DNA barcoding of true limpets (Order Patellogastropoda) along coast of China: a case study

Jiping Lin, Lingfeng Kong, and Qi Li

Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao, China

Abstract
In this study, we applied a partial sequence of mitochondrial COI gene as DNA barcode to assess the viability of DNA barcoding for distinguishing Patellogastropoda. One-hundred thirty-five COI gene sequences were obtained from 13 species belonging to Nacellidae (Cellana) and Lottiidae (Lottia, Patelloida and Nipponacmea) along the coast of China. The alignment result of these sequences indicated the existence of insertions in mitochondrial COI gene of Patellogastropoda. The Kimura 2-parameter (K2P) distances within species and genera were 0.00–1.01% (average 0.07%) and 18.09–37.80% (average 24.07%), respectively, an obvious barcoding gap existed. All species in our study were clearly discriminated in all trees (neighbor-joining (NJ), Bayesian, and maximum likelihood (ML) tree) with a highly supported clade node. The character-based barcode method successfully identified 100% of the Patellogastropod species included, and performed well in discriminating Patellogastropod genera. The results of this study affirm that DNA barcoding based on the COI gene can identify species belonging to Patellogastropoda rapidly and accurately.

Keywords
COI, DNA barcoding, Patellogastropoda, true limpets

Introduction
Patellogastropod limpets which are called true limpets are common and prominent species which could be seen colonizing in intertidal rocky shores throughout the world’s oceans from the Polar Regions to the tropics. They belong to a wide variety of species and play an important role in the coastal marine ecosystems (Branch, 1985). The number of extant Patellogastropod species recorded around the world is about 930 and shown in World Register of Marine Species (Costello et al., 2013, http://www.marinespecies.org/). To discriminate Patellogastropod species, a few morphological characters have been used (e.g. sperm, radular morphology and headfoot). However, the principal character used for the species-level taxonomy of Patellogastropoda was the external feature of the shell. In spite of this, it is known that shells are highly variable and usually result in taxonomy confusions (Mauro et al., 2003).

In the last two decades, molecular short standardized DNA fragments, termed DNA barcodes, have been developed for species discrimination (Hebert et al., 2003). Relative to the traditional morphological species-identification method alone, DNA barcoding has three main advantages: (i) more accuracy and efficiency in specie identification (Rivera & Currie, 2009); (ii) simplicity in operation; (iii) less time-consuming when identifying a huge number of specimens. Since DNA barcoding is an excellent species-identification technique, a lot of taxonomists had applied DNA barcoding to identify species accompanying with the traditional morphological method, such as in birds (Hebert et al., 2004), fish (Holmes et al., 2009) and marine molluscs, including chitons (Kelly et al., 2007), Thecosomata (Maas et al., 2013), etc.

In China, Patellogastropod limpets distribute almost throughout the whole coastline of China. Some Patellogastropod species, like Cellana toreuma, play quite important role in the Chinese coastal marine ecosystem, such as promoting the energy flow and matter cycling of rocky intertidal zones, regulating the rocky intertidal zones’ biology community structure (Wang et al., 2001). However, research about Patellogastropod species discrimination in China mainly remained in the morphological level. Because of the high variability in shells and the existence of phenotypic plasticity and synonyms, in China, the classification and the identification of extant Patellogastropod species based on the traditional morphological method have been challenging, except in a few species (Cellana toreuma, Cellana grata, Nipponacmea schrenckii and Lottia dorsuosa) which possess distinctive features like texture, color and shape (Huang & Lin, 2012; Zhang, 2008).

Considering the difficulty in discriminating Patellogastropod species by morphological features, more rapid and accurate species-identification method is needed. In light of the success that DNA barcoding has achieved in other taxa including marine molluscs around the world (Layton et al., 2014; Sun et al., 2012), in this study, we employed the most widely used “barcode” COI gene to evaluate the utility of DNA barcoding for species identification, discovery of possible cryptic species and correction of morphological misidentification in Patellogastropoda. We sequenced COI gene from 135 specimens of 13 species in four genera which belonged to two families of Patellogastropoda collected from the coast of China.
Materials and methods

Specimens sampling
Patellogastropod limpets were collected during June 2007–April 2014 from intertidal rocky shores along the coast of China (Table S). Specimens were identified based on shell characters (Huang & Lin, 2012; Okutani, 2000). Several specimens could not be identified to species reliably because of their varying shells, they were subjected to molecular analysis using COI sequences, and we referred to these specimens as sp.

