Reassessment of the mitochondrial 12S-rRNA gene for DNA barcoding of museum specimens of shelled marine gastropods from Japan

Nao Fukunaga1, Moe Shimizu1, Shinnosuke Teruya2, Nazifa Naziha Razali1, Satoko Nakashima1,4, Naoto Tomobuchi1, Takenori Sasaki3, and Davin H. E. Setiamarga1,3,*

1National Institute of Technology (KOSEN), Wakayama College, Gobo, Wakayama, Japan, 644-0023
2Okinawa Prefectural Deep-sea Water Research Center, Kumejima, Okinawa, Japan
3The University Museum, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, 133-0023
4Present address: Graduate School of Frontier Science, The University of Tokyo, Kashiwa-no-ha, Chiba, Japan, 277-8561

Abstract. DNA barcoding is an effective and powerful tool for taxonomic identification and thus very useful for biodiversity monitoring. This study investigated the usefulness of the mitochondrial 12S-rRNA gene for the DNA barcoding of shelled marine gastropods. To do so, we determined partial 12S-rRNA sequences of 75 vouchered museum specimens from 69 species of shelled gastropods from Japan. The specimens have been identified morphologically, and natural history data catalog. Sequence analyses through BLAST searches, maximum likelihood phylogenetic analysis, and species delimitation analysis suggested that the 12S-rRNA gene is helpful for barcoding shelled marine gastropods. They thus could be helpful to complement barcoding studies using other markers such as COI. The analyses successfully confirmed all samples’ identity at higher taxonomy (subfamily and above), but much less so at the species level. Our result thus also underlines the lingering problem of DNA barcoding: The lack of comprehensive reference databases of sequences. However, since we provided sequences of properly curated, vouchered museum specimens in this study, our result reported here has thus also helped to give taxonomically reliable reference sequences for biodiversity monitoring and identifications of shelled gastropods which include many important fisheries species.

1 Introduction

With its ca. 35,000 of recorded species, Gastropoda, a class of shelled mollusks (Conchifera), is one of the most prominent invertebrate groups, in which, during its long evolutionary history, has radiated and occupied a diverse array of ecological niches in marine, freshwater, and biofouling/invasive organisms [2]. Many shelled members of this group are also greatly influenced by the recent ocean acidification event caused by global warming [1] [3-4]. As such, it is essential to monitor the diversity of this taxa constantly.

* Corresponding author: davin@wakayama.kosen-ac.jp

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
Recent development in DNA sequencing and the building up of DNA sequence databases have allowed for the usage of DNA sequences for the quick and effective identifications and classifications of samples collected from the field with relatively high accuracy, a method called “DNA barcoding” [5]. Further development of DNA sequencing technology (e.g., Next Generation Sequencing) has allowed for the development of a non-invasive method of biodiversity monitoring using DNA sequences, called the eDNA (environmental DNA = eDNA) method, which is essential barcoding of DNA fragments shed off into the environment by living organisms [6-8]. For DNA barcoding and e-DNA to work, the availability of a robust, reliable, and exhaustive reference database of DNA sequences collected from target taxa, is essential. However, at present, biases in biodiversity studies might have caused such a reliable database to be not available for some taxa, including gastropods (e.g. [9]). In addition, Machida et al. [10] and Page [11] reported that currently, some data are not adequately curated. As a result, some taxa might become unidentifiable and thus become “dark taxa.” This becomes very problematic in monitoring studies, especially those conducted by people with inadequate taxonomical skills and resources, or if target taxa contain many possible undescribed or cryptic species, or when conducting eDNA-based monitoring for which morphological samples are simply unavailable.

On the technical side, the marker genes must have enough base substitutions to distinguish different species of the same genus [5], but not different enough to dramatically differentiate sympatric individuals of the same species. Several effective universal primers to amplify a region of the mitochondrial COI gene have been developed for metazoans, causing a tremendous amount of DNA barcoding studies to be conducted and thus empirically shown that the amplified COI segment has enough substitution rates to distinguish animals at the species level (e.g. [12-17]). However, in a previous study, we have suggested that doing DNA barcoding with only one genetic marker could be risky because of primarily technical problems (limited primer efficacies and the inability of a single quality to place samples properly at higher taxonomies) [16-17]. In that study [16-17], we also suggested that using multiple markers would help to alleviate the problems because it would allow for the collection of a more complete genetic data (primers of different genes might work on samples not amplifiable with those of one marker gene, and a combination of markers would allow for a more robust phylogeny), and thus allowing researchers to collect a more complete picture reflecting the actual biodiversity. Therefore, in that study, we evaluated and thus proposed the usefulness of the nuclear gene Histone-H3 (H3) as a marker for DNA barcoding of shelled gastropods, using previously developed primers [18].

