The Clam *Cyclina sinensis* (Gmelin) Phylogeography Study with 28S rRNA Gene and Potential of Nuclear rRNA Genes in Genetic Assessments of Molluscs

NI Gang[1], LI Qi[1), 2), 3), *], KONG Lingfeng[1), 2)], and YU Hong[1), 2)]

1) *The Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao* 266003, *China*
2) *Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao* 266003, *China*
3) *Laboratory of Tropical Marine Germplasm Resources and Breeding Engineering, Sanya Oceanographic Institution, Ocean University of China, Sanya* 572000, *China*

Abstract Intraspecific diversity of molluscan species is usually studied based on maternally inherited mitochondrial DNA, from which only part of the evolutionary history can be reflected. Some nuclear ribosomal RNA genes such as 28S rRNA represent potential candidates that can be easily applied in phylogeography because of lacking intraindividual variation. However, considering their low polymorphism, genetic appraisals on whether and how they can be used in population studies are necessary. Here, we applied a short 28S rRNA to assess genetic patterns of the clam *Cyclina sinensis* along the coast of China and compared the results with a former study based on COI and ITS-1 analyses. The results revealed the 28S rRNA data set was characterized by an extremely low level of variation, with only seven haplotypes defined for 93 individuals. Haplotype and nucleotide diversity for each population was almost the lowest when compared with the other two markers. However, the distribution of two dominant haplotypes showed clear geographic patterns, and significant population differentiation was revealed between the East China Sea and the South China Sea. These patterns were highly concordant with findings of the former study that populations of *C. sinensis* were historically separated by land bridges among sea basins. Our study suggested that although the nuclear rRNAs have shortcomings such as low variation, they have advantages including lack of intraindividual variation and high amplification rates. Applying rRNA genes can enrich the toolbox of nuclear markers in molluscan phylogeographic studies.

Key words marine phylogeography; molluscs; nuclear gene; genetic break

1 Introduction

The basic level of marine biodiversity is genetic diversity, which provides a natural variation of the raw material of evolution and ensures the survival of populations withstanding environmental changes (May and Godfrey, 1994). Studies on genetic diversity can provide valuable biological and evolutionary information that is essential for the successful conservation and/or management of marine species (Xiao *et al.*, 2009).

Phylogeography is a powerful tool for using genetic data to understand intraspecific diversity patterns both in space and time (Avise, 2009). In recent decades, massive studies have been carried out in diverse taxonomic groups (see Hickerson *et al.*, 2010; Shafer *et al.*, 2010; Turchetto-Zolet *et al.*, 2013). However, research effort on marine species, especially on the largest marine phylum Mollusca, is still not comparable to that on terrestrial and freshwater groups (Beheregaray, 2008). Moreover, the previous studies were logeography is a genomic scan based on next-generation sequencing, which can offer extremely high resolution on population differentiation (Emerson *et al.*, 2010). However, in practice, the cost for sequencing hundreds of specimens, which is typically required in phylogeographic studies, is a heavy burden for small research projects. Traditional nuclear markers like ribosomal RNA (rRNA) genes are still useful in phylogeographic studies, especially for

---

* Corresponding author. Tel: 0086-532-82031622
  E-mail: qili66@ouc.edu.cn

numerous marine invertebrates that have not been studied thoroughly. The rRNA genes tend to evolve in a concerted way, which is known as 'concerted evolution' that result in nucleotide homogeneity among multigene families (Dover, 1982). Specific rRNA regions can be selectively amplified by universal markers followed by direct DNA sequencing (Freeland *et al.*, 2011). Compared with other nuclear markers such as microsatellites, this approach does not require species-specific development for species that are never examined before (Zane *et al.*, 2002). The transferable markers also facilitate comparisons of evolutionary patterns among different species (Schlötterer, 2004). However, to employ this method, it is important to know whether a conserved nuclear gene comprises sufficient polymorphic sites to reflect the evolutionary history of a species.