DNA isolation, amplification and sequencing
About 50–100 mg of tissue was removed from specimens’ foot tissue and used for DNA isolation by the CTAB method, slightly modified from Winnepennixx et al. (1993). Isolated DNA was dissolved and stored in TE buffer, then frozen at −20°C until used. A partial region of mitochondrial COI gene was amplified by polymerase chain reaction (PCR) using universal primers LCO1490/HCO2198 (Folmer et al., 1994). For the specimens that were not successfully amplified by the universal COI primers, the other pair of primers jgLCO1490/jgHCO2198 (Geller et al., 2013) was used. As jgLCO1490/jgHCO2198 were redesigned from the universal COI primers, the pair of primers jgLCO1490/jgHCO2198 (Geller et al., 2013) was used. As jgLCO1490/jgHCO2198 were redesigned from LCO1490/HCO2198, both sets of primers amplify the same segment. PCR amplification was performed in a total volume of 20 μL, including two U Taq DNA polymerase (Takara, Tokyo, Japan), about 30–100 ng of template DNA, 1 μM of forward and reverse primers, 200 μM of each dNTP, 1× PCR buffer and 2 mM MgCl2. The PCR was carried out under the following conditions: 94°C (3 min); five cycles of 94°C (40 s), 40°C (40 s), 72°C (1 min); 30 cycles of 94°C (40 s), 42–52°C (universal primers, 40 s), or 48–52°C (primers of Geller, 40 s), 72°C (1 min), and 72°C (10 min).

PCR products were first visualized on 1.5% agarose gels with ethidium bromide and then purified by EZ Spin Column DNA Gel Extraction kit (EZ BioResearch, St. Louis, MO). The purified products were sequenced using ABI PRISM 3730 (Applied Biosystems, Waltham, MA).

Data analysis
The complementary DNA sequences were viewed and edited manually by comparing both strands through SEQMAN software (DNASTar 7.2.1, DNASTAR Lasergene, Inc, Madison, WI). Sequences were aligned using CLUSTALW (Thompson et al., 1994) in BioEdit 7.2.5 (Qiagen Inc., Valencia, CA) (Hall, 1999) and deposited in GenBank with the accession numbers KM221033–KM221167. COI barcodes were 658 bp (Folmer primers) and 661 bp/658 bp (Geller primers). As some specimens possessed long segment of uncertain bases, uncertain bases at both ends were all trimmed, final COI data set was aligned to 498 bp for the high quality of the data.

To determine the sequence divergence, the Kimura 2-parameter (K2P) model (Kimura, 1980) was used. To visualize these distances, a neighbor-joining (NJ) tree with bootstrap analysis (1000 replicates) was performed using 37 haplotypes in our study. MEGA 6 software (MEGA Inc., Englewood, NJ) was used to build NJ tree here (Tamura et al., 2013). For the reliability and the validity of the results, the Bayesian tree and the maximum-likelihood (ML) tree were established in this study using 37 haplotypes. The best-fit model of DNA sequence evolution was evaluated using jModelTest 2.1.4 (Software Foundation, Inc., Boston, MA) (Darriba et al., 2012) under the Akaike information criterion (AIC). The TVM+I+G model was selected to generate the Bayesian tree, which was obtained with MrBayes 3.2 (Software Foundation, Inc., Boston, MA) (Ronquist & Huelsenbeck, 2003). The parameters set for the Bayesian tree were random trees = 2; mcmc ngen = 2,000,000; samplefreq = 100; sump burnin = 5000. The ML tree was built by online program PhyML 3.0 (PhyML 3.0) (Guindon & Gascuel, 2003, http://www.atgc-montpellier.fr/phyml/).

We also employed the character-based DNA barcode method to measure barcode success in our study. The characteristic attribute organization system (CAOS) was applied for this method (Sarkar et al., 2008). CAOS identified pure unique diagnostics, termed as “characteristic attributes” (CAs), at the target branching nodes. Herein, sequences of 37 haplotypes belonging to 13 species of the Patellogastropoda were used to constitute a reference dataset. As recorded in literatures, 13 species represent main limpet species along the coast of China. Phylogenetic trees of the reference sequences were produced in MEGA 6 software (MEGA Inc., Englewood, NJ) (Tamura et al., 2013) using the K2P model and incorporated into NEXUS files with the DNA data matrix in MacClade v4.08 (Sinauer Associates, Mishawaka, IN) (Maddison & Maddison, 2005). Then the datasets were executed in P-Gnome. The most variable sites that distinguish all the taxa were chosen and the character states at these nucleotide positions were listed. Unique combinations of character states (character-based DNA barcodes) were identified.