In this study, to develop and assess the utility of another molecular marker for DNA barcoding-based studies of shelled gastropods, we investigated the usefulness of the mitochondrial gene 12S-rRNA using previously reported primers [19-20]. The 12S-rRNA gene has been used in previous molecular phylogenetics, phylogeography, and DNA barcoding of various metazoans, including gastropods. We sequenced the 12S-rRNA of different vouched samples of shelled marine gastropods stored at the University Museum of The University of Tokyo, including some old museum samples (the oldest was sampled in 1999, and the latest was tested in 2015). Our result presented here, which was based on BLAST searches, phylogenetic analysis, and species delimitation analysis, has confirmed the usefulness of the 12S-rRNA as a genetic marker for DNA barcoding of shelled gastropods, which also include many fisheries species. The result also highlighted the lack of a robust and comprehensive reference database of this gene if it is to be used as a marker for gastropods. Meanwhile, because we used properly curated and vouched museum specimens as samples, the natural history data of our representatives are reliable. Therefore, we have also contributed a set of reference sequence data from taxonomically reliable samples through this study, which is crucial for biodiversity monitoring using DNA barcoding and e-DNA.
2 Materials and methods

2.1 Sample collection

A total of 75 individuals of 69 species of shelled gastropods were used in this study. All samples used in this study were vouchered specimens stored in the University Museum, The University of Tokyo. These samples were initially collected from various locations in Japan, and then fixed and stored in 95% EtOH. Morphological identifications (based on [21]) of collected samples were conducted before or after fixation. Representatives were chosen at random, with one or two individuals per species. Most specimens are at least nine years old, with the most senior sample collected in 1999, and the latest in 2015. The list of samples is provided in Table 1.

2.2 DNA sequencing and sequence data acquisition

A piece of the muscle tissue from the mantle or the foot (about 0.25 mg) was cut out from each sample. Total genomic DNA was extracted using the standard CTABphenol-chloroform method. PCR was performed using a standard protocol but with an annealing temperature of 52°C. Three combinations of previously published three primers (Table 2) [19-20] were used to amplify the 12S-rRNA fragment of samples. Sanger sequencing of amplicons (using both the forward and reverse primers) was outsourced (FASMAC Co. Ltd., Kanagawa, Japan). For comparison, we also sequenced a fragment of the COI gene of all samples, using previously published primers [22-24]. The list of primers used in this study is shown in Table 2. Obtained sequences were then checked for contamination by BLASTn searches [25]. Sequence fidelity was confirmed and edited manually on the software MESQUITE ver. 3.61 [26-27], by also simultaneously checking the chromatograms by eye (visualized on ApE ver.2.0.61 [28]). After sequence editing, all forward and reverse sequences were assembled manually by eye.

2.3 Sequence identification (DNA barcoding) through BLASTn searches

In order to confirm if the obtained sequences were homologous to previously published arrangements, and thus to get taxonomic information of the organisms from which the lines were obtained, we performed BLASTn searches on the assembled sequences. We consider a sample as correctly identified if the morphological identification matches the BLASTn search result. We also confirmed at which taxonomic level a particular sequence was identified (species, genus, family, order) to check the availability of reference sequences on GenBank and the fidelity of the GenBank sequences. BLASTn searches were conducted for both the 12S-rRNA and COI gene sequences.
Table 1. List of analyzed OTU in this study, along with the BLASTn search results for DNA

