Fish Sci
DOI 10.1007/s12562-015-0928-x

ORIGINAL ARTICLE

Biology

Genetic variation and population structure of the Pacific oyster *Crassostrea gigas* in the northwestern Pacific inferred from mitochondrial COI sequences

Shuang Li[1] · Qi Li[1] · Hong Yu[1] · Lingfeng Kong[1] · Shikai Liu[2]

Received: 27 May 2015 / Accepted: 8 September 2015

Abstract The Pacific oyster *Crassostrea gigas* as one of the important commercial bivalves in the world, is native to the coast of East Asia. The wild populations of this species have declined due to natural or anthropogenic factors in recent years. To provide valuable insights into management and conservation of *C. gigas* we investigated the genetic variation and population structure of *C. gigas* in the northwestern Pacific by analyzing partial mitochondrial cytochrome oxidase I gene (COI) with samples from 12 populations. Moderate levels of genetic diversity and low nucleotide diversity were found within the populations. No significant population structure existed in such a big spatial scale; only low but significant genetic differentiation between some of the Chinese and Japanese populations was observed, which highlighted the important roles of current gene flow and historical re-colonization in deceasing genetic differentiation for marine invertebrates with sedentary adults. The results obtained in this study provide useful information on the genetic diversity and phylogeographic pattern of *C. gigas* and reveal the complex interactions of multiple factors (strong genetic connectivity by dispersal of larvae and ocean currents in this region) in the northwestern Pacific.

Keywords *Crassostrea gigas* · Genetic structure · Mitochondrial DNA · Ocean current

✉ Qi Li
qili66@ouc.edu.cn

[1] Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 266003, China

[2] The Fish Molecular Genetics and Biotechnology Laboratory, Aquatic Genomics Unit, School of Fisheries, Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL 36849, USA

Introduction

Many marine invertebrates have developed a biphasic life history strategy, wherein planktonic larvae disperse for a period of time before metamorphosing into sedentary adults [1, 2]. Therefore, these species have traditionally been recognized to display mild or low geographical differentiation and strong genetic connectivity because of larvae dispersal and insufficient geographical barriers in marine environments [3–5]. The oceanographic processes (e.g., ocean currents) are related to genetic distribution in marine organisms over large spatial scales [6], which is particularly true for adult sedentary marine organisms [7]. There has been a growing number of studies highlighting the relationships among dispersal, gene flow, ocean currents and Pleistocene glacial-interglacial cycles in marine systems, including many species such as fishes [8, 9], crustaceans [10, 11], and mollusks [12–15].

Climatic oscillations during the Pleistocene period had a great influence on the contemporary distribution, abundance and evolutionary histories of biota, which could be expected to have genetic consequences [16, 17]. Several marginal seas in East Asia–the sea of Japan (JPS), the East China sea—the Yellow sea (ECS/YS), and the South China sea (SCS), which separate the East Asian coasts from the Northwestern (NW) Pacific were dramatically impacted by Pleistocene glacial–interglacial cycles [18]. For example, lowered sea levels (with a maximum decline of 120–150 m) during the glacial maxima resulted in the separation of these three seas, which, consequently, fragmented marine ecosystems and isolated populations of aquatic species [19]. The relict fauna in the NW Pacific remained in three marginal seas that have been shown to be separated refugia during glacial periods, accumulating substantial genetic

Published online: 30 September 2015

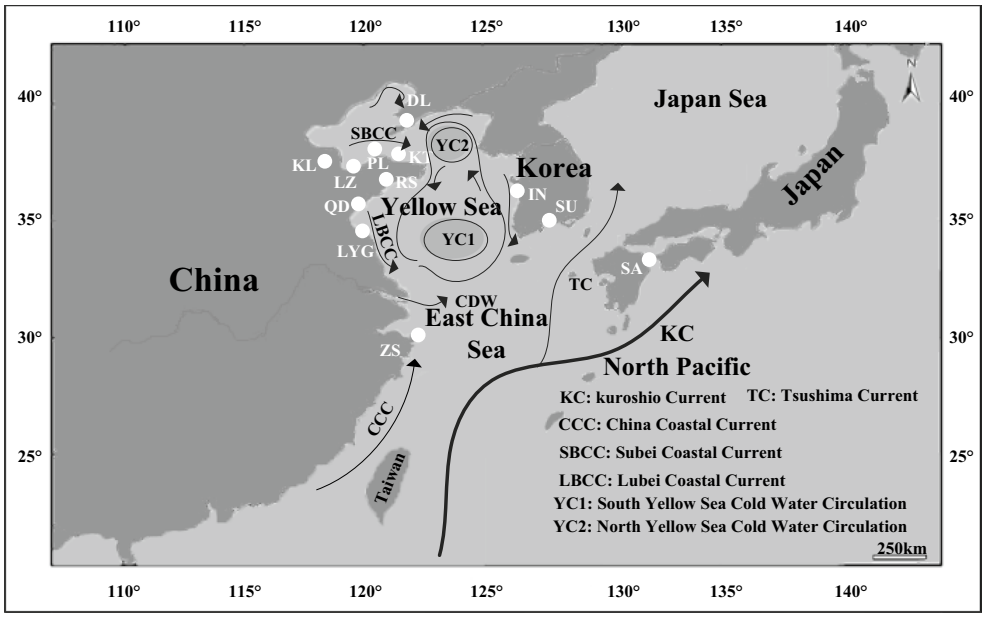

Fig. 1 Map of sampling sites of *C. gigas* from 12 populations and illustration of main ocean currents along the coast of East Asia. *KC* Kuroshio current, *TC* Tsushima current, *CCC* China coastal current, *SBCC* Subei coastal current, *LBCC* Lubei coastal current, *YC1* south Yellow Sea cold water circulation, and *YC2* north Yellow Sea cold water circulation

differentiation [20]. Within interglacial periods, the sea level rose and quickly submerged the land bridges, leading to the reconnection of these marginal seas and the population-level expansion of species [21]. Such recurrent fluctuations in sea levels and water temperature caused by glacial/interglacial cycles inevitably affected the spatial distribution and genetic patterns of marine organisms [9, 11].