In this study, we performed a phylogeographic study on the bivalve *Cyclina sinensis* (Gmelin) using a partial 28S rRNA gene. *C. sinensis* is a common bivalve with a wide distribution in China and Japan, and is usually found from the muddy sand bottom of the intertidal zone (Qi, 2004). This species has a short planktonic larval duration of about 7 days on average in laboratory conditions (Zeng and Li, 1991), indicating there is limited gene flow among distant populations. The indication was verified by a former study on populations along the coast of China (Ni *et al.*, 2012), in which significant population differentiation was revealed by both COI and ITS-1 genes between populations of the East China Sea (ECS) and the South China Sea (SCS). This background provides a chance for us to compare the performance of the 28S rRNA gene in illustrating phylogeographic break within the *C. sinensis* populations. According to our results, we also discussed the potential and pitfalls of using rRNA genes in molluscan phylogeographic studies.

2 Materials and Methods

2.1 Sample Collection

Specimens of *C. sinensis* used in this study are a subset of populations from a former study of Ni *et al.* (2012). They are selected to represent both the ECS and SCS lineages according to former results. A total of 93 individuals from 11 localities were used here, with a range of 5–10 individuals per site (Fig.1 and Table 1). Samples were
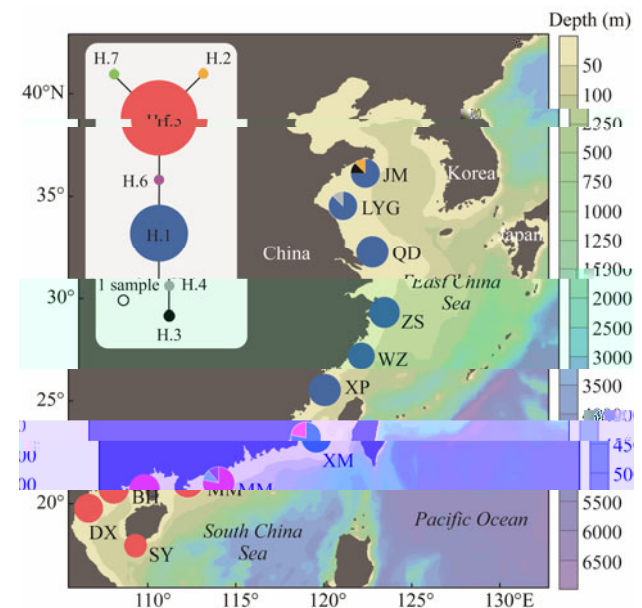

Fig.1 Sampling localities for the clam *C. sinensis* along the coast of China and the distribution of haplotypes in each population. Haplotype network is shown (top left) with the sizes of circles being proportional to their frequencies.

Table 1 Sampling information and diversity indices for the 11 populations of *C. sinensis*

Sampling site	Latitude, longitude	Individuals	Haplotypes	Haplotype diversity (h)	Nucleotide diversity (π)
Jimo (JM)	36°20′N, 120°43′E	8	3	0.464	0.004209
Lianyungang (LYG)	34°37′N, 119°31′E	8	2	0.250	0.000847
Qidong (QD)	31°52′N, 121°40′E	10	1	0.000	0.000000
Zhoushan (ZS)	30°01′N, 122°04′E	9	1	0.000	0.000000
Wenzhou (WZ)	27°57′N, 120°54′E	7	1	0.000	0.000000
Xiapu (XP)	26°50′N, 120°02′E	10	1	0.000	0.000000
Xiamen (XM)	24°25′N, 118°12′E	9	2	0.389	0.002619
Maoming (MM)	21°29′N, 111°06′E	10	3	0.378	0.001347
Beihai (BH)	21°25′N, 109°07′E	9	1	0.000	0.000000
Dongxing (DX)	21°27′N, 107°58′E	8	1	0.000	0.000000
Sanya (SY)	18°13′N, 109°38′E	5	1	0.000	0.000000

collected from October 2005 to January 2010 from public access areas without specific permit. The adductor muscle was either incised and stored in 95% ethanol or frozen at −30℃ immediately for DNA extraction in the near future. Total genomic DNA was extracted from about 50 mg muscle tissue following a phenol-chloroform purification procedure described by Li *et al.* (2002).