| Sample No. | Species | Genebank ID | BLASTn result | Homology Level | E-value | Identity (%) | Gaps % |
|------------|---------|-------------|---------------|----------------|---------|---------------|--------|
| SS45       | Batillaria multiformis | NC_047187.1 | Batillaria utramentaria | Genus | 0 | 98 | 0 |
| B308       | Batillaria zonalis | HQ833835.1 | Batillaria cumangeti | Genus | 0 | 98 | 0 |
| SS104      | Planaxis sulcatus | HQ833835.1 | Planaxis sulcatus | Species | 0 | 99 | 0 |
| B309       | Semialuca spirina libertina | NC_037771.1 | Semialuca spirina coreana | Genus | 0.0E-160 | 93 | 0 |
| B303       | Cerithidea cingulata | AB351593.1 | Cerithidea adiastrensis | Genus | 8.0E-126 | 88 | 3 |
| SS57       | Nerita albicilla | LC215360.1 | Nerita albicilla | Species | 0 | 99 | 0 |
| SS124      | Nerita helicinoides | MT66161.1 | Nerita chamaeleon | Genus | 6.0E-162 | 94 | 0 |
| SS59       | Nerita japonica | LC565707.1 | Nerita japonica | Species | 0 | 99 | 0 |
| SS126      | homosodoris japonica | MT161611.1 | Nerita chamaeleon | Genus | 1.0E-173 | 96 | 0 |
| SS120      | Nerita striata | KF728888.1 | Nerita fulgarus | Genus | 2.0E-162 | 94 | 0 |
| SS103      | Cassidula mustelina | KJ903919.1 | Cassidula nucleus | Species | 2.0E-171 | 96 | 0 |
| SS141      | Eronnea erroneae | HQ33835.1 | Eronnea erroneae | Species | 0 | 98 | 0 |
| SS83       | Echinolittorina radiata | AJ623151.1 | Echinolittorina radiata | Species | 0 | 99 | 0 |
| SS142      | Notocochlis gualteriana | MK507895.1 | Notocochlis sp. | Genus | 0 | 99 | 0 |
| SS143      | Canarium mutabile | MW244820.1 | Tridentarius dantatus | Family | 1.0E-163 | 94 | 0 |
| SS110      | Conomurex luhananus | KY853669.1 | Conomurex luhanamus | Species | 0 | 99 | 0 |
| B349       | Mitrella burchardi | HQ833864.1 | Mitrella bincincta | Genus | 0 | 99 | 0 |
| B366       | Fusinus ferrugineus | NC_039906.1 | Fusinus longicandus | Genus | 0 | 98 | 0 |
| SS115      | Pleuroplaca trapezium | MN32235.1 | Turritilanus turbines | Subfamily | 2.0E-137 | 90 | 1 |
| B355       | Nassarius comoidalis | NC_041310.1 | Nassarius comoidalis | Species | 0 | 99 | 0 |
| B342       | Nassarius fraterculus | NC_037040.1 | Nassarius fraterculus | Species | 0 | 99 | 0 |
| SS102      | Nassarius alboecus | KY489098.1 | Nassarius fenestratun | Genus | 0 | 99 | 0 |
| SS144      | Nassarius coronatus | KY489997.1 | Nassarius coronatus | Species | 0 | 99 | 0 |
| SS44       | Cantharbus mollis | HQ833863.1 | Cantharbus cecilei | Genus | 0 | 97 | 0 |
| B344       | Exzinoespis menkeana | FM990997.1 | Pisonia striata | Family | 5.0E-158 | 93 | 0 |
| SS140      | Prierygia dactylus | KR087379.1 | Prierygia dactylus | Species | 0 | 100 | 0 |
| SS123      | Coralliphila nertisoidae | A2939379.1 | Coralliphila nertisoidae | Species | 0 | 98 | 0 |
| SS106      | Drupella cornus | FR85390.1 | Drupella cornus | Species | 0 | 99 | 0 |
| B356       | Opporomus fascicolor | HS08382.1 | Opporomus fascicolor | Species | 0 | 99 | 0 |
| B318       | Nacella lima | KJ903800.1 | Nacella huxeyana | Genus | 0 | 100 | 0 |
| B336       | Ocenebra inornatus endemontis | NC_034032.1 | Ocenebra falcata | Genus | 0 | 98 | 0 |
| SS81       | Mancinella echinata | HS484099.1 | Mancinella echinata | Species | 0 | 99 | 0 |
| SS70       | Mancinella siro | HS484091.1 | Mancinella gryosa | Genus | 0 | 99 | 0 |
| SS86       | Menathais tuberosa | KJ743972.1 | Menathais tuberosa | Species | 0 | 99 | 0 |
| B302       | Reishia bronni | NC_039165.1 | Thais lutostoma | Genus | 0 | 98 | 0 |
| SS75       | Reishia bronni | HS033578.1 | Thais lutostoma | Genus | 0 | 99 | 0 |
| B293       | Thais clavigera | HE584119.1 | Reishia clavigera | Species | 0 | 99 | 0 |
| SS34       | Thais clavigera | HE584119.1 | Reishia clavigera | Species | 0 | 99 | 0 |
| Sample No. | Genus | Species |
|-----------|-------|---------|
| SS116     | Vasum | turbinellum |
| B363      | Homoiodoris | japonica |
| B232      | Acmaea | pallida |
| B253      | Collisella | dorsuosa |
| B34       | Tectura | ensylva |
| B357      | Lepeta | kuragiensis |
| SS95      | Lottia | luchuana |
| B396      | Nipponacmea | radula |
| SS14      | Nipponacmea | radula |
| SS91      | Nipponacmea | schrenckii |
| SS17      | Patelloida | pygmaea |
| SS91      | Patelloida | ryukyuensis |
| B256      | Patelloida | saccharina |
| SS51      | Patelloida | saccharina |
| SS2       | Cellana | grata |
| SS209     | Cellana | grata |
| SS18      | Cellana | nigrolineata |
| SS174     | Cellana | radiata |
| SS84      | Cellana | testudinaria |
| SS3       | Cellana | toreuma |
| SS4       | Cellana | toreuma |
| B247      | Scutellastra | flexuosa |
| B257      | Siphonaria | sirius |
| B359      | Siphonaria sp. | sp. |
| SS131     | Tectus | pyramis |
| B268      | Chlorostoma | argyrostoma |
| B285      | Chlorostoma | lischkei |
| SS72      | Chlorostoma | turbonatum |
| SS73      | Tegula | pfefferi |
| SS67      | Omphalina | nigerrimus |
| SS69      | Omphalina | rusticus |
| B254      | Cantharidius | japonicus |
| B145      | Cantharidius | jessoevici |
| B332      | Lamella | coreensis |
| SS54      | Lamella | coreensis |
| SS101     | Lamella | coronata |
| B316      | Turbo | stenogyrus |

