Phylogenetic Relationships of the Genus *Meretrix* (Mollusca: Veneridae) Based on Mitochondrial COI Gene Sequences

CHEN Ai-hui1, 2,*, LI Zhao-xia2, FENG Gong-neng2

(1. Jiangsu Provincial Laboratory of Coastal Wetland Bio-resources and Environmental Protection, Yancheng 224002, China; 2. School of Chemical and Bioengineering, Yancheng Institute of Technology, Yancheng 224003, China)

Abstract: Fifteen sequences from the mitochondrial cytochrome c oxidase subunit I gene (COI) were determined for 4 species of the genus *Meretrix*, with the homologous sequences of *M. petechialis* obtained from the GenBank data library. The alignment length of the sequences was 574 bp after excluding ambiguous sites, including 93 parsimony informative sites. In the fragments, the percentages of A, T, C and G were 21.15%, 44.71%, 14.05% and 20.09% respectively. There were 12 haplotypes identified: 4 *M. meretrix*, 2 *M. lamarckii*, 3 *M. lusoria*, 1 *M. lyrata* and 2 *M. petechialis*. Furthermore, it was revealed that *M. meretrix*, *M. petechialis* and *M. lusoria* shared some haplotypes. Phylogeny trees were reconstructed by Maximum-parsimony (MP) and Bayesian method using *Cylina sinensis* as the outgroup. Our results indicated that *M. lusoria*, *M. petechialis* and *M. meretrix* are closely related species. This is in accordance with the viewpoint that *M. petechialis* and *M. lusoria* should be treated as a junior synonym of *M. meretrix*.

Key words: *Meretrix*; *M. meretrix*; COI gene; Phylogeny

Members of the genus *Meretrix* are highly appreciated table food for internal and export markets, fetching a high price. Some species have already become regional economic mainstay in coastal areas of China. However, previous research on species of the *Meretrix* mainly focused on aquaculture and fishery and paid little attention on its classification. Thus, the taxon of the *Meretrix* is still in argument over several discrepancies among systematists in the world. Traditionally hard clams are mainly identified based on visible shell characters. Commonly, taxonomic ambiguities exist due to morphological variability. Modern taxonomic work includes analysis of a host of other traits, including anatomy, physiology, behavior, and genetics. Citations of clam studies include many synonyms, indicating ambiguities in *Meretrix* species
identification. Accurate species identification is very important in the case of multiple morphology-based classification and determining similar species for aquaculture management, biodiversity studies, and population dynamics.

The mitochondrial cytochrome c oxidase subunit I gene (COI) shows distinct divergence and provides valuable information in species identification to complete taxonomic data and validation of systemic position, phylogeny (Machordom et al, 2003; Smith et al, 2004; Donald et al, 2005). In the present study, the COI gene is selected to elucidate taxonomy and phylogenetic relationships among 5 Meretrix species from the Chinese coast.

1 Materials and Methods

1.1 Samples and DNA extraction

Geographical origins of the Meretrix used here are shown in Tab. 1. All the clams were identified morphometrically, and then fifteen of each species’ adductor muscle tissue samples were collected and preserved in 100% alcohol.

| Genus | Species and haplotype number identifying the sequence | Locality | GenBank accession numbers |
|-------|------------------------------------------------------|----------|----------------------------|
| *Meretrix* | M. meretrix 1 | Jiangsu, China | FJ434675 |
|          | M. meretrix 2 | Jiangsu, China | FJ434676 |
|          | M. meretrix 3 | Jiangsu, China | FJ434677 |
|          | M. meretrix 4 | Jiangsu, China | FJ434678 |
|          | M. lusoria 1 | Jiangsu, China | FJ434679 |
|          | M. lusoria 2 | Jiangsu, China | FJ434680 |
|          | M. lusoria 3 | Jiangsu, China | FJ434681 |
|          | M. lyrata | Hainan, China | FJ434682 |
|          | M. lamarecki 1 | Hainan, China | FJ434683 |
|          | M. lamarecki 2 | Hainan, China | FJ434684 |
|          | M. petechialis 1 | | AB280785* |
|          | M. petechialis 2 | | AY874530* |
| *Cylina* | C. sinensis 1 | Jiangsu, China | FJ434685 |
|          | C. sinensis 2 | Jiangsu, China | FJ434686 |

*Sequences downloaded from GenBank.

