Population subdivision of the surf clam *Mactra chinensis*6.66 024.378(A)40a0.2 3e

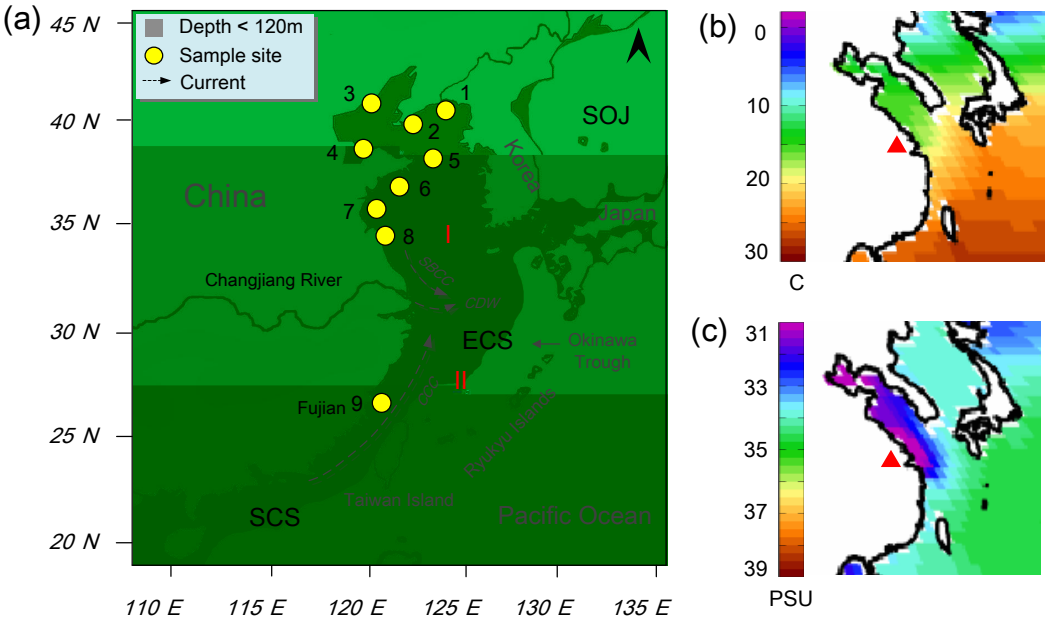
Submitted 31 July 2015
Accepted 22 August 2015
Published 29 September 2015

Corresponding author
Qi Li, qili66@ouc.edu.cn

Academic editor
Peter Prentis

Additional Information and Declarations can be found on page 15

DOI **10.7717/peerj.1240**


(a)
45 N
40 N
35 N
30 N
25 N
20 N
110 E
115 E
120 E
125 E
130 E
135 E
Depth < 120m
Sample site
Current
SOJ
China
Japan
Changjiang River
ECS
Okinawa Trough
Fujian
Ryukyu Islands
Taiwan Island
SCS
Pacific Ocean
KSW
SBCC
CDW
ZCC
I
II
1
2
3
4
5
6
7
8
9
(b)
0
10
20
30
C
(c)
31
33
35
37
39
PSU

species occurring in the ECS, especially for intertidal species having experienced a direct loss of habitat (*Ni et al., 2012*). They would have been forced into the refugium during the

Table 1 Sampling details and diversity indices (with standard deviation) for the nine populations of Mactra chinensis.

Sample site (Abbr.)	Latitude, longitude	Mitochondrial diversity				Microsatellite diversity			
		N	n	h	π	N	A_R	H_E	H_O
1. Dandong (DD)	39°50′N, 124°10′E	23	6	0.632 (0.090)	0.00230 (0.00161)	42	16.56 (3.60)	0.87 (0.07)	0.56 (0.09)
2. Zhuanghe (ZH)	39°38′N, 122°59′E	16	5	0.533 (0.142)	0.00193 (0.00145)	60	17.07 (3.42)	0.89 (0.05)	0.67 (0.07)
3. Qinhuangdao (QH)	39°54′N, 119°36′E	18	5	0.549 (0.127)	0.00148 (0.00119)	50	17.38 (4.24)	0.88 (0.07)	0.58 (0.11)
4. Penglai (PL)	37°50′N, 120°44′E	21	9	0.757 (0.088)	0.00224 (0.00160)	60	17.63 (4.08)	0.90 (0.04)	0.54 (0.12)
5. Wendeng (WD)	36°54′N, 122°03′E	14	9	0.879 (0.079)	0.00373 (0.00243)	55	16.49 (3.59)	0.88 (0.08)	0.56 (0.11)
6. Haiyang (HY)	36°34′N, 121°01′E	15	10	0.942 (0.040)	0.00305 (0.00206)	54	17.11 (3.25)	0.90 (0.04)	0.53 (0.17)
7. Rizhao (RZ)	35°06′N, 119°23′E	20	9	0.790 (0.086)	0.00250 (0.00174)	60	16.59 (3.81)	0.87 (0.08)	0.48 (0.11)
8. Lianyungang (LY)	34°34′N, 119°41′E	19	10	0.784 (0.098)	0.00350 (0.00226)	60	17.17 (3.35)	0.88 (0.06)	0.63 (0.07)
9. Pingtan (PT)	25°30′N, 119°54′E	20	9	0.705 (0.114)	0.00390 (0.00245)	60	14.27 (5.11)	0.81 (0.09)	0.57 (0.09)

Notes.

N, number of individuals analysed; n, number of haplotypes; h, haplotype diversity; π, nuclear diversity; A_R, mean allele richness; H_E, expected heterozygosity; H_O, observed heterozygosity; Abbr, site abbreviation.

survival rate of nearly 100% during the pelagic stage from day 0 to day 10 at about 23 °C. Higher temperature (>30 °C) can cause high mortality. When the salinity is <15 PSU, the development of larvae will be significantly refrained. These abiotic and biotic factors may have complex interactions on the population structure of *M. chinensis*, which need to be carefully disentangled. Here we will first examine whether the current ECS populations have a single origin, and then test whether there is genetic subdivision among the populations. We will discuss the results in the context of potential influence of diverse factors.

