL	20	3	0.402 ± 0.388	4	0.195 ± 0.115	0.0006 ± 0.0007	-	-	-	-	-	-	-
QZL	23	9	2.052 ± 1.194	11	0.585 ± 0.122	0.0031 ± 0.0020	-	-	-	-	-	-	-
NBL	22	6	2.339 ± 1.327	9	0.693 ± 0.080	0.0036 ± 0.0023	-	-	-	-	-	-	-
NBT	18	5	1.715 ± 1.049	6	0.614 ± 0.118	0.0026 ± 0.0018	-	-	-	-	-	-	-
RZT	20	13	3.204 ± 1.727	14	0.937 ± 0.035	0.0049 ± 0.0029	-	-	-	-	-	-	-
WHT	24	13	3.154 ± 1.692	16	0.855 ± 0.065	0.0048 ± 0.0029	-	-	-	-	-	-	-
TGT	20	5	0.844 ± 0.625	5	0.600 ± 0.101	0.0013 ± 0.0011	-	-	-	-	-	-	-
PJT	24	8	0.882 ± 0.640	7	0.656 ± 0.105	0.0013 ± 0.0011	-	-	-	-	-	-	-
DDT	22	5	0.603 ± 0.498	4	0.528 ± 0.118	0.0009 ± 0.0008	-	-	-	-	-	-	-

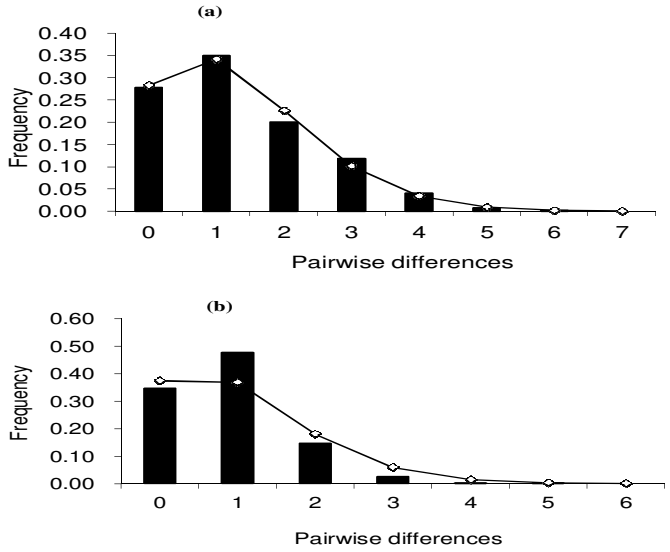

Figure 4. The observed pair-wise differences (bars) and the expected mismatch distributions under the sudden expansion model (solid line) for COI haplotypes in lineage: (a) A of *H. tientsinensis* and *H. latimera;* (b) B of *H. tientsinensis* and *H. latimera*. COI, cytochrome c oxidase I.

East China Sea populations, the frequency is very low in the South China Sea populations and the Bohai and north Yellow Sea populations, which may reflect limited effective dispersal of individuals moving into already saturated habitats.

In conclusion, our results show that *H. tientsinensis* and *H. latimera* formed an unresolved “clade”, with two lineages and some genetic differentiation at the mitochondrial DNA level. The conspecific status of *H. tientsinensis* and *H. latimera* was further supported by our published morphological analyses (Xu et al., 2010). Therefore, it would be suitable to move both species down to subspecies level, that is - *Helice latimera latimera* Parsis, 1918 and *Helice latimera tientsinensis* Rathbun, 1931. To further tackle this issue, it is desirable to use different methods such as multilocus nuclear DNA, morphological and crossbreeding experiments to evaluate the species status, with the combined efforts of a greater number of scientists.

ACKNOWLEDGMENTS

We are grateful to Dr. Pierre De Wit for his insightful comments on the manuscript. This work was supported by the investigation and assessment on marine medicinal bioresources in China of 908 special program from the State Oceanic Administration under contract No.908-01-ST12 and the special grant of Chongqing University of Arts and Sciences for the person with ability.

REFERENCES

Bandelt HJ, Macaulay V, Richards M (2000). Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Mol. Phylogenet. Evol. 16: 8-28.

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995). Mitochondrial portraits of human populations using median networks. Genetics, 141: 743-753.

Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA, 91:6491-6495.

Chen D (1991). Fisheries Ecology of the Yellow Sea and Bohai Sea. Ocean Press, Beijing,Chn. pp. 48-485.

Dai AY, Yang SL (1991). Crabs of the China seas. China Ocean Press, Beijing, Chin. pp. 551-555.

Dai AY, Song YZ (1977). On Grapsidae (Crustacea: Brachyura) of China. Acta Zool. Sin. 23(4): 360-376.

Daniels SR, Stewart BA, Gouws G (2002). Phylogentic relationships of the southern African freshwater crab fauna (Decapoda: Potamonautidae: *Potamonautes*) derived from multiple data sets reveal biogeographic patterning. Mol. Phylogenet. Evol. 25(3): 511-523.