Total genomic DNA was extracted from the stored tissue samples by the standard Proteinase-K/Phenol-Chloroform-ethanol method. Before incubation the samples were soaked in ultra pure water for 2 days. Scissored tissue samples were re-suspended in 400 μL digestive system, which contained Tris (pH 8.0) 0.01 mol/L, EDTA (pH 8.0) 0.1mol/L, NaCl 0.05mol/L, SDS 1% and 10μL Proteinase K, and was incubated at 52°C for 12 h. The digested samples were phenol-extracted, ethanol-precipitated once more, and redissolved in 10 mmol/L Tris-HCl, pH 8.0. The concentration of DNA was estimated using a UV spectrophotometer. The DNA was diluted to a final concentration of 50ng/μL. To avoid DNA damage caused by multigelation, DNA samples were stored at 4°C.

1.2 PCR amplification

The partial region of the mitochondrial COI gene was amplified from 15 individuals of each species using the universal invertebrate primers COI-F 5'-GGTCAACTACATATAAGATATTGG-3' and COI-R 5'-TAAACTTCAAGGGTGACCCAAAATCA-3' (Folmer et al, 1994) synthesized by Shanghai Sangon Company, China. The polymerase chain reaction (PCR) is employed in this study to amplify the COI region. A total 30μL of each PCR reaction comprised the following: PCR buffer (10×) 3.0 μL, MgCl2 (25 mmol/L) 2.0 μL, dNTPs (2 mmol/L) 2.0 μL, forward primer (10 μmol/L) 1.0 μL, reverse primer (10μmol/L) 1.0μL, Taq polymerase 0.2 μL, DNA (50ng/μL) 1.0μL, ultra pure water 19.8μL. The Hot-Start PCR was employed with initial denaturation of 5min at 94°C followed by 30 cycles of denaturation for 30s at 94°C, annealing at 52°C for 30s and an extension of 72°C for 45s. After the completion of 30 cycles, a final extension step of 7min at 72°C was performed. The PCR product was then kept at 15°C until removed from the machine. The amplified product was tested in 2% agarose gel and visualized using the Gel Doc system. Those products with distinct and intense bands were selected after a series of sequencing protocols. The sequencing was done both in the forward and reverse directions by Shanghai Sangon Company.

1.3 Data analysis

Fifteen COI gene sequences of 4 species in the *Meretrix* (except sequences of *M. petechialis* were
sequenced, and COI sequences of *M. petechialis* were downloaded from GenBank. The raw nucleotide sequences obtained were assembled using SeqMan (DNASTAR package) and the sequence information checked entirely for consistency from both directions manually. We examined 574 base pairs (bp) and identified 12 haplotypes: 4 of *M. meretrix*, 2 of *M. petechialis*, 3 of *M. lusoria*, 1 of *M. lyrata* and 2 of *M. lamarckii* (Tab. 1). In which, the sequence of *M. lusoria* 3 was identical to the sequence of *M. petechialis* 2, and the sequence of *M. petechialis* 1 was identical to the sequence of *M. meretrix* 2. The collated sequences of 12 haplotypes were aligned using Clustal X1.8 (Thompson et al, 1997) with parameters on default. No insertions/deletions and stop codons were examined. These sequences were translated according to the invertebrate mitochondrial genetic code to the expected 191 amino acids. Nucleotide variation and substitution patterns were examined using the software MEGA3 (Kumar et al, 2004) on the basis of uncorrected pairwise genetic distance (p-distance).

Maximum-parsimony (MP) analyses were performed on PAUP 4.0b10 (Swofford, 2002), using *Cylina sinensis* as an outgroup species. Heuristic MP Search was performed using the simple addition sequence and the tree bisection-reconnection (TBR) branch swapping algorithm. Both weighted and unweighted analyses were performed. Weighting was carried out to equalize contributions from each codon position: positions weighted 3: 11: 1 for first, second, and third positions, respectively. Levels of support for individual relationships were estimated through 1,000 bootstrap replicates.