The genetic structure of marine species along the coast of East Asia can also be attributed to complex hydrology factors, including both ocean currents and freshwater outflow such as Changjiang diluted water (CDW) [22]. The coast of East Asia areas is dominated by two main ocean systems during summer, namely the Kuroshio and the coastal currents (Fig. 1), both of which are characterized by cyclonic circulation [23, 24]. The Kuroshio featured by high temperature and high salinity consists of a strong main stream and the Tsushima warm current (TC). The main stream flows northward along the east coast of Taiwan, then turns northeast to the southern coast of Kyushu, Japan, whereas the TC enters the JPS. The coastal currents are composed of the China coastal current (CCC) from the SCS, the Subei coastal current (SBCC) and Lubei coastal current (LBCC), all of which are characterized by low salinity flow into the ECS/YS [22]. Theoretically, oceanic currents operating during the pelagic phase should facilitate larvae dispersal and promote genetic exchange [25]. Another important factor is the CDW, which carries huge freshwater outflow into the ECS and profoundly influences the hydrological and ecological features of the ECS [26]. The CDW has been considered to be an extrinsic barrier for genetic connectivity for some intertidal species, such as *Cyclina sinensis* [27] and *Cellana toreuma* [28], while the barrier effect of the Changjiang River outflow on gene flow of intertidal species failed to be detected in some other cases [13]. Therefore, it is of great interest and importance to investigate the historical and contemporary interplay among a complex set of ecological, genetic, demographic, oceanographic and climatic processes in the coast of East Asia during Pleistocene glacial–interglacial cycles.

Crassostrea oysters are of great economic and ecological importance to global fisheries and aquaculture industries [29]. The Pacific oyster, *C. gigas*, with a spatially wide-range distribution, is native to the coast of East Asia, including China, Korea and Japan. As one of the intertidal species, *C. gigas* has strong capability to adapt to various environmental conditions and temperature fluctuations [30]. Despite its abundance and economic and ecological importance, population genetic studies of *C. gigas* in its native habitats are limited. In recent decades, the wild populations of *C. gigas* have declined due to natural or anthropogenic reasons in East Asia (e.g., diseases, overfishing and the deterioration of ocean conditions caused by anthropogenic activities) [31, 32]. Therefore, investigating the population genetic structures of *C. gigas* can facilitate monitoring the genetic diversity and lay the foundation for management and conservation [33]. Given this, an assessment of the genetic diversity of *C. gigas* becomes more urgent.

In this study, we conducted a comprehensive investigation on *C. gigas* sampled from widely distributed natural habitats along the coasts in East Asia using partial mitochondrial cytochrome oxidase I (mtCOI) fragments, with the objective to provide information on the genetic structure and demographic history of *C. gigas*.

Materials and methods

Sample collection

Samples were collected from 12 sites along the coast of Northern China, South Korea and Japan from 2004 to 2014 (Fig. 1; Table 1), including Dalian (DL, China), Kenli (KL, China), Laizhou (LZ, China), Penglai (PL, China), Kongtong (KT, China), Rushan (RS, China), Qingdao (QD, China), Lianyungang (LYG, China), Zhoushan (ZS, China), Incheon (IN, Korea), Suncheon Bay (SU, Korea), and Saiki (SA, Japan). These sampling sites covered the major geographical distribution of *C. gigas* in East Asia. A total of 228 samples were obtained with 20 individuals from each site, except for SU (Korea) where only 8 samples were obtained. Fresh adductor muscles were dissected and flash frozen at −80 °C for DNA extraction.

DNA extraction, species diagnosis, polymerase chain reaction (PCR) amplification and sequencing

Genomic DNA was extracted following a standard phenol–chloroform procedure, as described by Li et al. [34]. The species identity was confirmed by molecular diagnosis using multiplex species-specific PCR analysis as described in Wang and Guo [35], since species identification solely based on phenotypic characteristics (e.g., shell) is not always reliable in oysters.

The universal primers LCO1490 and HCO2198 [36] were used to amplify a mtCOI fragment. PCR amplification was carried out in a 30-μl volume with an initial denaturation step of 3 min at 94 °C; 35 cycles of 94 °C for 45 s, 51 °C for 45 s and 72 °C for 45 s; and a final extension at 72 °C for 10 min. The 30-μl amplification reaction contained 2.0 mM $MgCl_2$, 25 mM of each dNTP, 0.33 μM each primer, 100 ng of template DNA, 3 μl of 10 × buffer, and 1.5 U *Taq* polymerase (TaKaRa). PCR products were purified and sequenced from both directions on an ABI 3730 XM DNA analyzer (Applied Biosystems).

Data analysis

Sequence alignments were performed with CLUSTAL W using default parameters, and manual adjustments were

Table 1 Sampling information, diversity indices and values of neutral tests (Tajima's D & Fu's Fs) for each population of *C. gigas* and the mismatch distribution parameter SSD

Sampling site	Time	Abbreviation	*n*	*s*	*H*	*h*	*π*	*k*	Tajiam's D	Fu's Fs	Goodness-of-fit test	
											SSD	*P*
China												
Dalian	2013	DL	20	6	8	0.5211 ± 0.1346	0.000966 ± 0.000897	0.6	−2.05624[a]	−5.65535[b]	0.0106	0.4
Kenli	2013	KL	20	8	8	0.5895 ± 0.1297	0.001288 ± 0.001084	0.8	−2.17464[a]	−6.10151[b]	0.00251	0.64
Laizhou	2013	LZ	20	3	4	0.2842 ± 0.1284	0.000483 ± 0.000587	0.3	−1.72331[a]	−2.74926[a]	0.00584	0.41
Penglai	2013	PL	20	5	5	0.3684 ± 0.1351	0.000805 ± 0.000799	0.5	−1.97429[a]	−2.99084[a]	0.00025	0.89
Kongtong	2013	KT	20	8	9	0.6526 ± 0.1225	0.001432 ± 0.001165	0.9	−2.04091[a]	−7.38334[b]	0.00774	0.4
Rushan	2013	RS	20	7	7	0.5211 ± 0.1346	0.001127 ± 0.000991	0.7	−2.12144[a]	−4.98359[b]	0.00029	0.99
Qingdao	2008	QD	20	6	6	0.5158 ± 0.1316	0.001110 ± 0.000982	0.7	−1.88764[a]	−3.46687[a]	0.00036	0.94
Lianyungang	2013	LYG	20	7	7	0.5211 ± 0.1346	0.001127 ± 0.000991	0.7	−2.12144[a]	−4.98359[a]	0.00029	0.95
Zhoushan	2013	ZS	20	2	3	0.1947 ± 0.1145	0.000322 ± 0.000466	0.2	−1.51284[a]	−1.86305[a]	0.0016	0.41
Korea												
Incheon	2014	IN	20	5	6	0.5158 ± 0.1316	0.000949 ± 0.000887	0.6	−1.78003[a]	−4.01486[a]	0.01115	0.25
Suncheon Bay	2012	SU	8	1	2	0.2500 ± 0.1802	0.000403 ± 0.000571	0.3	−1.0548	−0.182	0.27914	0.05
Japan												
Saiki	2004	SA	20	4	4	0.4316 ± 0.1262	0.001187 ± 0.001026	0.7	−1.003	−0.6996	0.02561	0.37