2.2 Sequence Acquisition

A fragment of the 28S rRNA gene was amplified using

the primer pair 28SMIDF (5'-CTTGAAACACGGACCA AGG-3' forward) and 28SD6R (5'-CCAGCTATCCTGAG GGAAACTTCG-3' reverse) (Mikkelsen *et al.*, 2006). Each polymerase chain reaction (PCR) was carried out in 50-μL volumes including 2 U Taq DNA polymerase (Takara Co.), about 100 ng template DNA, 0.25 μmol L^{-1} of each primer, 0.2 mmol L^{-1} dNTPs, 1×PCR buffer, and 2 mmol L^{-1} MgCl$_2$. The PCR amplification was conducted on a GeneAmp® 9700 PCR System (Applied Biosystems), and the cycling conditions were as follows: initial denaturation

at 94℃ for 3 min, followed by 35 cycles of denaturation at 94℃ for 1 min, annealing at 56℃ for 1 min, and extension at 72℃ for 1 min, with a final extension at 72℃ for 5 min. The products were checked using 1.5% TBE agarose gel electrophoresis stained with ethidium bromide. The target fragments were purified with EZ Spin Column PCR Product Purification Kit (Sangon) and sequenced using Big-Dye chemistry (Applied Biosystems Inc., USA) on an ABI PRISM 3730 (Applied Biosystems) automatic sequencer. We did not pursue cloning of the gene because multiple peaks indicating the intraindividual variations were not observed in all specimens.

2.3 Data Analyses

Sequences were edited and aligned using the DNA-STAR software (DNASTAR, Inc.) and then refined manually. Haplotypes were defined using the DnaSP 5 under the setting of sites with gaps considered (Librado and Rozas, 2009), and deposited in GenBank with accession numbers HQ881576-HQ881582. Their phylogeny was inferred using a maximum parsimony network in the TCS 1.21 package (Clement *et al.*, 2000). The gaps were treated as a fifth state because we think that haplotypes separated by an indel event are different. We used ARLEQUIN 3.5 (Excoffier and Lischer, 2010) to calculate molecular diversity indices including haplotype diversity (h) and nucleotide diversity (π). Spatial analysis of molecular variance (SA-MOVA) implemented in the software SAMOVA 1.0 (Dupanloup *et al.*, 2002) was performed to define groups of populations (K) that were spatially homogeneous. The group number was tested from $K=2$ to $K=10$ with 100 simulated annealing processes. Partitioning of the genetic variation was accomplished by a hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) in ARLEQUIN 3.5 with 10000 permutations. Pairwise Φ_{ST} was also calculated in the same software under the setting of pairwise difference with 10000 permutations followed by a standard Bonferroni correction (Rice, 1989).

3 Results

We successfully amplified and sequenced the partial 28S rRNA gene for all 93 individuals. As predicted, the data set was characterized by an extremely low level of sequence variation. The final alignment was 297 bp long, including two polymorphic sites and four gaps, defining a total of seven haplotypes. Seven of the 11 populations were dominated by only one haplotype, and haplotype and nucleotide diversities for each of them were zero (Table 1). The rest populations had two or three haplotypes. JM showed the highest variation with haplotype diversity (h) of 0.464 and nucleotide diversity (π) of 0.004209, while the LYG was the lowest with $h=0.250$ and $\pi=0.000847$ (Table 1). The parsimony network of the haplotypes showed a simple topology with shallow genetic divergence. The haplotype H.1, as the most abundant one with 56 copies (accounting for 60.2%), showed significant geographic distribution in seven ECS populations. The second common haplotype H.5 had 32 copies (34.4%), with major distribution in four SCS populations, but also appeared in a neighbouring ECS population (XM) with 2 copies (Fig.1).