Bayesian inference was done using MrBayes 3.0b4 (Huelsenbeck & Ronquist, 2001) with the best-fitting model TrN+G estimated by Modeltest 3.7 (Posada, 1998). Trees saved below the burn-in generations were discarded, and a majority-rule consensus of the remains were calculated in MrBayes 3.0b4, providing posterior probabilities for clades. The MrBayes 3.0b4 was run with the following specifications: TrN model with a gamma distribution (invgamma), Markov’s chains started from a random tree for 400,000 generations, and sampling of the Markov chains at intervals of 100 generations. Four chains were run simultaneously with the initial 200 cycles discarded as burn-in.

### 2 Results

#### 2.1 Description of data

After aligning, using Clustal X1.8, the length of COI was 574 bp. The obtained COI gene sequences contained 149 variable sites and 93 parsimony-informative sites. The average base composition were 21.15% A, 44.71% T, 14.05% C, and 20.09% G. The A + T contents were higher than those of G+C, which is a pattern that has been repeatedly seen in the mtDNA of Molluscs.

| All sites | 1st codon position | 2nd codon position | 3rd codon position | Amino acids |
|----------|-------------------|-------------------|-------------------|-------------|
| Total    | 574               | 192               | 191               | 191         |
| Invariant| 425               | 158               | 184               | 83          | 150         |
| Variable | 149               | 34                | 7                 | 108         | 41          |
| Parsimony| 93                | 20                | 6                 | 67          | 29          |

Nucleotide variation and substitution patterns were examined using the software package MEGA 3.0. The average values of intraspecific pair-wise sequences divergence was 4.5%. The highest conspecific divergences were among individuals of *M. lamarckii*, which showed uncorrected divergence of 8.54%. Divergence could not be calculated for *M. lyrata* as it was represented by a single haplotype. Sequence variations between different species within the *Meretrix* ranged from about 1.83% to 14.81%. The uncorrected divergences of *M. petechialis* and *M. lusoria* with *M. meretrix* are 1.83%, 4.60%, which are lower the average uncorrected congeneric divergence of 11.08% of the genus *Meretrix*. (Tab. 3, Tab. 4).

Among the 12 haplotypes, the third codon positions were invariant in only 83 out of the 191 sequenced samples in the *Meretrix* (Tab. 2). In contrast, more than four-fifths of the first codon positions (158 of 192) and about 95% of the second codon positions (184 of 191) were monomorphic. The number of parsimony informative sites at each of the codon positions for the whole data set, are also given in Tab. 2. There were approximately eleven times as many parsimony informative sites at the third codon positions (67) than the second codon positions (6), and about three times as many as at first positions (20). For the data set as a whole,
the ratios are 20: 6: 67 for first, second, and third positions, respectively—a ratio of about 3: 1: 11. To offset the preponderance of the third position variation in the reconstruction of phylogenetic trees, weighted analysis using a weighting of sites of 3: 11: 1 was included.

2.2  Phylogenetic relationships

Figs. 1-3 present the phylogenies recovered under parsimony and Bayesian analysis, respectively. The values of interior branch test for some of the nodes in the parsimony trees and the posterior probability values for some nodes in the Bayesian tree were low. However, the topologies of the trees were very similar in most clusters.

In this analysis, haplotypes of *M. petechialis* and *M. lusoria* are nested within *M. meretrix* forming clade I. The difference among all the trees is the position of *M. lyrata*. In parsimony tree (weighted), *M. lyrata* clustered with clade I, but is sister to the assemblage (clade I + *M. lamarckii*) in another parsimonious tree and Bayesian tree.

3  Discussion

3.1  Taxonomic position of *M. petechialis* and *M. lusoria*

*M. lusoria* and *M. petechialis* were previously considered as independent species, different from *M.*

### Tab. 3  Intra specific nucleotide P distances for the *Meretrix* species

| Species          | Mean P distance |
|------------------|-----------------|
| 1. *M. meretrix* | 0.0325±0.0050   |
| 2. *M. petechialis* | 0.0052±0.0030   |
| 3. *M. lusoria* | 0.0697±0.0082   |
| 4. *M. lyrata* | n/c             |
| 5. *M. lamarckii* | 0.0854±0.0113   |

Average intraspecific P distance in the genus = 0.045

### Tab. 4  Inter specific nucleotide P distances among the *Meretrix* species

| Species       | 1     | 2     | 3     | 4     | 5     |
|---------------|-------|-------|-------|-------|-------|
| 1. *M. meretrix* |       |       |       |       |       |
| 2. *M. petechialis* | 0.0183 |       |       |       |       |
| 3. *M. lusoria* | 0.0460 | 0.0438 |       |       |       |
| 4. *M. lyrata* | 0.1407 | 0.1385 | 0.1440 |       |       |
| 5. *M. lamarckii* | 0.1448 | 0.1420 | 0.1414 | 0.1481 |       |