Sampling year (Time); Abbr, site abbreviation; number of individuals sampled per site (*n*); number of haplotype per locality (*H*); number of polymorphic sites per locality (*s*); haplotype diversity (*h*); nucleotide diversity (*π*); and mean number of pairwise differences (*k*)

[a] Significant Tajima's *D* ($P < 0.05$) and Fu's Fs values ($P < 0.02$)

[b] Significant Tajima's *D* and Fu's Fs values ($P < 0.0001$)

made in MEGA 5.0 [37]. DNASP 5 [38] was used to estimate the total number of haplotypes (H) and polymorphic sites (S). All of haplotypes for each locus were deposited in the Genbank database (accession numbers: KP099007–KP099052). Molecular diversity indices such as haplotype diversity (h), polymorphic sites (s), nucleotide diversity (π) for each population, and mean number of pairwise differences (k) were calculated using ARLEQUIN 3.5 [39].

The genetic structure among all populations was assessed by analysis of molecular variance (AMOVA) [40] using ARLEQUIN 3.5. Population pairwise Φ_{ST} values were also computed with ARLEQUIN. The significance of each pairwise comparison was tested with 10,000 random replicates, and a standard Bonferroni correction was applied for multiple tests [41]. Additionally, the isolation by distance (IBD) pattern was tested for mtCOI based on pairwise geographical distance (km) and genetic distances (F_{ST}) with the Mantel test using the IBD web service program [42]. The significance of the correlation was tested using permutation methods (10,000 randomizations).

The demographic history of populations was estimated by conducting Tajima's D [43] and Fu's Fs tests [44]. Both of which were useful to measure deviations from neutrality. Statistics in mismatch distributions was obtained to test for a departure from sudden population expansion [sum of squared deviation (SSD) to test the validity of the sudden expansion model] [45, 46]. The significance of SSD was assessed by 1000 parametric bootstrap replicates. Both neutrality tests and mismatch analyses were calculated in ARLEQUIN 3.5. A mismatch frequency histogram was generated in DNASP 5 to further detect the historical expansion model. The real time since population expansion (t) was calculated through the formula $t = \tau/2\ \mu k$, where τ is the parameter tau estimated under the sudden expansion model in ARLEQUIN 3.5, μ is per locus per year substitution rate (divergence time/2) and k is the sequence length. We adopted an mtCOI divergence time estimated for molluscan species from 2 % [47] to 2.4 % [33] for the lack of a clear fossil or geological record.

Bayesian inference (BI) and neighbor joining (NJ) analyses were performed to infer the phylogenies of the haplotypes. We used *Crassostrea angulata* (KC170323) as an out group. The software jModelTest [48, 49] was applied to select the best-fit nucleotide substitution model for COI fragments using the Akaike information criterion (AIC). The Hasegawa, Kishino and Yano (HKY) model was selected for Bayesian analysis that was performed with MrBayes 3.1.2 [50]. Each analysis used four Markov chains with default heating values and was run for 10 million generations while sampling trees every 1000 generations. The first 70 % of generations were discarded as burn-in, and then a 50-% majority-rule consensus tree with branch lengths was constructed with the remaining generations. The NJ analysis was performed in MEGA 5.0. A haplotype network was generated using the 95 % statistical parsimony algorithm implemented in TCS 1.21 [51] to evaluate the genetic relationships of *C. gigas* haplotypes.

Results

Genetic diversity

A total of 621 bp fragments of mtCOI gene generated from 228 specimens were used for the analyses. Sequences harbored 46 polymorphic sites, including 44 transitions and two transversions with no indels being detected. We found 46 haplotypes, 38 of which were private. The number of haplotypes for each population ranged from two to nine (Table 1). The unique conspicuous high-frequency haplotype (H·1), also the most abundant haplotype, was observed in all populations accounting for 73.25 % (167/228) of the total haplotypes identified (Table 2), followed by H·8 that was present in five populations (KL, LZ, KT, QD and ZS) with only six copies. Other low-frequency haplotypes except for the private ones were shared by only 2–3 populations with at most two copies. Existence of private haplotypes was observed in all sampling sites except for Suncheon Bay (SU) in Korea.

All populations were characterized by low nucleotide diversity (0.001432–0.000322; Table 1), and moderate haplotype diversity (mean $h = 0.4569$). The mean number of pairwise differences between individuals within sampling sites ranged from 0.2 in ZS to 0.9 in KT. Molecular diversity indices indicated that KT samples scored the highest value of genetic diversity, while ZS samples occupied the lowest.