MATERIALS AND METHODS

Sample collection and DNA extraction

Sampled sites were chosen to cover the main natural range of *M. chinensis* in the ECS according to the description of *Xu (1997)*. As it is not endangered or protected species, no specific permit is required for collection. Finally, a total of 501 specimens of nine natural populations were sampled from public access areas between May 2006 and September 2010 (Table 1). We failed to collect any populations from Lianyungang (LY) to Pingtan (PT) with a straight-line distance of approximately 1,000 km during our several years' fieldwork, which was mainly due to unsuitable habitat types (mud substrate) around the estuary. The adductor muscle was excised from each individual and preserved in 100% ethanol immediately until DNA preparation. Genomic DNA was extracted from approximate 50 mg muscle tissue using a standard phenol-chloroform extraction method as described by *Li, Park & Kijima (2002)*.

Mitochondrial DNA sequencing and analysis

The mitochondrial COI gene was amplified for a subset of individuals (14 to 23 specimens) in each population with the primers LCO-1490 and HCO-2198 (*Folmer et al., 1994*). Each polymerase chain reaction (PCR) was performed in 50-μL volumes containing 2 U *Taq* DNA polymerase (Takara, Otsu, Shiga, Japan), 50–100 ng of genomic DNA, 0.25 μM of each primer, 0.2 mM dNTP mix, 2 mM $MgCl_2$ and 5 μL 10× PCR buffer. PCR was carried

out on a GeneAmp® 9700 PCR System (Applied Biosystems, Carlsbad, California, USA) based on the conditions in *Ni et al. (2012)*. Amplification products were confirmed by 1.5% TBE agarose gel electrophoresis and then purified using EZ Spin Column PCR Product Purification Kit (Sangon, Shanghai, China) following described protocol. The cleaned product was prepared for sequencing using the BigDye Terminator Cycle Sequence Kit (ver. 3.1; Applied Biosystems) and finally analysed on an ABI PRISM 3730 automatic sequencer. Sequences were assembled and aligned using the DNAstar software suite (DNASTAR, Madison, Wisconsin, USA). Haplotypes were defined using the DnaSP 5 (*Librado & Rozas, 2009*), and their relationships were inferred using a maximum parsimony network in the TCS 1.21 package (*Clement, Posada & Crandall, 2000*). The best-fit model of sequence evolution was determined with jModelTest 2 (*Darriba et al., 2012*). GTR + I model was selected under the Akaike information criterion and used in subsequent analysis. Molecular diversity indices such as number of haplotypes (n), haplotype diversity (h) and nucleotide diversity (π) were calculated for each population in the ARLEQUIN 3.5 (*Excoffier & Lischer, 2010*).

A spatial analysis of molecular variance (SAMOVA) was used to define the best population grouping strategy based on F_{CT} values (number of groups K: 2 to 8) in SAMOVA 1.0 (*Dupanloup, Schneider & Excoffier, 2002*). A hierarchical analysis of molecular variance (AMOVA; *Excoffier, Smouse & Quattro, 1992*) was conducted with 10,000 permutations in ARLEQUIN 3.5 to estimate the partitioning of genetic variation. As the GTR model was not available in ARLEQUIN, the closest model *Tamura & Nei (1993)* was used. Pairwise Φ_{ST} was also calculated under this model with 1,000 random replicates followed by a standard Bonferroni correction (*Rice, 1989*). The mantel test (1,000 randomizations) for isolation by distance (IBD) was performed between genetic similarity $(F_{ST}/(1-F_{ST}))$ (*Slatkin, 1995*) and Euclidean geographical distances using IBDWS version 3.23 (*Jensen, Bohonak & Kelley, 2005*).

Historical demography of each population was investigated using Tajima' D (*Tajima, 1989*) and Fu's F_S neutrality tests (*Fu, 1997*) as implemented in ARLEQUIN 3.5 with 10,000 bootstrap replicates. Once a test yielded a value that was significantly different from zero, mismatch distribution was performed to further characterize the expansion. The sum-of-squared-differences (SSD) statistic was used to test the goodness-of-fit between the observed mismatch distribution and that expected under a sudden expansion model (10,000 replicates).

The coalescent approach implemented in IMa (*Hey & Nielsen, 2007*) was used to parameterize gene flow and divergence time among three groups (namely G1, G2 and G3, see results). Divergence time (t) was individually estimated for two group pairs G1:G2 and G1:G3 because G2 and G3 seemed to derive from G1 separately. Initial runs were analysed to determine the appropriate upper bounds for the migration (m) and divergence time (t). Twenty heated Markov chains were run with a burn-in period of 10 million steps, and all runs were consisted of 100 million steps (recording every 1,000 steps). Heating parameters ($g1 = 0.8$ and $g2 = 0.9$) were used to provide good mixing of the Markov chains. Each procedure was repeated at least for three times with different random seeds.

The analyses were considered to converge upon a stationary distribution if the independent runs reported similar posterior distributions (*Hey, 2005*) and the ESS for each run was >200 (*Won & Hey, 2005*). The mutation-scaled parameter t can be converted into the real time (T) based on following formula: $T = t \times g/u$ (g, generation time; u, mutation rate per locus per year) (*Hey & Nielsen, 2004*). For *M. chinensis*, the generation time was 1 year (*Wang et al., 1984*). However there was neither an accurate mutation rate nor a clear fossil record available for the species. Former molluscan studies using mutation rates estimated from deep splits of interspecific phylogeny were recently questioned because accelerated molecular rate estimates were suggested in short evolutionary timescales, known as the "time dependency molecular rates" hypothesis (*Ho et al., 2005*). Under the hypothesis, mutation rate can be an order of magnitude faster than that based on a phylogenetic calibration (*Grant et al., 2012*). So we adopted here a tenfold faster mutation rate of 12% myr^{-1} than the upper boundary (1.2% myr^{-1}) used in former studies (e.g., *Marko et al., 2010*; *Liu et al., 2011*) to shed light on a recent demographic scenario.

Microsatellite genotyping and analysis

Microsatellite data were analysed to validate the population structure revealed by mitochondrial COI. In our previous study, we had genotyped eight populations (DD, ZH, QH, PL, WD, HY, RZ and LY) with nine polymorphic microsatellites (*Ni, Li & Kong, 2011*). Here we screened 60 individuals in PT population with the same microsatellite loci. A detailed methodology of PCR and genotyping conditions can be found in *Ni, Li & Kong (2011)*.