Average interspecific P distance in the genus =0.1108

Fig. 1  MP tree resulting from the analysis of the COI gene sequences of 12 haplotypes

Numbers on nodes correspond to percentage bootstrap values for 1000 replicates.
Fig. 2 MP tree resulting from analysis of the COI gene sequences of 12 haplotypes
The tree is based on a weighting of codon positions 3:11:1 for first, second, and third codon positions, respectively. Numbers on nodes correspond to percentage bootstrap values for 1000 replicates.

Fig. 3 Bayesian tree resulting from the analysis of the COI gene sequences of 12 haplotypes
Numbers on nodes correspond to the value of posterior values.

meretrix, in the genus Meretrix (Jukes-Browne, 1914; Habe, 1997; Zhuang, 2001). It was shown that the two species differed from *M. meretrix* by the following shell characteristics: Anterior and ventral margins, posterior end of shell. But the results from COI gene sequences analysis strongly disagreed with their viewpoints. According to the phylogenetic trees obtained by this study, *M. petechialis* and *M. lusoria* nested within *M. meretrix*, of them, *M. meretrix* has rounded posterior end, *M. petechialis* and *M. lusoria* are bluntly angled.
Therefore, the shape of the shell end is a homoplasious character, and species determination in the Meretrix using this character should be investigated further. The results obtained in this study do not support the present taxonomic status of M. lusoria and M. petechialis. This is in accordance with the ideas from allozyme analysis that M. lusoria and M. petechialis were the most closely related species within the genus Meretrix (Ayako et al., 2008). This viewpoint also supported that M. lusoria and M. meretrix belong to different geographic subspecies of one species (Pan et al, 2006).

In our study, the COI gene fragment of M. lusoria 3 is identical to M. petechialis 2, and M. petechialis 1 is identical to M. meretrix 2. Moreover, the unweighted divergences (1.83%, 4.60% and 4.38%) among these three species are much lower than the average interspecific divergence (11.08%). Prashad (1932) pointed out that M. meretrix is a species which experienced the greatest variation in the group of bivalves, and because of shades of shells and shell colors, it was wrongly divided into many species. Furthermore, Fischeer-Piette (1941) also suggested that M. petechialis and M. lusoria should be treated as the same species of M. meretrix.

In summary, the views of taxonomic status of M. lusoria and M. petechialis are ambiguous based on morphology. In our study, the viewpoint that M. petechialis and M. lusoria should be treated as a junior synonym of M. meretrix was supported.

3.2 Classification of the genus Meretrix

The Meretrix is widely scattered through coastal areas of the Indian Ocean, Southeast Asia, China, Korean Peninsula and Japanese Archipelago. Because of the remarkable variation of shapes and patterns of shells, early researchers divided this genus into many species, such as M. meretrix, M. typical, M. petechialis, M. castanea, M. graphica, M. labiosa, M. inpudina, M. zonaria, M. lusoria, M. lyrata, M. lamarckii, M. mophina, M. exilis, etc. When modern scholars (Dautzenberg & Fischer, 1905; Dautzenberg, 1906; Fischer-Piette, 1941, 1976) merged the clams into one genus, but retained the aforementioned names, there was still discrepancies on the species names used. To date, 9 species are generally recognized in Meretrix (Ayako et al, 2008). Zhuang (2001) pointed out that Meretrix has a little species and only three species, M. meretrix, M. lusoria and M. lamarckii, have been found in China. He admitted M. lyrata but didn't incorporate it into Chinese fauna. As M. lyrata occurs on the west coast of southern China and M. petechialis occurs from the west coast of the Korean Peninsula to southern China and Vietnam, the author considered that they should be included into Chinese fauna. This means that there are 5 species (including subspecies) in the Chinese fauna: M. meretrix, M. lusoria, M. lamarckii, M. petechialis and M. lyrata.