Population genetic structure

Population pairwise Φ_{ST} values were generally low with non-significant value, only low but significant genetic differentiation was observed in some samples between Chinese and Japanese populations (Table 3; $P < 0.05$), indicating high genetic connectivity among these populations. However, remarkably, AMOVA analysis using two grouping strategies (Table 4) revealed that a majority of the total molecular variance was distributed within groups (I: 94.97 %, II: 98.98 %) with significant values ($P < 0.05$), followed by variance distributed among groups (5.31 %) and among populations within groups (0.52 %), respectively. The variance components among groups and among populations within groups were not statistically significant ($P > 0.05$) using both grouping strategies in AMOVA analysis. IBD analyses showed that genetic distances were not correlated with geographic distances ($P = 0.886$,

Table 2 Haplotype frequencies of mitochondrial COI fragments observed among 12 populations

Haplotypes	Haplotype frequencies												
	DL	KL	LZ	PL	KT	RS	QD	LYG	ZS	IN	SU	SA	Total
	20	20	20	20	20	20	20	20	20	20	8	20	228
H·1	14	13	17	16	12	14	14	14	18	14	7	15	168
H·2	1												1
H·3	1												1
H·4	1												1
H·5	1												1
H·6	1												1
H·7	1												1
H·8		1	1		1		2		1				6
H·9							1						1
H·10							1						1
H·11		1					1						2
H·12							1						1
H·13						1							1
H·14						1				1			2
H·15						1							1
H·16						1							1
H·17						1							1
H·18		1				1							2
H·19										2			2
H·20										1			1
H·21										1			1
H·22										1			1
H·23		1											1
H·24		1											1
H·25		1											1
H·26		1			1			1					3
H·27			1										1
H·28			1										1
H·29												1	1
H·30												1	1
H·31												3	3
H·32					1								1
H·33					1								1
H·34					1								1
H·35					1								1
H·36					1						1		2
H·37					1								1
H·38				1				1					2
H·39								1					1
H·40								1					1
H·41								1					1
H·42								1					1
H·43									1				1
H·44				1									1
H·45				1									1
H·46				1									1

Table 3 Pairwise Φ_{ST} based on mitochondrial COI fragments among the 12 populations (see Table 1 for abbreviations)

	DL	KL	LZ	PL	KT	RS	QD	LYG	ZS	IN	SU
KL	0.00011										
LZ	0.00005	−0.00926									
PL	0.00010	0.00006	0.00010								
KT	0.00718	−0.01154	−0.00814	0.00766							
RS	−0.00772	−0.00661	0.00000	0.00018	0.00678						
QD	0.00824	−0.02012	−0.00972	0.00882	−0.01210	0.00760					
LYG	−0.00011	−0.00682	−0.00022	−0.00834	0.00036	0.00011	0.00736				
ZS	−0.00006	−0.01033	−0.02035	0.00008	−0.00902	−0.00007	−0.01098	−0.00024			
IN	0.00883	0.00757	0.00117	0.00977	0.01419	0.00058	0.01622	0.00824	0.01313		
SU	−0.02764	−0.03404	−0.00660	−0.02290	−0.04699	−0.03122	−0.02120	−0.03137	0.00933	−0.01618	
SA	0.04531*	0.03965*	0.05756	0.04133	0.02631	0.04229*	0.04931*	0.04219	0.06324	0.04249	−0.04296

* Significant Φ_{ST} values at $P < 0.05$

Table 4 Analysis of molecular variance (AMOVA) among *C. gigas* mitochondrial COI sequences

Source of variation	d.f.	Φ statistics	% of variation	P value
I (China and Korea/Japan)				
Among groups	1	$\Phi_{CT} = 0.05310$	5.31	0.08525
Among populations within groups	10	$\Phi_{SC} = -0.00291$	−0.28	0.63059
Within populations	216	$\Phi_{ST} = 0.05034$	94.97	0.04891*
II (Yellow sea/East China sea)				
Among groups	1	$\Phi_{CT} = 0.00499$	0.50	0.21485
Among populations within groups	10	$\Phi_{SC} = 0.00526$	0.52	0.17871
Within populations	216	$\Phi_{ST} = 0.01023$	98.98	0.04832*

* Significant Φ_{ST} values at $P < 0.05$

$r^2 = 0.0692$; Fig. 2). These results indicated that significant phylogeographic structures of *C. gigas* were rejected in the study areas.

Historic demographic expansions

Two neutrality test values for all but two populations (SU and SA, $P > 0.05$, Table 1) were negative and most of them were highly significant ($P < 0.02$), demonstrating a demographic expansion event under the neutral model within these groups. The mismatch distribution displayed a negative binomial curve (Fig. 3) that was likely related to a recent population expansion. Furthermore, the SSD values of mismatch analyses were statistically non-significant (all $P > 0.05$, Table 1). All the results above supported the null hypothesis of recent population expansion of *C. gigas*. Based on the formula $t = \tau/2\mu\kappa$, population expansion time estimated for *C. gigas* started 52–40 Kya (thousand years ago).

Phylogenetic analyses

The Bayesian phylogenetic tree was shallow with no significant genealogical branches or clusters of samples corresponding to sampling sites (Fig. 4), which was also consistent with the observation that a population expansion event occurred after a bottleneck. A further NJ tree

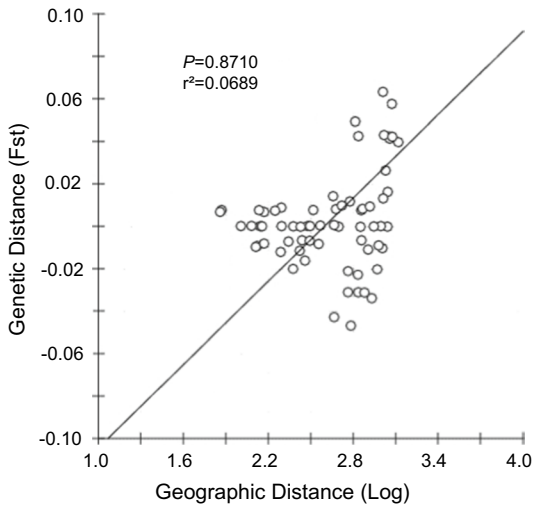

Fig. 2 IBD plots for *C. gigas*. The *plot* shows a relationship of pairwise population genetic and geographic distance among 12 populations along the coast of East Asia

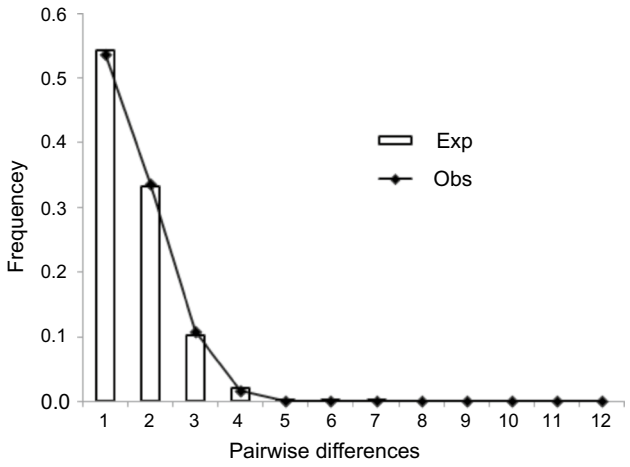

Fig. 3 The observed (*line*) and expected (*bars*) mismatch distributions under the recent expansion model for the mitochondrial cytochrome oxidase I (COI) of *C. gigas*

(not displayed) also supported the BI analysis results. The haplotype network showed a shallow and typical 'star-like' topology dominated by a major haplotype shared by all 12 populations (Fig. 5). The major haplotype (H·1) lay at the center, with many singletons or low-frequency haplotypes being linked by short branches.