The expected heterozygosity (H_E) and observed heterozygosity (H_O) were calculated for each population using the program MICROSATELLITE ANALYSER (MSA; *Dieringer & Schlotterer, 2003*), and the mean allele richness (A_R) was calculated with FSTAT version 2.9.3 (*Goudet, 2001*). *Weir & Cockerham*'s (*1984*) F_{ST} (θ) was calculated in the program MSA (*Dieringer & Schlotterer, 2003*). Significance of values was tested using 1,000 permutations with the Bonferroni correction (*Rice, 1989*). The mantel test (1,000 randomizations) for IBD was performed using the combined data set of 501 individuals in IBDWS.

We used two different kinds of analyses to assess the population structure. First, the *Cavallis-Sforza & Edwards (1967)* chord distance D_C was calculated and an unrooted neighbor-joining (NJ) tree was generated using the software POPULATIONS v1.2.31 (*Langella, 2011*). Supports for nodes were assessed by bootstrapping with 1,000 replicates. Second, population structure was inferred with a Bayesian algorithm as implemented in STRUCTURE v.2.3 (*Pritchard, Stephens & Donnelly, 2000*). Tested K ranged from 1 to 12 (sampled populations plus three). For each value, 20 replicates were run using the admixture model, correlated allele frequencies and the prior population information with a burn-in period of 10,000, followed by 100,000 steps. The most appropriate value of K was determined by the statistic ΔK introduced by *Evanno, Regnaut & Goudet (2005)* using Structure Harvester v0.6.92 (*Earl & von Holdt, 2012*). We averaged all 20 replicates for the best K by Greedy method implemented in CLUMPP (*Jakobsson & Rosenberg, 2007*), and finally visualized the results with DISTRUCT (*Rosenberg, 2004*). AMOVA analysis

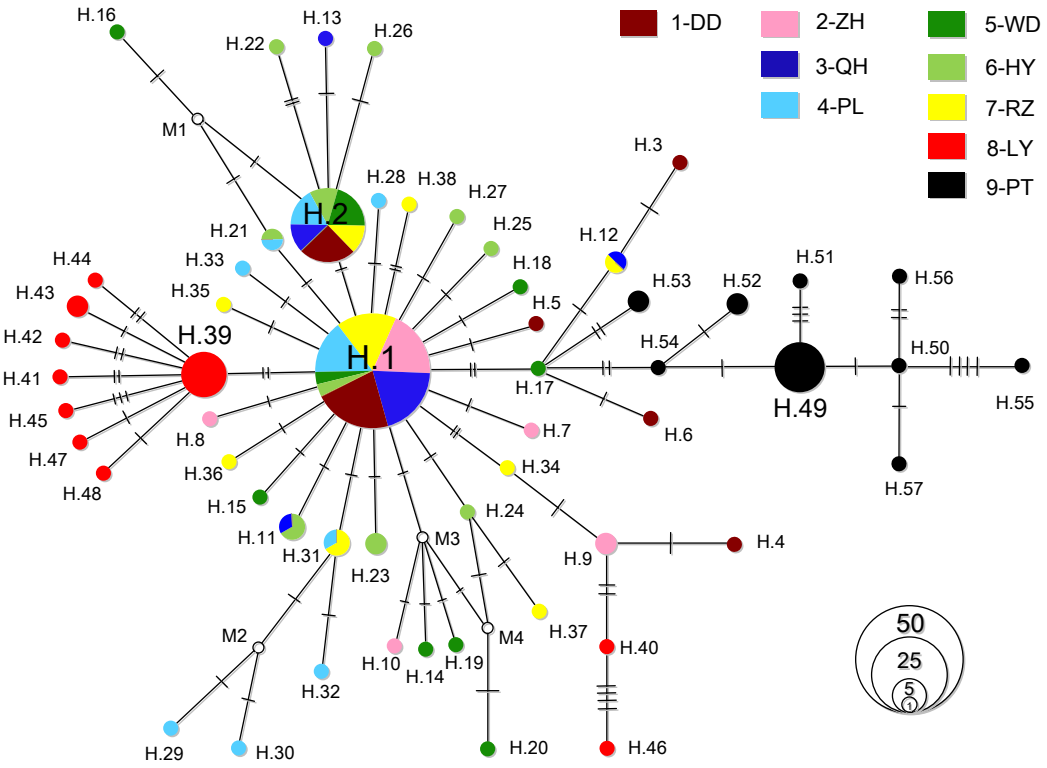


Figure 2 Haplotype network. Haplotype relationships of *M. chinensis*. Each circle represents a single haplotype sized in proportion to its frequency. Bars indicate mutation steps between haplotypes. Small empty circles (M1-4) represent missing haplotypes.

was conducted with 10,000 replicates in ARLEQUIN to check for hierarchical structure of variability.

RESULTS

Mitochondrial COI analyses

An alignment of 627 bp COI gene was analysed for the 166 individuals, with 59 variable sties yielding a total of 57 unique haplotypes (deposited in GenBank with accession numbers KC205870–KC205926). The number of haplotypes per population ranged from 5 to 10 with an average of 8. Haplotype diversity ranged from 0.533 in ZH to 0.942 in HY, and nuclear diversity ranged from 0.00148 in QH to 0.00390 in PT (Table 1). The parsimony network roughly showed a 'star-like' topology with significant geographic patterns (Fig. 2). The central haplotype (H.1) was the most common one with 59 copies (accounting for 35.5%), and appearing in seven of nine populations, it was the most likely ancestral haplotype according to *Posada & Crandall (2001)*. H.1 was separated from the other haplotypes by at most 10 mutation steps. The second haplotype H.2, appearing in six populations, was also abundant (24 copies, 14.5%). Haplotypes H.39-48, however, were spatially restricted to LY with no sharing with other populations noticed. The situation was also observed for PT with all nine haplotypes H.49-57 private (Fig. 2).

Table 2 AMOVA analyses. Results from analysis of molecular variance (AMOVA) of population structure in Mactra chinensis.