Table 1. (Continued).
| Sample No. | Species                        | GenBank ID | BLASTn result | COI        | Homology Level | E-value | Identity (%) | Gaps (%) |
|------------|--------------------------------|------------|---------------|------------|----------------|---------|---------------|----------|
| SS45       | Batillaria multiformis         | AB845520.1 | Batillaria    | Species    | 0              | 99      | 0             | 0        |
| B308       | Batillaria zonalis            | MN389845.1 | Batillaria    | Species    | 0              | 99      | 0             | 0        |
| SS104      | Planaxis sulcatus             | MT628956.1 | Planaxis      | Species    | 0              | 99      | 0             | 0        |
| B309       | Semisulcospira libertina      | KM031760.1 | Semisulcospira| Species    | 0              | 98      | 0             | 0        |
| B303       | Cerithidea cingulata          | HE600370.1 | Cerithidea    | Species    | 0              | 100     | 0             | 0        |
| SS57       | Nerita albicilla              | EU253356.1 | Nerita        | Species    | 0              | 97      | 0             | 0        |
| SS124      | Nerita helicoides             | EU732232.1 | Nerita        | Species    | 0              | 99      | 0             | 0        |
| SS59       | Nerita japonica               | EU732262.1 | Nerita        | Species    | 0              | 98      | 0             | 0        |
| SS126      | homosquamosus japonica        | MW277094.1 | Nerita        | Species    | 0              | 99      | 0             | 0        |
| SS120      | Nerita striata                | EU732235.1 | Nerita        | Species    | 0              | 100     | 0             | 0        |
| SS103      | Cassidula musselina           | MN389193.1 | Cassidula     | Species    | 0              | 99      | 0             | 0        |
| SS141      | Erronea erroneae              | MK507895.1 | Notocochlis   | -          | 0              | 96      | 0             | 0        |
| SS83       | Echinolitorina radiata        | HM500064.1 | Echinolitorina| Species    | 0              | 99      | 0             | 0        |
| SS142      | Notocochlis gualteriana       | MK507895.1 | Notocochlis   | Species    | 0              | 95      | 0             | 0        |
| SS143      | Canarium mutabile             | DQ325261.1 | Strombus      | Species    | 0              | 96      | 0             | 0        |
| SS110      | Conomurex lubuamazonck        | KY853669.1 | Conomurex     | Species    | 0              | 96      | 2             | 0        |
| B349       | Mitrella burchardi            | JN052989.1 | Mitrella      | Genus      | 0              | 100     | 0             | 0        |
| B366       | Fusinus ferrugineus           | HM180885.1 | Fusinus       | Genus      | 0              | 100     | 0             | 0        |
| SS115      | Pleuroplaca trapezium         | KT733962.1 | Pleuroplaca   | Species    | 0              | 97      | 0             | 0        |
| B355       | Nassarius comoidalis          | KY783915.1 | Nassarius     | Species    | 0              | 99      | 0             | 0        |
| B342       | Nassarius fraterculus        | KX099666.1 | Nassarius     | Species    | 0              | 100     | 0             | 0        |
| SS102      | Nassarius albecens            | KY499727.1 | Nassarius     | Species    | 0              | 99      | 0             | 0        |
| SS144      | Nassarius coronatus           | KY451287.1 | Nassarius     | Species    | 0              | 97      | 0             | 0        |
| SS44       | Cantharus molitis             | JN053007.1 | Cantharus     | Genus      | 0              | 88      | 1             | 0        |
| B344       | Ectinosipis menkeana          | KX209514.1 | Engina        | Genus      | 0              | 93      | 0             | 0        |
| SS140      | Pterygia dacytus              | KR087291.1 | Pterygia      | Species    | 0              | 96      | 0             | 0        |
| SS123      | Coralliophila nitidus         | MG917504.1 | Coralliophila | Species    | 0              | 99      | 0             | 0        |
| SS106      | Drupella coronas              | FR853843.1 | Drupella      | Species    | 0              | 99      | 0             | 0        |
| B356       | Oppomorus funiculatus         | HE854845.1 | Morula        | Species    | 0              | 99      | 0             | 0        |
| B336       | Ocinebra inornata endermontis| HM180637.1 | Lunella       | Species    | 0              | 97      | 0             | 0        |
| SS81       | Mancinella echinata           | KJ466695.1 | Mancinella    | Species    | 0              | 97      | 2             | 0        |
| SS70       | Mancinella siro               | HE853441.1 | Mancinella    | Species    | 0              | 98      | 0             | 0        |
| SS86       | Menathais taherona            | KU747972.1 | Menathais     | Species    | 0              | 99      | 0             | 0        |
| B302       | Reishia bronni                | HM180825.1 | Thais         | Genus      | 0              | 99      | 0             | 0        |
| SS75       | Reishia bronni                | NC039165.1 | Thais         | Genus      | 0              | 99      | 0             | 0        |
| E293       | Thais clavigera               | HM180819.1 | Reishia       | Species    | 0              | 98      | 1             | 0        |
| SS34       | Thais clavigera               | MH400316.1 | Reishia       | Species    | 0              | 99      | 0             | 0        |
Table 1. (Continued).