For SAMOVA analysis, the method determined the optimal grouping K, for which the among-group component (F_{CT}) of the total genetic variance was the highest. In this study, the highest value was observed at two groups, group 1 including JM, LYG, QD, ZS, WZ, XP and XM, and group 2 including MM, BH, DX and SY. Under this grouping strategy, hierarchical analyses of AMOVA indicated a significant level of differentiation among groups ($\Phi_{CT}=0.873$, $P=0.003$; Table 2), explaining 87.32% of the overall genetic variation. Variation within populations accounted for 12.29% ($\Phi_{ST}=0.877$), while variation among populations within groups ($\Phi_{SC}=0.031$) only accounted for 0.39% of the total variation. Pairwise analyses revealed uniformly large and significant Φ_{ST} values between populations in different seas after a standard Bonferroni correction, while none of the values between populations within the ECS or SCS was significant (Fig.2).

Table 2 Result of analysis of molecular variance (AMOVA)

	Grouping	Source of variation	*d.f.*	Φ-statistics	% of variation	*P* value
Group 1	JM, LYG, QD, ZS, WZ, XP, XM	Among groups	1	$\Phi_{CT}=0.873$	87.32	**0.003**
Group 2	MM, BH, DX, SY	Among populations within groups	9	$\Phi_{SC}=0.031$	0.39	0.145
		Within populations	82	$\Phi_{ST}=0.877$	12.29	**0.000**

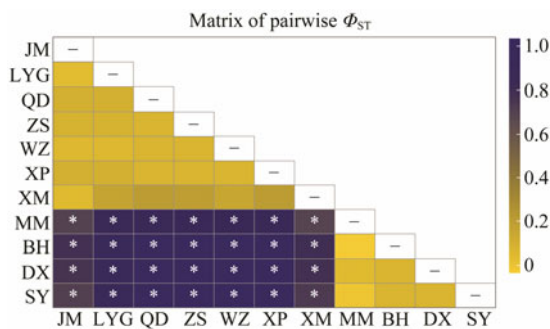Note: Significant differences (*P*-value < 0.05) are indicated in bold.

Fig.2 Heatmap visualizing pairwise Φ_{ST} values between the *C. sinensis* populations. Asterisk (*) indicates the significant *P* value after a standard Bonferroni correction.

4 Discussion

4.1 Low Variation of the 28S rRNA Gene in *C. sinensis* Populations

C. sinensis is one of the most widely distributed bivalves in China, and its genetic pattern has been studied using diverse genetic markers including amplified fragment length polymorphism (Zhao *et al.*, 2007) and sequence data (Ni *et al.*, 2012). In this study, we accessed its genetic patterns using a partial conserved 28S rRNA gene. Unsurprisingly, the data set was characterized by an extremely low level of polymorphism. Only seven haplotypes were defined and no more than five mutations were required to

NI

nal of Biogeography, **36**: 3-15.

Beheregaray, L. B., 2008. Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**: 3754-3774.

Clement, M., Posada, D., and Crandall, K. A., 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, **9**: 1657-1659.

Cutter, A. D., 2013. Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution*, **69**: 1172-1185.

Dover, G., 1982. Molecular drive: A cohesive mode of species evolution. *Nature*, **299**: 111-117.

Dupanloup, I., Schneider, S., and Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**: 2571-2581.

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E., *et al.*, 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences*, **107**: 16196-16200.

Excoffier, L., and Lischer, H. E. L., 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.

Excoffier, L., Smouse, P. E., and Quattro, J. M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**: 479-491.

Freeland, J. R., Kirk, H., and Petersen, S., 2011. *Molecular Ecology*. John Wiley & Sons, Ltd., Chichester, UK, 35-76.

Funk, D. J., and Omland, K. E., 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **34**: 397-423.

Hare, M. P., 2001. Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution*, **16**: 700-706.

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., *et al.*, 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**: 291-301.

Hurry, C. R., Schmidt, D. J., Ponniah, M., Carini, G., Blair, D., and Hughes, J. M., 2014. Shared phylogeographic patterns between the ectocommensal flatworm *Temnosewellia albata* and its host, the endangered freshwater crayfish *Euastacus robertsi*. *PeerJ*, **2**: e552.