Marker	Grouping	Source of variation	*df*	Statistics	% of variation	*P* value
Mitochondrial COI	(DD, ZH, QH, PL,WD, HY, RZ); (LY); (PT)	Among groups	2	$\Phi_{CT} = 0.6574$	65.74	0.0262*
		Among populations within groups	6	$\Phi_{SC} = 0.0124$	0.43	0.0257*
		Within populations	157	$\Phi_{ST} = 0.6617$	33.83	<0.0001*
Microsatellites	(DD, ZH, QH, PL,WD, HY, RZ, LY); (PT)	Among groups	1	$F_{CT} = 0.0521$	5.21	<0.0001*
		Among populations within groups	7	$F_{SC} = 0.0234$	2.22	<0.0001*
		Within populations	993	$F_{ST} = 0.0743$	92.57	<0.0001*

Notes.
* significant at $P < 0.05$.

Table 3 Population comparison. Pairwise Φ_{ST} (COI, below diagonal) and F_{ST} (θ) values (microsatellites, above diagonal) between the *M. chinensis* populations.

Sites	DD	ZH	QH	PL	WD	HY	RZ	LY	PT
DD	—	0.010*	0.012*	0.018*	0.016*	0.027*	0.024*	0.037*	0.056*
ZH	0.041	—	0.014*	0.009*	0.017*	0.018*	0.022*	0.037*	0.067*
QH	−0.031	0.058	—	0.013*	0.021*	0.024*	0.036*	0.038*	0.066*
PL	0.021	0.061	0.012	—	0.005	0.012*	0.016*	0.025*	0.073*
WD	0.016	0.104*	0.026	0.051	—	0.012*	0.008*	0.026*	0.079*
HY	0.008	0.084*	−0.003	0.010	−0.001	—	0.015*	0.020*	0.085*
RZ	−0.013	0.013	−0.020	−0.008	0.039	0.011	—	0.024*	0.078*
LY	0.529*	0.531*	0.559*	0.532*	0.484*	0.498*	0.511*	—	0.085*
PT	0.680*	0.686*	0.706*	0.697*	0.616*	0.667*	0.674*	0.726*	—

Notes.
* Indicates significant difference after Bonferroni correction ($P < 0.05/36$).

For SAMOVA analysis, the grouping for which the F_{CT} was highest was chosen as the optimal grouping method. Here the highest F_{CT} appeared at three groups ((DD, ZH, QH, PL, WD, HY, RZ); (LY); (PT), referred to hereafter as G1, G2 and G3 correspondingly). Based on this strategy, hierarchical analyses of AMOVA revealed a significant level of geographic structuring among groups ($\Phi_{CT} = 0.6574$, $P = 0.0262$; Table 2). Genetic differences among groups explained 65.74% of the total variation, followed by variation within populations of 33.83% ($\Phi_{ST} = 0.6617$). Only 0.43% of the variation was attributed to variation among populations within groups ($\Phi_{SC} = 0.0124$). These results were corroborated with the pairwise Φ_{ST} values: genetic structure was indicated as 17 comparisons were significant after Bonferroni corrections (Table 3). LY and PT showed remarkably differentiation from each other and all other sites with rather large Φ_{ST} (0.484–0.726 and 0.616–0.726, respectively). The IBD analysis suggested significantly positive correlation between genetic and geographic distance matrices ($R^2 = 0.394$, $P = 0.004$; Fig. 3A).

Only two populations ZH and PT showed no signal of historical expansion in both Tajima' *D* and Fu' F_S tests ($P > 0.05$, Table 4). For the rest, at least one of the tests yielded significantly negative value, indicating each population had experienced a demographical