Results
Totally, 91 individuals were amplified by LCO1490/HCO2198 (Folmer et al., 1994) and 44 individuals including whole individuals of Lottia luchuana, L. doro-suosa, Patelloidea heroldi, P. saccharina lanx and partial individuals of Nipponacmea schrenckii, N. sp., P. ryukyuensis, P. conulus and L. cassisi were amplified by the primers jgLCO1490/jgHCO2198 (Geller et al., 2013). Four individuals of Nipponacmea schrenckii and five individuals of Patelloidea ryukyuensis encountered amplifying failure using both pairs of primers. Insertions were detected in the four species of Patelloidea and four species of Lottia which were both belonged to family Lottiidae. Only a single codon insertion occurred in all these eight species.

Intraspecific genetic divergence based on the Kimura 2-parameter ranged from 0.00% to 1.01%, with an average of 0.07%. The greatest intraspecific genetic divergence (1.01%) was found in Lottia luchuana as well as Patelloidea ryukyuensis. Interspecific genetic divergence within genera ranged from 18.09% to 37.80% with a mean of 24.07%. The lowest interspecific divergence within genera was found between Patelloidea heroldi and P. ryukyuensis (18.09%), while Lottia cassisi and L. doro-suosa contributed the highest interspecific divergence within genera (37.80%). The barcoding gap was very clear.

Topologies revealed by all analysis methods (NJ, ML and Bayesian) were not identical (Figure 1), but the recovered tip clad (i.e. putative species) of all trees were the same. In all the tree topologies of COI data set, sequences obtained from three genera (Lottia, Patelloidea and Nipponacmea) of Lottiidae and the other genus (Cellana) of Nacellidae, respectively, grouped together by taxonomic affinity and formed two main phylogenetic clad. Eleven putative species (Cellana grata, C. toreuma, Lottia doro-suosa, L. cassisi, L. luchuana, Patelloidea ryukyuensis, P. heroldi, P. conulus, P. saccharina lanx, Nipponacmea schrenckii and N. concinna) were recognized among the 37 haplotypes, while two species were identified to genus level (L. sp. and N. sp.) using the COI phylogeny reconstruction. The haplotypes of two unidentified species (L. sp. and N. sp.) were blasted to determine if their sequence was already present in Genbank. However, there were not suitable blast results found in Genbank which indicate that they might represent new molecular operational taxonomic units (MOTUs).
The character states at 21 nucleotide positions of the COI gene region for 13 species of the Patellogastropoda are provided in Table 1. The particular nucleotide positions were chosen due to the high number of CAs at the important nodes or because of the presence of CAs for groups with highly similar sequences. All the 13 species revealed a unique combination of character states at 21 nucleotide positions with at least four CAs for each species. As to genus level, all the genera revealed a unique combination of character states at the 10 nucleotide positions with at least four CAs for each genus.

**Discussion**

Barcoding analysis presents an extremely useful tool to identify species because of the difficulty in diagnosing species from morphological characters. In our study, 124 specimens were successfully identified to 11 species according to phylogenetic approach and distance-based DNA barcoding, while several samples were identified only to genus level. Meyer & Paulay (2005) proposed that the extent of the separation between intraspecies and interspecific divergence in the selected marker determined the accuracy of distance-based DNA barcoding, chiefly. A 10× threshold rule (Hebert et al., 2004) and even the smallest over mean interspecific distances to compute barcoding gaps (Meier et al., 2008) were recommended to identify cases where a current species represented two or more taxa. In our study, the maximum intraspecific distance (1.01%, *Patelloida ryukyuensis*) was much lower than the minimum interspecific distance (18.09%, between *Patelloida ryukyuensis* and *P. heroldi*) and the 10× rule was fulfilled in all the species. To the character-based method, DNA barcoding is effective for the identification of genetic entities at species and genus levels in this study. On the species level, all species revealed a unique combination of character states at 21 nucleotide positions with at least four CAs for each species. Every species even closely related species *Cellana toreuma* and *C. grata* could be distinguished by this method. On the genus level, we find character-based barcodes with at least four CAs for four genera in the COI gene region by which confirmed the classification about *L. sp.* and *N. sp.* in our study. Therefore, the resolution of COI was distinct and made Patellogastropod species separated into distinctive taxonomic groups availably.