Discussion

Population genetic diversity

In this study, a current genetic background of *C. gigas* in East Asia's coastal areas was provided. The level of genetic diversity of *C. gigas* was characterized by moderate haplotype diversity (4 of 12 populations had low values, including LZ, PL, ZS and SU) and low nucleotide diversity, which was partly consistent with the previous conclusions reported by Sekino et al. [52]. In addition, the level of genetic variation for *C. gigas* was relatively lower than that of some co-distributed bivalves (Table 5), such as *C. sinensis* [13], *Tegillarca granosa* [14], *Crassostrea ariakensis* [53] and *Atrina pectinate* [21]. A different population history of these species was assumed to be responsible for the differences in genetic variations.

Population structure

High levels of gene flow and weak genetic differentiation among *C. gigas* populations were revealed by the AMOVA and pairwise Φ_{ST} analyses. The genetic pairwise differentiations (Φ_{ST}) between *C. gigas* populations were low, which is frequently observed in high dispersal marine invertebrates [54]. The weak differentiation was further revealed by the Bayesian phylogenetic tree and haplotypes networks. The major haplotype occurred in all sampling sites, demonstrating strong historical or current gene flow among all the populations at this spatial scale [16, 20]. In addition, no evidence of IBD was found in this study. Weak genetic differentiation of *C. gigas* was also detected by Sekino et al. [52] in a fine scale. Moreover, similar population structures have been revealed in some other marine species distributed in these coastal areas, such as *Chelon haematocheilus* [8], *Eriocheir* sensu *strict* [11], *A. pectinate* [21], Chinese shrimp *Feneropenaeus chinensis* [55] and *Rapana venosa* [56, 57].

The possibility of anthropogenic habitat expansion that leads to gene flow can be ruled out by the observation that the *C. gigas* population samples shared very few mtDNA haplotypes [52]. For geographically isolated marine invertebrate species, pelagic larval duration (PLD) is one of the key parameters in shaping the pattern of genetic population differentiation [58]. The longer the PLD of a species, the farther distance it can spread via passive transport by ocean currents flow, though some exceptions exist [59]. A 2- to 3-week long larval stage of *C. gigas* combined with insufficient geographical barriers, theoretically, allows the larvae to migrate long distances with the enhancement of ocean currents.

The intricate hydrology conditions along the coast of East Asia could contribute to the genetic homogeneity among geographically isolated species. The KC (Fig. 1) flows northward along the east coast of Taiwan, then turns northeast to the southern coast of Kyushu, Japan; the CCC flows into the ECS from the SCS; the LBCC and SBCC flows southward to the ECS. Moreover, there is a local circulation called the Yellow Sea cold water circulation (including YC1 and YC2) in the YS during summer season. These currents operating in the regions potentially enhance dispersal of marine larvae and, consequently, decrease population structures among seas [10]. In contrast, the CDW seemingly fails to hinder the dispersal of the *C. gigas* larvae. No significant genetic differentiation of *C. gigas* was detected in line with the Changjiang River discharge, which may mainly benefit from its strong adaptability to various environments such as a euryhaline adaptation mechanism [60, 61]. Obviously, weak genetic structure, as revealed by mtCOI data in this study, is consistent with the view that marine organisms with larval stages often show low genetic differentiation among populations as a result of high dispersal capabilities and large-scale oceanic mixture [62, 63]. However, low, but significant, genetic differentiation (only 4 of 66 pairwise Φ_{ST} values) observed between some of the Chinese samples and Japanese samples might be the result of demographic fluctuation, which could be caused by any of several other factors, including genetic drift before recruitment and natural selection during early life stages [21, 64, 65].