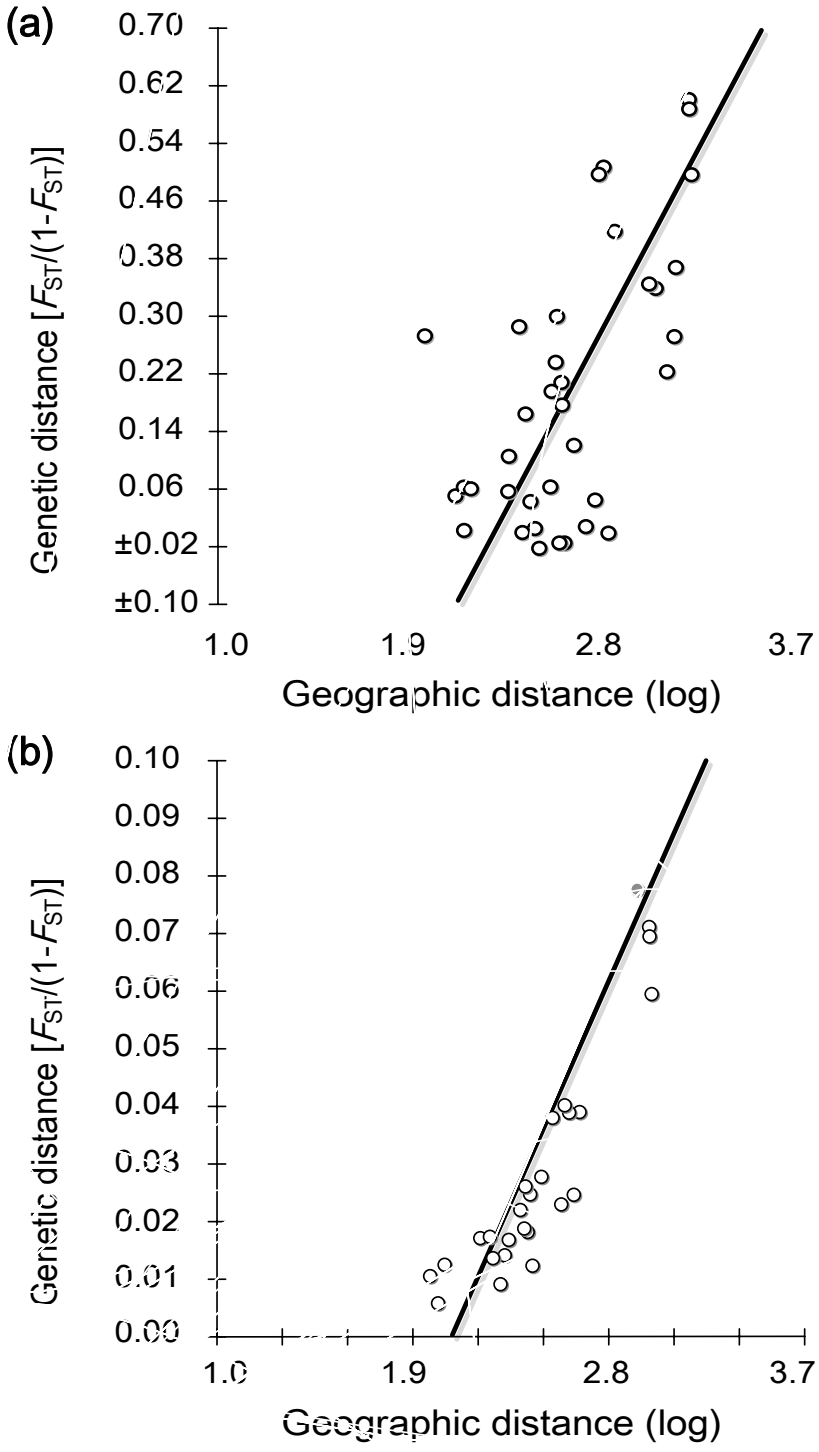


Figure 3 Isolation by distance analyses for *M. chinensis* populations. (A) Relationship between geographic (log-transformed) and COI genetic distances. (B) Relationship between geographic (log-transformed) and microsatellite genetic distances.

Table 4 Demongraphy analyses. Estimates of neutral tests (Tajima' *D* and Fu' F_S) for each population and the mismatch distribution parameter SSD.

Population	Tajima' *D*		Fu' F_S		Goodness-of-fit test	
	D	*P*	F_S	*P*	SSD	*P* (Sim. Ssd ≥ Obs. Ssd)
1. DD	−1.7792	0.0187	−0.9806	0.2718	0.0383	0.05700
2. ZH	−1.5192	0.0556	−0.9669	0.2009	—	—
3. QH	−1.5531	0.0463	−1.4663	0.0975	0.0043	0.7510
4. PL	−1.4943	0.0572	−4.7706	0.0006	0.0083	0.3840
5. WD	−1.5354	0.0511	−4.1843	0.0053	0.0115	0.4300
6. HY	−1.4488	0.0695	−6.4180	0.0000	0.0438	0.0740
7. RZ	−2.0852	0.0064	−4.4322	0.0016	0.0011	0.9000
8. LY	−2.2006	0.0024	−4.4153	0.0058	0.0024	0.9540
9. PT	−1.4197	0.0660	−2.5527	0.0721	—	—

expansion under the neutral model (*Rogers & Harpending, 1992*). The subsequent goodness-of-fit tests for them also supported the null hypothesis of sudden expansion model with nonsignificant values for SSD (all $P > 0.05$, Table 4).

Independent runs of IMa2 converged on similar marginal posterior probability distributions. Plots of the probability for the migration parameters of pairwise groups were uniformly unambiguous and very narrow: each of them was very close in position and height to the lowest migration value in the histogram (Fig. 4A). While it was theoretically possible that a nonzero peak might be found with a finer resolution, these estimates were considered effectively zero (see also *Won & Hey, 2005*; *Dixon, Kapralov & Filatov, 2011*). The mean time estimate with statistical confidence (95% highest posterior density) for t_1 (G1:G2) was 1.610 (0.908–2.768), while t_2 (G1:G3) was 1.354 (0.568–3.328). When applying a mutation rate of 12% myr^{-1}, the real time T_1 amounted to 21.4 (12.1–36.8) and T_2 was 18.0 (7.5–44.2) kyr (Fig. 4B), roughly corresponding to the LGM happening about 19–22 kyr ago (*Yokoyama et al., 2001*).

Microsatellite-based population structure

For microsatellites, genetic diversity indices for each population were summarized in Table 1. The mean allele richness varied from 14.27 in PT to 17.63 in PL, and the observed and expected heterozygosities ranged from 0.81 to 0.90 and from 0.48 to 0.67, respectively. All pairwise F_{ST} tests except one were significant after Bonferroni corrections (Table 3). Extremely high F_{ST} values ($\theta = 0.056$–0.085) were observed between PT and all other populations, suggesting significant genetic differentiation between them.

Significant positive correlation was detected between genetic and geographic distance matrices for the nine populations ($R^2 = 0.705$, $P = 0.001$; Fig. 3B). Population structure revealed by two different analyses displayed a similar pattern as two genetically different groups were observed. According to the NJ tree, the populations were divided into two groups: one consisting of the southern PT, and the other including all other populations

Figure 4 IMa analyses. (A) Gene flow and (B) divergence time among the three groups (G1, G2 and G3) estimated using coalescent approach IMa.

northern of the estuary (Fig. 5). For the STRUCTURE analyses, the highest ΔK in Structure Harvester was found for $K = 2$ (Fig. 6A). At this value, the result indicated PT dominated by the purple cluster was significantly differentiated from all other sites (Fig. 6B). At $K = 3$, although no additional population group was detected, two clusters (namely pink and green clusters) were recovered with clear gradations in allele frequencies across the northern populations (Figs. 6B and 6C).