Former studies show that the COI gene sequences are rich in variations and fit for systematic analysis (Machordom, 2003; Donald, 2005). Bayesian tree indicated that the clade M. lamarckii is a sister group to the assemblage (M. lusoria + M. petechialis + M. meretrix). This is in accordance with the ideas from 16S rDNA and ITS1 that the phylogenetic relationships of four Meretrix species: ((M. lusoria + M. meretrix) + M. lamarckii) + M. lyrata) (Pan et al, 2006).

A result that is insensible to differential weighting of the data is to be regarded as the most robust hypothesis of phylogeny. As mentioned above, the position of M. lyrata in the most parsimonious tree is indeed sensitive to the differential weighting schemes (Figs. 1, 2). However, if one chooses to focus on supported groups (>70%), we prefer to do, there is no major difference between the weighted supported results. Thus, weighting did not seem to provide any further help in clarifying relationships of the selected taxa.

References:

Ayako YY, Masashi Y, Hideyuki I. 2008. Genetic Relationships among Species of Meretrix (Mollusca: Veneridae) in the Western Pacific Ocean [J]. Pacific Science, 62(3): 385-394.

Dautzenberg PH, Fischer H. 1905. Liste des mollusques recoltes par M. le capataine de fregate blaise au Tonkin, et description d’espèces nouvelle [J]. Journ Conchyliol, 4(53): 213-224.

Dautzenberg PH. 1906. Contribution a la faune malacologique de l’Indo-Chine [J]. Journ Conchyliol, 54: 215-219.

Donald KM, Kenned YM, Spencer HG. 2005. The phylogeny and taxonomy of austral monodontine topshells (Mollusca: Gastropoda: Trochidae) inferred from DNA sequences [J]. Molecular Phylogenetics and Evolution, 37: 474-483.

Fischer-Piette E. 1941. Revision des vivanuts de Meretrix s. s. du museum national d’ histoire naturelle [J]. Journ Conchyliol, 84: 315-344.

Fischer-Piette E. 1976. Les Veneridae indeterminate das collection de Calcutta [J]. Rec Zool Surv India, 70: 235-257.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit 1 from diverce metazoan invertebrates [J]. Mol Mar Biol Biotech, 3: 294-299.

Habe T. 1977. New and little known bivalves of Japan [J]. Venus, 36
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees [CP]. Bioinformatics, 17: 754-755.

Jukes-Browne AJ. 1914. A synopsis of the family Veneridae. Part I and II [J]. Proc Malac Soc Lond, 11: 58-94.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment [CP]. Briefings in Bioinformatics, 5: 150-163.

Posada D, Crandall KA. 1998. MODELTEST: Testing the model of DNA substitution [CP]. Bioinformatics, 14: 817-818.

Prashad B. 1932. The Lamellibranchia of the Siboga Expedition [M]. Systematic part. Siboga-Expedition, LII/1: 212-264.

Smith PJ, Mcveagh SM, Won Y, Vrijenhoek RC. 2004. Genetic heterogeneity among New Zealand species of hydrothermal vent mussels (Mytilidae: Bathymodiolus) [J]. Marine Biology, 144: 537-545.

Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony [CP]. Sunderland, MA: Sinauer Associates.

Zhuang QQ. 2001. Fauna sinica, phyllum Mollusca, class Bivalvia, family Veneridae [M]. Beijing: Science Press, 229-236.

三、会议费用（详见http://www.czs.ioz.ac.cn）

四、注册费付款方式（请通过电汇或邮政汇款）

1. 电汇帐号：（请注明“中国动物学会会议注册费”）
   开户名称：重庆动物学会（请注明，没有“市”字）
   银行帐号：3100028109200052809
   开户银行：工行重庆朝阳支行

2. 邮政汇款：（请注明“中国动物学会会议注册费”）
   汇款地址：重庆北碚天生路2号，西南大学生命科学学院
   收款人：蒲德永
   邮政编码：400715

3. 缴纳注册费联系人：
   蒲德永 E-mail：jia@swu.edu.cn 联系电话：023-68253005，13883220085

4. 会议组联系人：
   王志坚（西南大学生命科学学院副院长，重庆动物学会副秘书长）
   联系方式：电子邮件：wangzj@swu.edu.cn；
   联系电话：023-68253005；手机13983426426
   传真：023-68253005；03-68252365

   请参加会议的代表、科技工作者及学生于2009年6月20日前给重庆动物学会秘书处回执，以便会务工作安排等。

中国动物学会秘书处
2009年5月20日