The individuals which belong to *Lottia* (*Lottia dorsuosa, L. cassis, L. luchuana, L. sp.*) and *Patelloida* (*Patelloida ryukyuensis, P. heroldi, P. conulus, P. saccharina lanx*) of Lottiidae in our study all possessed one single codon insertion at same site in the barcode region of COI. Hebert (2003) established that indels are usually rare in the barcode region of COI, but relatively common in some classes of molluscs. Layton et al. (2014) proposed that insertions in COI gene were detected in bivalves (Arcidae, Astartidae, Carditidae, Glycymerididae, Mytilidae and Thyasiridae) and gastropods (*Limacina helicina, Lottiidae*). In some other current studies of marine molluscs, insertions were also found in the barcode region of COI genes, such as Vanikoro (Collin, 2003) and *Thyasira* (Mikkelsen et al., 2007). Insertions in COI gene of gastropods has been linked to accelerate rates of nucleotide substitution in these groups in prior work, but the functional significance of these changes in the COI of gastropods was unclear (Remigio & Hebert, 2003). Future work should focus on determining functional significance of this variation along with its association with rates of molecular evolution and the mechanisms responsible for.
Table 1. Character-based DNA barcodes for 37 haplotypes of 13 Patellogastropod species; character states (nucleotides) at 21 selected positions of the COI gene region (ranging from position 8–482).

| Species                        | Haplotypes          | Positions |
|-------------------------------|---------------------|-----------|
| Cellana grata (Gould, 1859)   | Hap11               | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                              | G G T G T G T G T G |           |
|                               | G T G G G A G T C C |           |
|                               | T T T T T T T T T |           |
| Cellana toreuma (Reeve, 1855) | Hap6                | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                              | T T C A A A A G A C |           |
|                               | C C T A A T T C A T |           |
|                               | A T A T A T T T T |           |
| Lottia cassis (Eschscholtz, 1833)| Hap2            | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | A C G C T A A T C A |           |
|                               | C A G A C C C T A |           |
|                               | A C T T T T T T |           |
| Lottia dorsuosa (Gould, 1859) | Hap8                | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | A G G T G T A T T |           |
|                               | T C G T T G G A |           |
|                               | T G G A T T G T |           |
| Lottia luchuana (Pilsbry, 1901) | Hap5            | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | A A T A A T A T |           |
|                               | T T T G T G T G |           |
|                               | G T G A T T G T |           |
| Lottia sp.                    | Hap34               | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | T T G G A T T T T |           |
|                               | T T A A A A G A G |           |
|                               | G A G A G T A A |           |
| Nipponacmea sp.               | Hap21               | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | A A A A A T G C G |           |
|                               | T A A T T G G A |           |
|                               | T G G A T T G T |           |
| Nipponacmea concinna (Lischke, 1870) | Hap16        | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | A G A G G T T G |           |
|                               | A T T G G G A |           |
|                               | C C C C C C C |           |
| Patelloidea conulus (Dunker, 1861) | Hap36       | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | G T T A C G G T |           |
|                               | A A C C G C G |           |
|                               | T G T G G C T |           |
| Patelloidea heroldi (Dunker, 1861) | Hap31       | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | T G T T A C G |           |
|                               | A G G T G T |           |
|                               | C G C C C C |           |
| Patelloidea ryukyuensis (Nakano & Ozawa, 2005) | Hap9       | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | T G C A T C A T |           |
|                               | A T G G C C C |           |
|                               | T G C C C C |           |
| Patelloidea saccharina lanx (Reeve, 1855) | Hap17       | 8 11 23 32 83 101 137 197 203 236 248 263 293 320 428 473 482 |
|                               | T G C A T C A T |           |
|                               | A T G G C C C |           |
|                               | T G C C C C |           |

Conclusion
One-hundred thirty-five specimens of Patellogastropoda here could be discriminated into 13 known or unknown species by morphology and COI-based DNA barcoding. Obvious barcoding gap and insertions in DNA barcode region were detected. However, the inefficiency of the universal primers in some individuals existed.

This study not only confirms that distance-based, tree-based and character-based DNA barcoding are powerful tools for rapidly identifying true limpet species but also provides a benchmark data for future studies.

Acknowledgements
The authors thank Dr. Akihiko Matsukuma for helping us in identifying the specimens.

Declaration of interest
The authors report that they have no conflicts of interest. The authors alone are responsible for the content and writing of the paper. This study was supported by research grants from National Natural Science Foundation of China (41276138 and 31372524), and Fundamental Research Funds for the Central Universities.