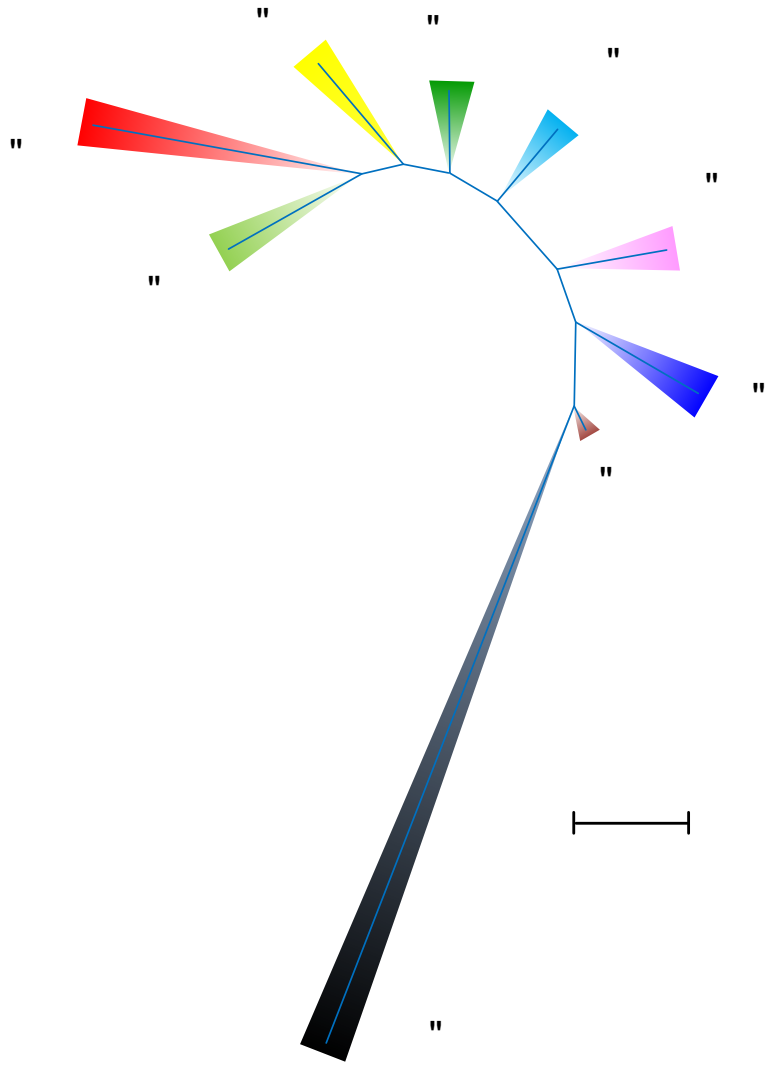


Figure 5 Neighbor-joining tree using microsatellites. Unrooted neighbor-joining tree based on microsatellite D_C distances among the nine populations. Numbers on braches indicate bootstrap support values.

DISCUSSION

In this study, a star-like haplotype topology with shallow divergence was recovered for COI haplotypes of the clam *M. chinensis*, suggesting the single origin of current populations from a common one. Ancestral individuals might have retreated into the ECS refugium when sea levels declined during periods of glaciations, forming a panmictic population. When sea levels rose as glaciers melted after the LGM, the survived individuals migrated out the basin (refugium) and repopulated along the ECS coastline quickly. This might represent a common evolutionary scenario for marine species inhabiting the ECS, and similar patterns were revealed in diverse organisms including fishes, molluscs and crustaceans (*Ni et al., 2014*). A caveat in this study was that the mitochondrial data (14–23 sequences) for each population had the potential of missing some low

Figure 6 STRUCTURE analyses. Microsatellite-based population structure analyses in STRUCTURE for *M. chinensis*. (A) Estimating the true number of clusters with ΔK. Here the uppermost level appears at $K = 2$; (B) the population structure of *M. chinensis* at $K = 2$, 3 respectively; (C) when $K = 3$, the percentage of individuals allocated to each cluster with membership >80% in each population, showing latitude gradations in allele frequencies.

frequency haplotypes. Although no divergent lineages were inferred for *M. chinensis*, substantial population subdivision was detected among populations using both COI and microsatellite markers, with somewhat different results: three groups of populations were defined in SAMOVA analysis of mitochondrial COI and a significant level of genetic structure was revealed among them with 65.7% of the total variation explained. Based on "time dependency molecular rates" hypothesis, coalescence analyses suggested the divergence between two pair groups happened about 21.4 and 18.0 kyr ago, respectively, with a close link to the LGM; two groups were supported by microsatellite analyses and the AMOVA result identified a significant among-group component, explaining 5.21% of the total variance (Table 2).

Population subdivision across the Changjiang freshwater boundary

The most striking result was the significant divergence of PT in the south of Changjiang estuary from the northern ones, congruent with the 'biogeography and phylogeography concordance' hypothesis. The mitochondrial data showed that no haplotypes were shared

between them and significant differentiation was revealed in pairwise Φ_{ST} analyses. The microsatellite results also supported their division in both F_{ST} and population structure analyses (the NJ tree and STRUCTURE). This sharp genetic discontinuity across the estuary most likely results from the influence of the biogeographic boundary associated with habitat discontinuity: several ocean currents and the Changjiang freshwater outflow meet around the estuary, causing striking physical and ecological gradients that may potentially limit gene flow (*Su & Yuan, 2005*); additionally, the PT site is geographically distant from other populations with the shortest straight-line distance $>$1,000 km, and there lacks stepping-stone populations (*Xu, 1997*), which may amplify the barrier effect. The similar effect was also reported for other sympatric species, such as gastropod *Cellana toreuma* (*Dong et al., 2012*) and two varieties of Sargassum (*Cheang, Chu & Ang, 2008*). However, it is not consistent across all species as no genetic break linked to the freshwater boundary was noticed for bivalves *Atrina pectinata* (*Liu et al., 2011*) and *Cyclina sinensis* (*Ni et al., 2012*). Difference in life-history characteristics, habitat requirement and historical distribution may be responsible for the discordance (*Ni et al., 2012*).

Population substructure of the northern group

Besides the divergence of PT, population subdivision was also observed among sites on the same side of the estuary. For the northern populations, the mitochondrial data revealed remarkable differentiation between LY and the rest of the populations both in phylogenetic and pairwise Φ_{ST} analyses, and all of its haplotypes were private. It was an interesting pattern as LY was only $\sim$50 km apart from the nearest population RZ, without known

large proportion of 48.1% in HY (Yates' chi-squared test: $\chi^2 = 12.7369, P = 0.00036$, corrected for continuity).

CONCLUSIONS

Using the surf clam *M. chinensis* as a case study, our results indicate that, although derived from a single ancestral population, significant subdivision can arise for intertidal species within the ECS. Genetic discontinuity was observed between distant populations across the Changjiang Estuary but also between adjacent northern populations, reflecting the influence of multiple driving forces including the Changjiang freshwater boundary, habitat discontinuity, environmental gradient factors and life-history traits on shaping the phylogeographic pattern of *M. chinensis*. Although phylogeographic studies on the NW Pacific have been increasing in recent years, the emerging picture is still far from complete, especially regarding the detailed within-sea evolutionary process (*Ni et al., 2014*). This study complemented our understanding of possible origins of marine biodiversity in the northwestern Pacific, and highlight the contribution of multiple ecological factors during the process.