Excoffier L (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Mol. Ecol. 13 (4): 853-864.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131: 479-491.

Felsenstein J (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39: 783-791.

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. Mol. Mar. Biol. Biotechnol. 3(5): 294-299.

Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147: 915-925.

Hasegawa M, Kishino H, Yano K (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160-174.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. USA, 101: 14812-14817.

Huelsenbeck JP, Rannala B (1997). Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science, 276: 227-232.

Knowlton N,Weigt LA, Solórzano LA, Mills DK, Bermingham E (1993). Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science, 260: 1629-1632.

Li CH, Li TW, Song LS, Su XR (2003). RAPD analysis on intra-species differentiation of *Tegillarca granosa* populations to the south and north of Fujian province. Zool. Res. 24(5): 362-366.

Liu JX, Gao TX, Yokogawa K, Zhang YP (2006). Differential population structuring and demographic history of two closely related fish species, Japanese sea bass (*Latolabrax japonicus*) and spotted sea bass (*Lateolabrax maculatus*) in Northwestern Pacific. Mol. Phylogenet. Evol. 39: 799-811.

Liu JX, Gao TX, Wu SF, Zhang YP (2007). Pleistocene isolation in the Northwestern Pacific marginal seas and limited dispersal in a marine fish, *Chelon haematocheilus* (Temminck & Schlegel, 1845). Mol. Ecol. 16: 275-288.

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

Nielsen R, Matz M (2006). Statistical approaches for DNA barcoding. Syst. Biol. 55: 162-169.

Pan BP, Song LS, Pu WJ, Sun JS (2005). Studies on genetic diversity and differentiation between two allopatric populations of *Cyclina sinensis*. Acta Hydrobiol. Sin. 29(4):372-378.

Planes S, Doherty PJ, Bernardi G (2001). Strong genetic divergence among populations of a marine fish with limited dispersal, *Acanthochromis polyacanthus*, within the great barrier reef and the coral sea. Evolution, 55: 2263-2273.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. Bioinformatics,14: 817-818.

Ray N, Currat M, Excoffier L (2003). Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol. 20: 76-86.

Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9: 552-569.

Rogers AR (1995). Genetic evidence for a Pleistocene population expansion. Evolution, 49: 608-615.

Saitou N, Nei M (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Sakai K, Yatsuzuka K (1980). Notes on some Japanese and chinese *Helice* with *Helice* (*Helicana*) n. subgen., including *Helice* (*Helicana*) *japonica* n. sp. (Crustacea: Decapoda). Senckenberg. Biol. 60: 393-411.

Sakai K, Türkay M, Yang SL (2006). Revision of the *Helice*/*Chasmagnathus* complex (Crustacea: Decapoda: Brachyura). Abh. Senckenberg. Naturforsch. Ges. 565: 1-76.

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

Schneider S, Roessli D, Excoffier L (2000). ARLEQUIN, version 2.0: a software for population genetic data analysis. Geneva: University of Geneva.

Schubart CD, Diesel R, Hedges SB (1998). Rapid evolution to terrestrial

life in Jamaican crabs. Nature, 393: 363-365.
Shih HT, Suzuki H (2008). Taxonomy, phylogeny, and biogeography of the endemic mudflat crab *Helice*/*Chasmagnathus* complex (Crustacea: Brachyura: Varunidae) from East Asia. Zool. Stud. 47(1): 114-125.
Slatkin M, Hudson RH (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics, 129: 555-562.
Sun HY, Jin YH, Zhang DZ, Yang SL, Li Q, Song DX, Zhou KY(2009). Mitochondrial sequence data reveals the phylogeny of the Asian *Helice* group of crabs (Decapoda: Brachyura: Varunidae). J. Zool. Syst. Evol. Res. 47(4): 322-327.
Tajima F (1989). The effect of change in population size on population DNA polymorphism. Genetics, 123: 597-601.
Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10: 512-526.
Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24(8): 1596-1599.
Wang CH, Li CH, Li SF (2008). Mitochondrial DNA-inferred population structure and demographic history of the mitten crab (*Eriocheir sensu stricto*) found along the coast of mainland China. Mol. Ecol. 17: 3515-3527.
Wang P (1999). Response of Western Pacific marginal seas to glacial cycles: paleoceanographic and sedimentological features. Mar. Geol. 156: 5-39.
Xu JM, Yang F, Sun SC(2010). Morphometrical studies on 10 populations of *Helice tientsinensis* and *H. latimera* (Crustacea, Decapoda) from coasts of china. Acta Hydrobiol. Sin. 34(3): 533-540.
Yin W, Fu CZ, Guo L, He QX, Li J, Jin BS, Wu QH, Li B (2009). Species delimitation and historical biogeography in the genus *Helice* (Brachyura: Varunidae) in the northwestern Pacific. Zool. Sci. 26(7): 467-475.