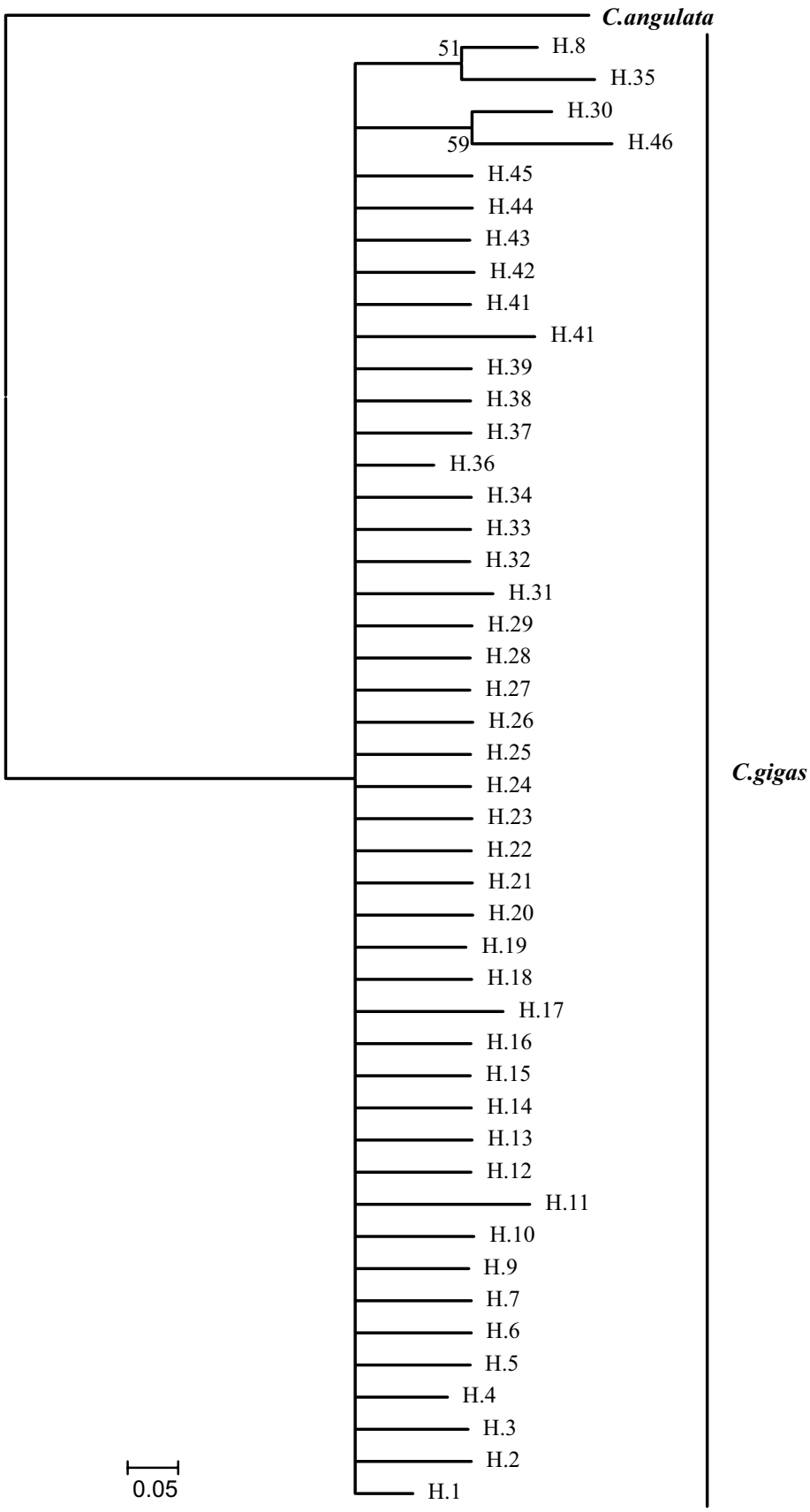

Fig. 4 Phylogenetic trees of *C. gigas* based on COI haplotypes using Bayesian analyses. The *numbers* above the branches are the Bayesian posterior probabilities

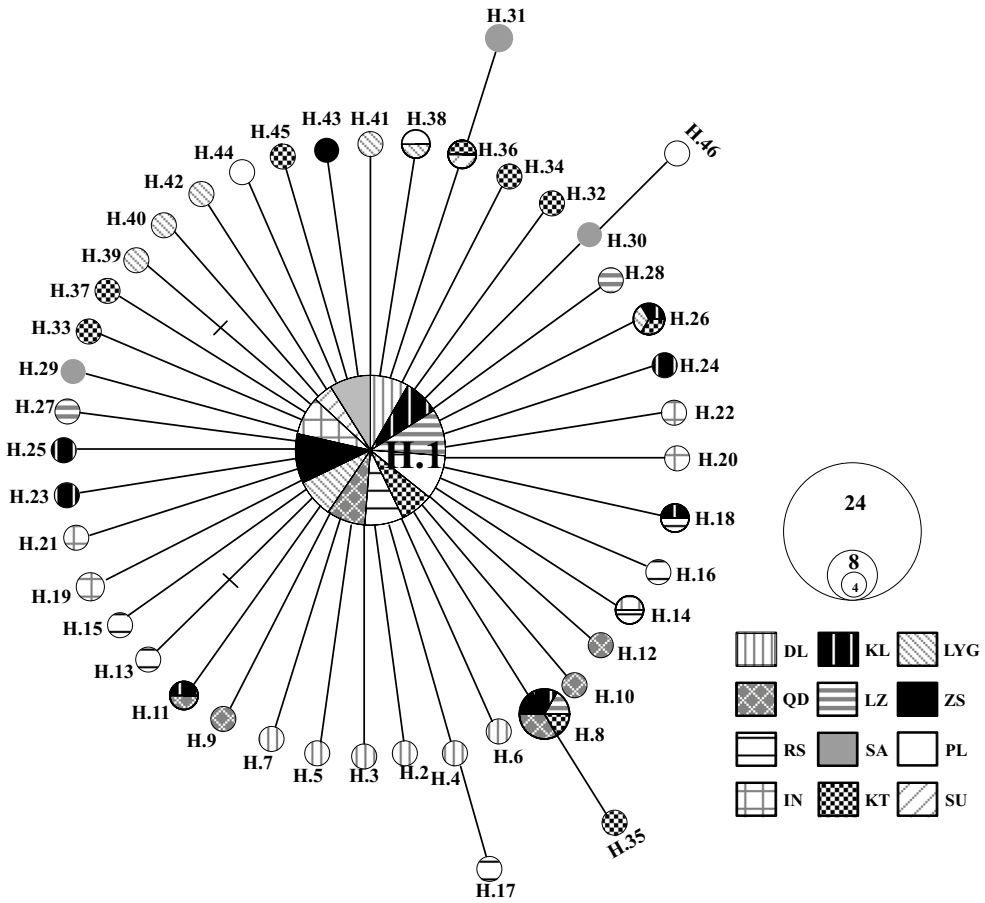

Fig. 5 Haplotype network for the 46 haplotypes of *C. gigas*. Different populations were marked by distinctive colors. Each *circle* corresponds to a unique haplotype sized in proportion to its frequency. The number of substitutions separating two haplotypes was indicated by the *vertical bars* on the *line*

Table 5 Two molecular diversity indices and molecular markers on five co-distributed bivalves

Bivalve	Molecular makers	$\bar{h}$	$\bar{\pi}$	Source
Tegillarca granosa	mtDNA: COI ncDNA: TIS-1	0.539	0.01240	Ni et al. [13]
Cyclina sinensis	mtDNA: COI ncDNA: TIS-1	0.571	0.00200	Ni et al. [13]
Crassostrea ariakensis	mtDNA: COI 16S,COII,COIII, Cytb	0.911	0.00100	Kim et al. [13, 53]
Atrina pectinata	mtDNA: COI Microsatellites	0.850	0.00380	Dong et al. [13, 28, 53]
Crassostrea gigas	mtDNA: COI	0.457	0.00097	Present study

$\bar{h}$ haplotype diversity, $\bar{\pi}$ nucleotide diversity, *mtCOI-COIII* mitochondrial cytochrome oxidase subunit I–III, *ncDNA* nuclear DNA, *Cytb* cytochrome b, *ITS*-1 internal transcribed spacer, and *SSR* microsatellite markers