ACKNOWLEDGEMENTS

We are indebted to Dr. Jun Chen from Linyi University and Dr. Hongtao Nie from Dalian Ocean University for their kind assistance in specimen collection.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by research grants from National Natural Science Foundation of China (41276138), National Marine Public Welfare Research Program (201305005), and Doctoral Program of Ministry of Education of China (20130132110009). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Natural Science Foundation of China: 41276138.
National Marine Public Welfare Research Program: 201305005.
Doctoral Program of Ministry of Education of China: 20130132110009.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Gang Ni conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

- Qi Li conceived and designed the experiments, wrote the paper, reviewed drafts of the paper.
- Lehai Ni performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Lingfeng Kong conceived and designed the experiments, analyzed the data, reviewed drafts of the paper.
- Hong Yu conceived and designed the experiments, performed the experiments, reviewed drafts of the paper.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:
DNA sequences were deposited in GenBank, accessions numbers KC205870–C205926;
Microsatellite loci were deposited in GenBank, accession number for Mc1: GQ499856; Mc2: GQ499857; Mc9: GQ499861; Mc10: GQ499862; Mc16: GQ499864; Mc21: GQ499866; Mc24: GQ499868; Mc28: GQ499872; Mc30: GQ499873.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.1240#supplemental-information.

REFERENCES

Brante A, Fernández M, Viard F. 2012. Phylogeography and biogeography concordance in the marine gastropod *Crepipatella dilatata* (Calyptraeidae) along the southeastern Pacific coast. *Journal of Heredity* **103**:630–637 DOI 10.1093/jhered/ess030.

Cavallis-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* **21**:550–570 DOI 10.2307/2406616.

Cheang CC, Chu KH, Ang Jr PO. 2008. Morphological and genetic variation in the populations of *Sargassum hemiphyllum* (Phaeophyceae) in the northwestern Pacific. *Journal of Phycology* **44**:855–865 DOI 10.1111/j.1529-8817.2008.00532.x.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**:1657–1659 DOI 10.1046/j.1365-294x.2000.01020.x.

Conkright ME, Locarnini RA, Garcia HE, O'Brien TD, Boyer TP, Stephens C, Antonov JI. 2001. *World Ocean Atlas 2001*. Washington, D.C.: National Oceanic and Atmospheric Administration.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**:772–772 DOI 10.1038/nmeth.2109.

Dieringer D, Schlotterer C. 2003. MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**:167–169 DOI 10.1046/j.1471-8286.2003.00351.x.

Dixon C, Kapralov M, Filatov D. 2011. Gene flow and species cohesion following the spread of *Schiedea globosa* (Caryophyllaceae) across the Hawaiian Islands. *Journal of Evolutionary Biology* **24**:1–11 DOI 10.1111/j.1420-9101.2010.02128.x.

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 DOI 10.1371/journal.pone.0036178.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**:2571–2581 DOI 10.1046/j.1365-294X.2002.01650.x.

Earl DA, Von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**:359–361 DOI 10.1007/s12686-011-9548-7.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**:2611–2620 DOI 10.1111/j.1365-294X.2005.02553.x.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**:564–567 DOI 10.1111/j.1755-0998.2010.02847.x.

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 DOI 10.1111/j.1755-0998.2010.02847.x.

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**:294–299.

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

Gilg MR, Hilbish TJ. 2003. The geography of marine larval dispersal: coupling genetics with fine-scale physical oceanography. *Ecology* **84**:2989–2998 DOI 10.1890/02-0498.

Goudet J. 2001. *FSTAT, a program to estimate and test gene diversities and fixation indices*. version 2.9.3. *Available at www.unil.ch/izea/softwares/fstat.html*.

Grant WS, Liu M, Gao TX, Yanagimoto T. 2012. Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographies in Pacific herring (and other species). *Molecular Phylogenetics and Evolution* **65**:203–212 DOI 10.1016/j.ympev.2012.06.006.

Hellberg ME. 2009. Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics* **40**:291–310 DOI 10.1146/annurev.ecolsys.110308.120223.

Hey J. 2005. On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology* **3**:965–975 DOI 10.1371/journal.pbio.0030193.

Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**:747–760 DOI 10.1534/genetics.103.024182.

Hey J, Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America* **104**:2785–2790 DOI 10.1073/pnas.0611164104.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* **22**:1561–1568 DOI 10.1093/molbev/msi145.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**:1801–1806 DOI 10.1093/bioinformatics/btm233.

Jensen J, Bohonak A, Kelley S. 2005. Isolation by distance web service. *BMC Genetics* **6**:13 DOI 10.1186/1471-2156-6-13.

Jiang L, Wu WL, Huang PC. 1995. The mitochondrial DNA of taiwan abalone *Haliotis diversicolor* Reeve, 1846 (Gastropoda: Archaeogastropoda: Haliotidae). *Molecular Marine Biology and Biotechnology* **4**:353–364.

Kong LF, Li Q. 2009. Genetic evidence for the existence of cryptic species in an endangered clam *Coelomactra antiquata*. *Marine Biology* **156**:1507–1515 DOI 10.1007/s00227-009-1190-5.

Kong LF, Li Q, Qiu ZX. 2007. Genetic and morphological differentiation in the clam *Coelomactra antiquata* (Bivalvia: Veneroida) along the coast of China. *Journal of Experimental Marine Biology and Ecology* **343**:110–117 DOI 10.1016/j.jembe.2006.12.003.

Langella O. 2011. POPULATIONS 1.2.30. *Available at http://bioinformatics.org/~tryphon/populations*.

Li Q, Park C, Kijima A. 2002. Isolation and characterization of microsatellite loci in the Pacific abalone, *Haliotis discus hannai*. *Journal of Shellfish Research* **21**:811–816.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451–1452 DOI 10.1093/bioinformatics/btp187.

Liu XQ, Fang JG, Bao ZM, Wang RC. 2007. A preliminary study on the reproductive biology of *Mactra chinensis* Philippi. *Periodical of Ocean University of China* **37**:89–92.

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). *Molecular Ecology* **16**:275–288 DOI 10.1111/j.1365-294X.2006.03140.x.