| Sample No. | Species                      | Genbank ID   | BLASTn result | Homology Level | E-value | Identity (%) | Gaps (%) |
|------------|------------------------------|--------------|---------------|----------------|---------|--------------|----------|
| SS116      | Vasmum turbinellum          | HQ834084.1   | Vasmum turbinellum | Species      | 0       | 99           | 0        |
| B363       | Homeodorida japonica        | KP9155442.1  | Homeodorida japonica | Species     | 0       | 99           | 0        |
| B323       | Acmaea pallida              | LC416617.1   | Niveotectra pallida | Species    | 0       | 97           | 0        |
| B253       | Collisella dorausta         | KM221108.1   | Lottia dorauesta | Species      | 0       | 91           | 0        |
| B334       | Tectura emylis              | MF81412.1    | Lottia inashibila | Genus        | 0       | 99           | 0        |
| B357       | Lepeta kuragiensis          | AH343974.1   | Cryptobranchia kuragiensis | Species   | 0       | 99           | 0        |
| SS95       | Lottia bouchua              | AH384371.1   | Lottia bouchua | Species     | 0       | 99           | 0        |
| B396       | Nipponacmea boninensis      | LC383956.1   | Japetulia ferea  | -            | 0       | 99           | 0        |
| SS14       | Nipponacmea radula          | KC344158.1   | Nipponacmea fuscoviridis | Genus    | 0       | 99           | 0        |
| SS91       | Nipponacmea schrenckii      | FR693394.1   | Lunella coreensis | -            | 0       | 100          |          |
| SS17       | Patelloidea pygmaea         | AH328519.1   | Patelloidea pygmaea | Species    | 0       | 98           | 0        |
| SS91       | Patelloidea ryukyuensis     | AH328520.1   | Patelloidea ryukyuensis | Species   | 0       | 99           | 0        |
| B256       | Patelloidea saccharina lanx | HM180776.3   | Patelloidea saccharina lanx | Species   | 0       | 99           | 0        |
| SS51       | Patelloidea saccharina lanx | HM180776.3   | Patelloidea saccharina lanx | Species   | 0       | 99           | 0        |
| SS2        | Cellana grata               | HM1B0722     | -              | Notoacmea schrenckii | -      | 100          |          |
| SS209      | Cellana grata               | M221072.1    | Cellana grata  | Species     | 0       | 100          |          |
| SS18       | Cellana nigrolineata        | LC600801.1   | Cellana nigrolineata | Species    | 0       | 100          |          |
| SS174      | Cellana radiata             | AH238554.1   | Cellana radiata orientalis | Species   | 0       | 99           | 0        |
| SS84       | Cellana testudinaria        | AH238563.1   | Cellana testudinaria | Species   | 0       | 99           | 0        |
| SS3        | Cellana toreuma             | HM1B0724.1   | Notoacmea schrenckii | -          | 0       | 100          |          |
| SS4        | Cellana toreuma             | KM221072.1   | Cellana grata  | Genus       | 0       | 99           | 0        |
| B247       | Scutellastra flexuosa       | KT1B4918.1   | Scutellastra flexuosa | Species   | 0       | 87           | 0        |
| B257       | Siphonaria sirius           | KF000832.1   | Siphonaria sp.  | Genus       | 0       | 99           | 0        |
| B359       | Siphonaria sp.              | MF632008.1   | Siphonaria bugisensis | Genus     | 1.0E-175 | 84           |          |
| SS113      | Tectus pyramis              | MN388932.1   | Tectus pyramis  | Species     | 6.0E-177 | 95           | 0        |
| B268       | Chlorostoma argyrostroma    | EU350144.1   | Chlorostoma lischi | Species   | 0       | 99           | 0        |
| B285       | Chlorostoma lischi          | EU350145.1   | Chlorostoma lischi | Species   | 0       | 99           | 0        |
| SS72       | Chlorostoma tarbutan        | LC413975.1   | Chlorostoma tarbutan | Species   | 0       | 99           | 0        |
| SS73       | Tegula pfefferi pfefferi   | HM1B0731.1   | Omphalus pfefferi carpenteri | Species   | 0       | 99           | 0        |
| SS67       | Omphalus nigerrimus         | HE830629.1   | Tegula aff. Argyrostroma | Genus     | 0       | 98           | 0        |
| SS69       | Omphalus rusticus           | NC. 06336.1  | Omphalus rusticus | Species     | 0       | 99           | 0        |
| B254       | Cantharidus japonicus       | EU350120.1   | Cantharidus callicroa | Genus    | 0       | 90           | 0        |
| B345       | Cantharidus jeksenensis     | AB505280.1   | Cantharidus jeksenensis | Species   | 0       | 100          |          |
| SS32       | Lunella coreensis           | MN604179.1   | Lunella coreensis | Species   | 0       | 99           | 0        |
| SS54       | Lunella coreensis           | MN604179.1   | Lunella coreensis | Species   | 0       | 99           | 0        |
| SS101      | Lunella coronata            | KC068908.1   | Lunella granulata | Genus     | 0       | 100          |          |
| B316       | Turbo stenogrypus           | AM403915.1   | Turbo stenogrypus | Species   | 0       | 99           | 0        |
**Table 2.** List the 12S-rRNA and COI gene primers used in this study.