Phylogenetic analyses

The simple topology of the Bayesian phylogenetic tree together with a shallow and "star-like" haplotype network revealed no significant genealogical branches corresponding to sampling sites. Such phylogenetic characteristics suggest that *C. gigas* originates from a common ancestral population, and the major haplotype found in all

populations is highly likely to be the ancestral haplotype. Notably, most investigations with respect to population genetics and phylogeographic pattern of marine species have indicated that three marginal seas along the NW Pacific coast have served as independent refugiums during the low sea level periods of Pleistocene glaciations [20]. Therefore, marine organisms and their progenies might be divided into ECS, SCS, and JPS lineages after the rising of post-glacial sea levels. This observation has been reported in several organisms, such as *C. haematocheilu* [8] and *C. sinensis* [13]. Based on our sampling scope and the population genetic structure revealed, we infer that the unique lineage of *C. gigas* in the present study might belong to the ECS lineage.

Historical demography

The NW Pacific Ocean has been greatly influenced by Pleistocene glacial–interglacial cycles, especially in the last glacial maximum (LGM, about 20,000 years ago) [47]. Most of the ECS/YS shelf was exposed during glacial periods, and marine organisms gathered into the Okinawa refugia [66]. Given the genetic patterns reflected by our data, a post-LGM expansion can be assumed for the *C. gigas* population. However, the inferred real expansion time starting from 42 to 51 Kya ($\tau = 0.645$), perhaps after the last interglacial maximum (LIG, 115 Kya) but not post-LGM, contradicts with the classic view of significant expansion after the LGM. Probably, the real expansion time was influenced by the short length of the analyzed COI region in our study. Similar to our findings, the pre-LGM expansion has been reported in some other marine species [47, 67, 68], but the specific reasons for this conflicted observations remain elusive, requiring further investigations.

From our results, moderate haplotype diversity and low nucleotide diversity were detected within native populations of *C. gigas* in East Asia. There being no significant genetic structure among the populations suggests a strong genetic connection among these populations. Only low, but significant genetic differentiation between some of the Chinese and Japanese populations might be attributed to both current gene flow (by larval dispersal interacting with oceanographic processes, such as ocean currents) and historical re-colonization (by population range expansion and demographic expansion in the late Pleistocene), suggesting a strong genetic connection among these populations. Consequently *C. gigas* distributed in the sampling areas could be considered as an ECS lineage. The results obtained in the present study provide useful information on the genetic diversity and phylogeographic patterns of *C. gigas* and reveal the complex interactions of multiple factors in the NW Pacific Ocean.

Acknowledgments This study was supported by grants from the National High Technology Research and Development Program (2012AA10A405-6) and the National Natural Science Foundation of China (31372524 and 41276138).

References

1. Bradbury IR, Snelgrove PV (2001) Contrasting larval transport in demersal fish and benthic invertebrates: the roles of behaviour and advective processes in determining spatial pattern. Can J Fish Aquat Sci 58:811–823
2. Purcell JF, Cowen RK, Hughes CR, Williams DA (2006) Weak genetic structure indicates strong dispersal limits: a tale of two coral reef fish. Proc R Soc B Biol Sci 273:1483–1490
3. Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Ann Rev Ecol Syst 25:547–572
4. Bird CE, Holland BS, Bowen BW, Toonen RJ (2007) Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp) with similar life histories. Mol Ecol 16:3173–3186
5. Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84:2007–2020
6. Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Annu Rev Ecol Evol Syst 40:291–310
7. Kinlan BP, Gaines SD, Lester SE (2005) Propagule dispersal and the scales of marine community process. Divers Distrib 11:139–148
8. 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
9. Shen KN, Jamandre BW, Hsu CC, Tzeng WN, Durand JD (2010) Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*. BMC Evol Biol 11:83
10. Tsang LM, Chan BK, Ma KY, Chu KH (2008) Genetic differentiation, hybridization and adaptive divergence in two subspecies of the acorn barnacle *Tetraclita japonica* in the northwestern Pacific. Mol Ecol 17:4151–4163
11. Xu J, Chan TY, Tsang LM, Chu KH (2009) Phylogeography of the mitten crab *Eriocheir* sensu stricto in East Asia: Pleistocene isolation, population expansion and secondary contact. Mol Phylogenet Evol 52:45–56
12. Kenchington E, Patwary M, Zouros E, Bird C (2006) Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). Mol Ecol 15:1781–1796
13. Ni G, Li Q, Kong L, Zheng X (2012) Phylogeography of bivalve *Cyclina sinensis*: testing the historical glaciations and Changjiang River outflow hypotheses in northwestern Pacific. PLoS One 7:e49487
14. Ni G, Li Q, Kong L, Zheng X (2012) Phylogeography of the bivalve *Tegillarca granosa* in coastal China: implications for management and conservation. Mar Ecol Prog Ser 452:119–130
15. Li L, Wu X, Yu Z (2013) Genetic diversity and substantial population differentiation in *Crassostrea hongkongensis* revealed by mitochondrial DNA. Mar Genom 11:31–37
16. Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge
17. Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proc Natl Acad Sci 97:9115–9120
18. Wang P, Bradshaw M, Ganzei S, Tsukawaki S, Hanssan K, Hantoro W, Poobrasert S, Burne R, Zhao Q (1997) West pacific