Liu H, He Q, Wang Z, Weltje GJ, Zhang J. 2010. Dynamics and spatial variability of near-bottom sediment exchange in the Yangtze Estuary, China. *Estuarine, Coastal and Shelf Science* **86**:322–330 DOI 10.1016/j.ecss.2009.04.020.

Liu J, Li Q, Kong LF, Zheng XD. 2011. Cryptic diversity in the pen shell *Atrina pectinata* (Bivalvia: Pinnidae): high divergence and hybridization revealed by molecular and morphological data. *Molecular Ecology* **20**:4332–4345 DOI 10.1111/j.1365-294X.2011.05275.x.

Marko PB, Hoffman JM, Emme SA, Mcgovern TM, Keever CC, Nicole C. 2010. The 'Expansion-Contraction' model of Pleistocene biogeography: rocky shores suffer a sea change? *Molecular Ecology* **19**:146–169 DOI 10.1111/j.1365-294X.2009.04417.x.

Mayr E. 1954. Geographic speciation in tropical echinoids. *Evolution* **8**:1–18 DOI 10.2307/2405661.

McManus J. 1985. Marine speciation, tectonics and sea-level changes in southeast Asia. In: *Proceedings of the Fifth International Coral Reef Congress, Tahiti*, vol. 4. 133–138.

Miglietta MP, Faucci A, Santini F. 2011. Speciation in the sea: overview of the symposium and discussion of future directions. *Integrative and Comparative Biology* **51**:449–455 DOI 10.1093/icb/icr024.

Ni LH, Li Q, Kong LF. 2011. Microsatellites reveal fine-scale genetic structure of the Chinese surf clam *Mactra chinensis* (Mollusca, Bivalvia, Mactridae) in Northern China. *Marine Ecology* **32**:488–497 DOI 10.1111/j.1439-0485.2011.00436.x.

Ni G, Li Q, Kong L, Yu H. 2014. Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology* **23**:534–548 DOI 10.1111/mec.12620.

Ni G, Li Q, Kong LF, Zheng XD. 2012. Phylogeography of bivalve *Cyclina sinensis*: testing the historical glaciations and Changjiang River outflow hypotheses in northwestern Pacific. *PLoS ONE* 7:e49487 DOI 10.1371/journal.pone.0049487.

Palumbi SR. 1992. Marine speciation on a small planet. *Trends in Ecology & Evolution* 7:114–118 DOI 10.1016/0169-5347(92)90144-Z.

Pedreschi D, Kelly-Quinn M, Caffrey J, O'Grady M, Mariani S. 2014. Genetic structure of pike (*Esox lucius*) reveals a complex and previously unrecognized colonization history of Ireland. *Journal of Biogeography* **41**:548–560 DOI 10.1111/jbi.12220.

Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**:37–45 DOI 10.1016/S0169-5347(00)02026-7.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959.

Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* **43**:223–225 DOI 10.2307/2409177.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**:552–569.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**:137–138 DOI 10.1046/j.1471-8286.2003.00566.x.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**:457–462.

Su JL, Yuan YL. 2005. *Coastal hydrology of China*. Beijing: Ocean Press. (in Chinese).

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585–595.

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**:512–526.

Taylor MS, Hellberg ME. 2003. Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science* **299**:107–109 DOI 10.1126/science.1079365.

Wang PX. 1999. Response of western Pacific marginal seas to glacial cycles: paleoceanographic and sedimentological features. *Marine Geology* **156**:5–39 DOI 10.1016/S0025-3227(98)00172-8.

Wang P, Bradshaw M, Ganzei S, Tsukawaki S, Hanssan K, Hantoro W, Poobrasert S, Burne R, Zhao Q, Kagami H. 1997. West Pacific marginal seas during last glacial maximum: amplification of environmental signals and its impact on monsoon climate. In: Wang B, ed. *Proceedings of 30th International Geological Congress: VSP, 13*. Utrecht, The Netherlands, 65–86.

Wang ZC, Liu JM, Zhu A, Li WJ, Shen YC. 1984. A preliminary study on the biology of the Chinese surf clam (*Mactra chinensis*) in Yalu River Estuary. *Journal of Fisheries of China* **8**:33–44 (in Chinese).

Wang RC, Wang ZP. 2008. *Science of marine shellfish culture*. Qingdao: China Ocean University Press.

Wang B, Wang X, Zhan R. 2003. Nutrient conditions in the Yellow Sea and the East China Sea. *Estuarine, Coastal and Shelf Science* **58**:127–136 DOI 10.1016/S0272-7714(03)00067-2.

Weersing K, Toonen RJ. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* **393**:1–12 DOI 10.3354/meps08287.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358–1370 DOI 10.2307/2408641.

Won YJ, Hey J. 2005. Divergence population genetics of chimpanzees. *Molecular Biology and Evolution* **22**:297–307 DOI 10.1093/molbev/msi017.

Xie C, Jian Z, Zhao Q. 1995. *Paleogeographic maps of the China Seas at the last glacial maximum WESTPAC Paleogeographic Maps*. Shanghai: UNESCO/IOC Publications. pp 75.

Xu FS. 1997. *Bivalve Mollusca of China Seas*. Beijing: Science Press (in Chinese).

Xu JW, 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. *Molecular Phylogenetics and Evolution* **52**:45–56 DOI 10.1016/j.ympev.2009.02.007.

Xu X, Oda M. 1999. Surface-water evolution of the eastern East China Sea during the last 36,000 years. *Marine Geology* **156**:285–304 DOI 10.1016/S0025-3227(98)00183-2.

Xu FS, Zhang SP. 2008. *An illustrated Bivalvia Mollusca fauna of China Seas*. Beijing: Science Press.

Yokoyama Y, Lambeck K, De Deckker P, Johnston P, Fifield IK. 2001. Timing of the Last Glacial Maximum from observed sea-level minima. *Nature* **406**:713–716 DOI 10.1038/35021035.

Zhang X, Qi ZY, Zhang FS, Ma XT. 1963. A preliminary study of the demarcation of marine molluscan faunal regions of and its adjacent waters. *Oceanologia et Limnologia Sinica* **5**:124–138 (in Chinese with English abstract).

Zhao Y, Li Q, Kong L, Bao Z, Wang R. 2007. Genetic diversity and divergence among clam *Cyclina sinensis* populations assessed using amplified fragment length polymorphism. *Fish Science* **73**:1338–1343.