| Gene   | Primer       | Sequence(5'-3')                  | Reference |
|--------|--------------|-----------------------------------|-----------|
| 12S    | 12Sma (F)    | CTGGGATTAGATACCTGCTTAT             | [19]      |
|        | 12S97L (F)   | AACYCAAGRACTTGGCCGTT             | [20]      |
|        | 12Smb (R)    | CAGAGAGTGCAGGGCGATTGT            | [19]      |
| COI    | LCO1490 (F)  | GGTCACAAATCATAAGATATTGG          | [22]      |
|        | LCOmod (F)   | TCTACTAACATAAGAYATGGNAC         | [24]      |
|        | HCO2198 (R)  | TAAACTCCAGGTGACAAAAATCA         | [22]      |
|        | H7005 (R)    | CCGGATCCACNACRTARTANGTRTCRTG  | [23]      |
|        | HCOmod (R)   | ACTTCTGGGTGTCCRAARAAYCARAA      | [24]      |

2.4 Sequence editing, dataset preparation, and phylogenetic analysis

Confirmed, edited, and assembled sequences were aligned using MAFFT v7 [29], the dataset containing aligned sequences was edited in Gblocks 0.91b (Online ver.) to exclude ambiguously aligned regions [30]. Furthermore, visualized in Mesquite ver. 3.61 [26-27]. Maximum likelihood (ML) phylogenetic inference was conducted in RAxML-GUI v.1.5 [31] under the GTR-GAMMA substitution model, partitioned per-gene, with 1000X bootstrap samplings carried out to assess the robustness of the obtained topology. Model selection before phylogenetic analysis was carried out in MEGA X [32]. Two Bivalves, *Pinctada fucata* (AB250258.1) and *Crassostrea gigas* (EF484878.1), were used as outgroups. Phylogenetic analysis was conducted only on the 12S-rRNA gene sequences.

2.5 Species delimitation analysis

Species delimitation analyses were conducted on the sequence datasets of COI and 12S-rRNA using the Automatic Barcode Gap Discovery (ABGD) software [33] to see if the 12S-rRNA fragment was used in our barcoding could differentiate species. The analyses used the aligned sequence data of the 12S-rRNA (length = 280 bp) and COI (length = 632 bp). Prior intraspecific divergence range (Pmin to Pmax) was set to 0.001 – 0.1, and the X value for the minimum relative gap width was set to 0.99 under the K2P model with TS/TV = 2.0.

3 Results

3.1 Sequence data acquisition

For the 12S-rRNA gene marker, we successfully amplified and thus obtained ca. 450 bp for 12Sma / 12Smb primer pairs (60 samples) and ca. 430 bp of 12S97L / 12Smb primer pairs (15 samples), making us successfully obtain the 12S-rRNA sequences of all individuals used in this study. After sequence editing and alignments, the 12S-rRNA sequence lengths used for phylogenetic analysis were 280 bp. We also successfully amplified the COI gene marker for all samples in this study (ca. 650 bp for LCO1490 / HCO2198 primer pairs = 54 samples; ca. 1100 bp for LCO1490 / H7005 primer pairs = 6 samples; ca. 650 bp for LCOmod_Kano2008 / HCOmod_Kano2008 primer pairs = 15 samples). The sequence length of the COI marker after sequence editing and alignments was 632 bp.
3.2 BLASTn searches-based DNA barcoding

BLASTn searches of the gene 12S-rRNA matched 39 species of 41 individuals (Table 1), with the sequence identities of 92%–100% and e-values of 0.00 to 6.00 e-117. Meanwhile, 30 samples of 29 species were confirmed only at the genus level (22 genera). The rest of the samples (four individuals) matched Genbank sequences at higher taxonomy (i.e., subfamily, family, and order levels). Meanwhile, for the COI gene, 47 species (50 individuals; 67%) were confirmed at the species level, with 25 individuals matched at the taxonomic levels of genera and above. Detailed results of the BLASTn searches are presented in Table 1.