marginal seas during last glacial maximum: amplification of environmental signals and its impact on monsoon climate. In: Proc 30th international geology congress pp 65–86
19. Kaiser TS, Neumann D, Heckel DG, Berendonk TU (2010) Strong genetic differentiation and postglacial origin of populations in the marine midge *Clunio marinus* (*Chironomidae, Diptera*). Mol Ecol 19:2845–2857
20. Ni G, Li Q, Kong L, Yu H (2014) Comparative phylogeography in marginal seas of the northwestern Pacific. Mol Ecol 23:534–548
21. Xue DX, Wang HY, Zhang T, Liu JX (2014) Population genetic structure and demographic history of *Atrina pectinata* Based on mitochondrial DNA and microsatellite Markers. PLoS One 9:e95436
22. Su J, Yuan Y (2005) Coastal hydrology in China. Ocean Press, Beijing, p 367 **(in Chinese)**
23. Zheng Q, Fang G, Song YT (2006) Introduction to special section: dynamics and circulation of the Yellow, East, and South China Seas. J Geophys Res Oceans 111:1978–2012
24. Liu J (2013) Status of marine biodiversity of the China Seas. PLoS One 8:e50719
25. Jelee H, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific *littorinid gastropods*: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. Mol Ecol 18:2165–2184
26. Wang Bd, Wang XL, Zhan R (2003) Nutrient conditions in the Yellow Sea and the East China Sea. Estuar Coast Shelf S 58:127–136
27. Zhao Y, Li Q, Kong L, Mao Y (2009) Genetic and morphological variation in the venus clam *Cyclina sinensis* along the coast of China. Hydrobiologia 635:227–235
28. Dong YW, Wang HS, Han GD, Ke CH, Zhan X, Nakano T, Williams GA (2012) The impact of Yangtze River discharge, ocean currents and historical events on the biogeographic pattern of *Cellana toreuma* along the China coast. PLoS One 7:e36178
29. Ren J, Liu X, Jiang F, Guo X, Liu B (2010) Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. BMC Evol Biol 10:394
30. Miossec L, Le Deuff RM, Goulletquer P (2009) Alien species alert: *Crassostrea gigas* (Pacific oyster). ICES Cooperative Res Rep 299:1017–6195
31. Aranishi F (2006) A novel mitochondrial intergenic spacer reflecting population structure of Pacific oyster. J Aappl Genet 47:119–123
32. Guo X, Zhang G, Qian L, Wang H, Liu X, Wang A (2006) Oysters and oyster farming in China: a review. J Shellfish Res 25:734
33. Lazoski C, Gusmão J, Boudry P, Solé-Cava AM (2011) Phylogeny and phylogeography of Atlantic oyster species: evolutionary history, limited genetic connectivity and isolation by distance. Mar Ecol Prog Ser 426:197–212
34. Li Q, Park C, Kijima A (2002) Isolation and characterization of microsatellite loci in the Pacific abalone, *Haliotis discus hannai*. J Shellfish Res 21:811–816
35. Wang H, Guo X (2008) Identification of *Crassostrea ariakensis* and related oysters by multiplex species-specific PCR. J Shellfish Res 27:481–487
36. Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299
37. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739
38. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452
39. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567
40. 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
41. Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225
42. Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. BMC Genet 6:13
43. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 23:585–595
44. Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925
45. Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66:591–600
46. 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
47. Marko PB, Hoffman JM, Emme SA, McGovern TM, Keever CC, Nicole Cox L (2010) The 'Expansion-Contraction' model of Pleistocene biogeography: rocky shores suffer a sea change? Mol Ecol 19:146–169
48. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256
49. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704
50. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574
51. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659
52. Sekino M, Sato S, Hong JS, Li Q (2012) Contrasting pattern of mitochondrial population diversity between an estuarine bivalve, the Kumamoto oyster *Crassostrea sikamea*, and the closely related Pacific oyster *C. gigas*. Mar Biol 159:2757–2776
53. Kim WJ, Dammannagoda ST, Jung H, Baek IS, Yoon HS, Choi SD (2014) Mitochondrial DNA sequence analysis from multiple gene fragments reveals genetic heterogeneity of *Crassostrea ariakensis* in East Asia. Genes Genom 36(5):1–14
54. Zulliger DE, Tanner S, Ruch M, Ribi G (2009) Genetic structure of the high dispersal Atlanto-Mediterreanean sea star *Astropecten aranciacus* revealed by mitochondrial DNA sequences and microsatellite loci. Mar Biol 156:597–610
55. Li YL, Kong XY, Yu ZN, Kong J, Ma S, Chen LM (2009) Genetic diversity and historical demography of Chinese shrimp *Feneropenaeus chinensis* in Yellow Sea and Bohai Sea based on mitochondrial DNA analysis. Afr J Biotechnol 8:1193–1202
56. Yang JM, Li Q, Zheng XD, Song ZL, Wang RC (2008) Genetic diversity in populations of *Rapana venosa* in coastal waters of China. Oceanol Limnol Sinica 39:257–262
57. Chandler E, McDowell J, Graves J (2008) Genetically monomorphic invasive populations of the rapa whelk, *Rapana venosa*. Mol Ecol 17:4079–4091
58. Selkoe K, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Mar Ecol Prog Ser 436:291–305

59. Shanks A (2009) Pelagic larval duration and dispersal distance revisited. Biol Bull 216:373–385
60. Meng J, Zhu Q, Zhang L, Li C, Li L, She Z, Huang B, Zhang G (2013) Genome and transcriptome analyses provide insight into the euryhaline adaptation mechanism of *Crassostrea gigas*. PLoS One 8:e58563
61. Zhao X, Yu H, Kong L, Liu S, Li Q (2014) Comparative transcriptome analysis of two oysters, *Crassostrea gigas* and *Crassostrea hongkongensis* provides insights into Adaptation to Hypo-Osmotic conditions. PLoS One 9:e111915
62. Reichow D, Smith M (2001) Microsatellites reveal high levels of gene flow among populations of the *California squid Loligo opalescens*. Mol Ecol 10:1101–1109
63. Katsares V, Tsiora A, Galinou MS, Imsiridou A (2008) Genetic structure of the endangered species *Pinna nobilis* (*Mollusca: Bivalvia*) inferred from mtDNA sequences. Biologia 63:412–417
64. Johnson M, Black R (1982) Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. Mar Biol 70:157–164
65. Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002) Genetic assessment of connectivity among marine populations. B Mar Sci 70:273–290
66. Wang P (1999) Response of Western Pacific marginal seas to glacial cycles: paleoceanographic and sedimentological features. Mar Geol 156:5–39
67. Derycke S, Remerie T, Backeljau T, Vierstraete A, Vanfleteren J, Vincx M, Moens T (2008) Phylogeography of the *Rhabditis* (*Pellioditis*) marina species complex: evidence for long-distance dispersal, and for range expansions and restricted gene flow in the North-East Atlantic. Mol Ecol 17:3306–3322
68. Laakkonen HM, Lajus DL, Strelkov P, Väinölä R (2013) Phylogeography of amphi-boreal fish: tracing the history of the Pacific herring *Clupea pallasii* in North-East European seas. BMC Evol Biol 13:67