Kong, L. F., and Li, Q., 2009. Genetic evidence for the existence of cryptic species in an endangered clam *Coelomactra antiquata*. *Marine Biology*, **156**: 1507-1515.

Lemer, S., Buge, B., Bemis, A., and Giribet, G., 2014. First molecular phylogeny of the circumtropical bivalve family Pinnidae (Mollusca, Bivalvia): Evidence for high levels of cryptic species diversity. *Molecular Phylogenetics and Evolution*, **75**: 11-23.

Li, Q., Park, C., and 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., and Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**: 1451-1452.

Liu, J. X., Gao, T. X., Wu, S. F., and Zhang, Y. P., 2007. Pleistocene isolation in the Northwestern Pacific marginal seas and limited dispersal in a marine fish, *Chelon haematocheilus* (Tem-minck & Schlegel, 1845). *Molecular Ecology*, **16**: 275-288.

May, R. M., and Godfrey, J., 1994. Biological diversity: Differences between land and sea [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **343**: 105-111.

Mikkelsen, P. M., Bieler, R., Kappner, I., and Rawlings, T. A., 2006. Phylogeny of Veneroidea (Mollusca: Bivalvia) based on morphology and molecules. *Zoological Journal of the Linnean Society*, **148**: 439-521.

Ni, G., Li, Q., Kong, L. F., and Yu, H., 2014. Comparative phylogeography in marginal seas of the northwestern Pacific. *Molecular Ecology*, **23**: 534-548.

Ni, G., Li, Q., Kong, L. F., and Zheng, X. D., 2012. Phylogeography of bivalve *Cyclina sinensis*: Testing the historical glaciations and Changjiang River outflow hypotheses in northwestern Pacific. *PLoS One*, **7**: e49487.

Qi, Z. Y., 2004. *Seashells of China*. China Ocean Press, Beijing, China, 608pp.

Qiu, F., Li, H., Lin, H., Ding, S., and Miyamoto, M. M., 2016. Phylogeography of the inshore fish, *Bostrychus sinensis*, along the Pacific coastline of China. *Molecular Phylogenetics and Evolution*, **96**: 112-117.

Rice, W. R., 1989. Analyzing tables of statistical tests. *Evolution*, **43**: 223-225.

Schlötterer, C., 2004. The evolution of molecular markers – Just a matter of fashion? *Nature Review Genetics*, **5**: 63-69.

Shafer, A., Cullingham, C. I., Cote, S. D., and Coltman, D. W., 2010. Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**: 4589-4621.

Turchetto-Zolet, A., Pinheiro, F., Salgueiro, F., and Palma-Silva, C., 2013. Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**: 1193-1213.

Voris, H. K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**: 1153-1167.

Wang, P. X., and Sun, X. J., 1994. Last glacial maximum in China: Comparison between land and sea. *Catena*, **23**: 341-353.

Xiao, Y. S., Zhang, Y., Gao, T. X., Yanagimoto, T., Yabe, M., and Sakurai, Y., 2009. Genetic diversity in the mtDNA control region and population structure in the small yellow croaker *Larimichthys polyactis*. *Environmental Biology of Fishes*, **85**: 303-314.

Xu, J., Chan, T. Y., Tsang, L. M., and Chu, K. H., 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.

Zane, L., Bargelloni, L., and Patarnello, T., 2002. Strategies for microsatellite isolation: A review. *Molecular Ecology*, **11**: 1-16.

Zeng, Z. N., and Li, F. X., 1991. The study of reproductive cycle of *Cyclina sinensis*. *Tropic Oceanology*, **10**: 86-93.

Zhang, A. B., Kubota, K., Takami, Y., Kim, J. L., Kim, J. K., and Sota, T., 2006. Comparative phylogeography of three Leptocarabus ground beetle species in South Korea, based on the mitochondrial COI and nuclear 28S rRNA genes. *Zoological Science*, **23**: 745-754.

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

(Edited by Qiu Yantao)