3.3 Obtained phylogeny and the taxonomic placement of the samples

The phylogenetic tree obtained from the maximum likelihood phylogenetic inference performed on the 12s-rRNA sequences is shown in Figure 1. Most samples were placed along with their morphologically identified species and genera levels with relatively reasonable support values (e.g., Cellana, BS = 72%; Chlorostoma, BS = 70%; Nerita, BS = 53%). However, Nassarius and Patelloida were not monophyletic, although both genera were placed in their proper higher taxonomies (Nassarius: Bucinidae; Patelloida: Lottiidae). Placements at the higher taxonomy levels (subfamily and above) for other genera and species also placed most of them in their valid taxa, despite the lack of strong bootstrap support on many clades, and improper taxonomic placements of some samples (Figure 1), for example, members of Muricoidea, Nucella lima, and Ocnebra inornatus endermonis, were placed in Bucinidae, while Notocochlis gualteriana (Naticoidea) was placed in Muricoidea; the Cypraeoid Erronea errones was placed in Bucinidae; some members of Littorinimorpha, Erronea errones, Notocochlis gualteriana, Canarium mutabile, Conomurex luwuamu, were instead put inside Neogastropoda, and Cassidula muscelina (Ellobiida) was included in Neogastropoda.

3.4 Species delimitations of target samples

Species delimitations conducted in ABGD on the 12S-rRNA gene dataset resulted in the identification of 61 species, while the same analysis on the dataset of the COI gene identified 65 species (Table 3). Meanwhile, our samples (75 OTUs) were morphologically identified as 69 species. As shown in Table 3, the result of ABGD on both genes generally agrees with the outcome of morphological identification. However, some closely related congeneric species were not properly delimited and identified as the same species in one of the markers or both. For example, Batillaria multiformis and Batillaria zonalis were identified as one species by the 12s-rRNA, while Cellana grata and Cellana toreuma were identified as one species by both gene markers.
4 Discussion

4.1 Sequence data acquisition success rate on preserved museum samples

In this study, we tested the 12S-rRNA primers on various museum samples across the whole Gastropods. The usefulness of the COI primers has been shown in multiple previous studies [13-14], including those of ours [16-17]. In general, our results here indicate that the 12S-rRNA primers tested in this study are helpful and can amplify the 12S-rRNA sequence fragments for DNA barcoding to complement results obtained using other markers such as COI. Meanwhile, we successfully got DNA sequences of both gene markers for all samples using standard PCR protocols, even though some of the samples were old museum samples, which were stored at conditions not ideal for molecular works (e.g., room temperature storage). Therefore, our result also suggests that these primers could probably be used for museums studies, such as museum samples barcoding and studies to obtain molecular data out of old museum samples [34].
4.2 The non-exhaustiveness of Genbank for the identification of gastropods

We performed BLASTn searches of the COI and 12S-rRNA sequences obtained from the samples to see if the sequences of our taxa are present on Genbank, besides checking if our morphological identification was consistent with the sequence data on Genbank. The result of our BLASTn searches suggested that most samples only 55% of our samples were correctly identified using the 12S-rRNA and only 67% even when using COI, which is the most commonly used DNA barcoding marker [35]. This result thus underlines the problem of taxonomic bias in biodiversity observation, causing the incompleteness and/or non-exhaustiveness of the database [9]. This could be problematic for studies depending on identification based on DNA sequences only, such as eDNA and metagenomics [6].

4.3 The phylogeny is relatively well-resolved for a single marker

We also conducted a phylogenetic analysis of the 12S-rRNA sequences to see if the gene could adequately place the samples in their proper taxa. We found that while most samples were properly grouped with their conspecifics or congeners, some samples were not (Figure 1). At present, we are unable to pinpoint the cause of these misplacements, which might include sequence errors, homoplasies and long-branch attractions, and the possible lack of identifying substitutions of the 12S-rRNA fragment used in our phylogenetic analysis. Meanwhile, classification at higher taxonomy, in general, is congruent with the recently proposed gastropod systematics [36-37]. This result is generally also in agreement with the result of our preliminary study [38]. However, detailed interrelationships did not agree entirely, and the statistical supports were low in most nodes of higher taxonomy. It is expected since the interrelationships among higher taxa (above genus) cannot usually be resolved using only single-gene data [16-17,38-39].

4.4 The 12S-rRNA marker was able to delimit most species in this study

We also conducted species delimitation analyses on the COI and 12S-rRNA to confirm if both characteristics could correctly identify/delimit the species of the samples, as identified morphologically. There was 69 morphospecies (out of 75 individuals) in our samples, which were identified by professional taxonomists/curators, which were also co-authors, of this study. Interestingly, however, both markers were unable to completely delimit all morphospecies (COI = 65 species, but 12S = 71 species), apparently having difficulties differentiating closely related congeneric species (Table 3).

The differences in species delimitation could probably be attributed to differences in the substitution rates of each taxon, which might be related to their different biology. This was also suggested by the species delimitation results of both genes before the removal of ambiguously aligned regions by GBlocks. These results indicated that the removal of ambiguously aligned regions might affect the detection of sequence diversity due to the removal of possible informative areas [40]. However, all in all, our result also indicated that both markers could delimit the samples at least at the genus level.

4.5 General conclusion and future directions

Our present results of species delimitation analysis, phylogenetic analysis, and BLASTn searches suggest that the short fragment of the 12S-rRNA gene used in this study is useful and effective enough to delimit various gastropod species, and thus useful for DNA barcoding and metabarcoding (eDNA studies) of shelled marine gastropods. The marker could thus be used to complement other DNA barcoding markers such as COI and 18S-rRNA. DNA
barcoding using multiple markers would allow researchers to capture a complete snapshot of biodiversity and avoid the numerous possible pitfalls caused by using only a single marker [16-17, 41-44].

Moreover, because our study presented appropriately used curated museum samples, our sequence data would become an essential addition to the reference database for future studies. Therefore, in the future, we will register our sequence data to an adequately curated database such as Genbank or DDBJ. We will also continue our study by testing more prospective markers on more properly curated museum samples of shelled marine gastropods from Japan to provide a comprehensive reference sequence database for further studies involving DNA barcoding, metabarcoding, and e-DNA.

5 Acknowledgements

We would like to thank previous and present members of the Setiamarga Lab at National Institute of Technology, Wakayama College, and the Sasaki Lab at The University Museum, The University of Tokyo, for constant support during this study. This study was financially supported by Nakatsuji Foresight Foundation FY2018 Research Grants for Basic Sciences [to DHES], FY2019 JSPS KAKENHI Grant-in-Aid for Scientific Research (C) Grant No. 19K12424 [to DHES and TS], and JSPS KAKENHI for Challenging Research (Exploratory) Grant No. 19K21646 [to TS and DHES].
Barcoding using multiple markers would allow researchers to capture a complete snapshot of biodiversity and avoid the numerous possible pitfalls caused by using only a single marker [16-17, 41-44]. Moreover, because our study presented appropriately used curated museum samples, our sequence data would become an essential addition to the reference database for future studies. Therefore, in the future, we will register our sequence data to an adequately curated database such as Genbank or DDBJ. We will also continue our study by testing more prospective markers on more properly curated museum samples of shelled marine gastropods from Japan to provide a comprehensive reference sequence database for further studies involving DNA barcoding, metabarcoding, and e-DNA.

### Acknowledgements

We would like to thank previous and present members of the Setiamarga Lab at National Institute of Technology, Wakayama College, and the Sasaki Lab at The University Museum, The University of Tokyo, for constant support during this study. This study was financially supported by Nakatsuji Foresight Foundation FY2018 Research Grants for Basic Sciences [to DHES], FY2019 JSPS KAKENHI Grant-in-Aid for Scientific Research (C) Grant No. 19K12424 [to DHES and TS], and JSPS KAKENHI for Challenging Research (Exploratory) Grant No. 19K21646 [to TS and DHES].

### Table 3. Results of analysis of 12S-rRNA and COI using ABGD.

| #  | Sample No. | Morphological | 12S species | COI species |
|----|------------|---------------|-------------|-------------|
| 1  | SS45       | Batillaria multiformis | 12S_species 1 | COI_species 1 |
| 2  | B308       | Batillaria zonalis | 12S_species 2 | COI_species 2 |
| 3  | SS104      | Planaxis sulcatus | 12S_species 3 | COI_species 3 |
| 4  | B309       | Semitapes cingulata | 12S_species 4 | COI_species 4 |
| 5  | SS77       | Cerithidea cingulata | 12S_species 5 | COI_species 5 |
| 6  | SS124      | Nerita albicilla | 12S_species 6 | COI_species 6 |
| 7  | SS120      | Nerita helicinoides | 12S_species 7 | COI_species 7 |
| 8  | SS59       | Nerita striata | 12S_species 8 | COI_species 8 |
| 9  | SS126      | Homosapta japonica | 12S_species 9 | COI_species 9 |
| 10 | SS103      | Cassidula mustelina | 12S_species 10 | COI_species 10 |
| 11 | SS142      | Erronea erronea | 12S_species 11 | COI_species 11 |
| 12 | SS141      | Notocochlis guateriana | 12S_species 12 | COI_species 12 |
| 13 | SS143      | Echinolittorina radiata | 12S_species 13 | COI_species 13 |
| 14 | SS130      | Canarium mutabile | 12S_species 14 | COI_species 14 |
| 15 | SS110      | Conomurex lahaanuus | 12S_species 15 | COI_species 15 |
| 16 | B349       | Mitrella burchardi | 12S_species 16 | COI_species 16 |
| 17 | B366       | Fusinus ferrugineus | 12S_species 17 | COI_species 17 |
| 18 | SS115      | Pleuroloca trapezium | 12S_species 18 | COI_species 18 |
| 19 | SS142      | Nassarius fraterculi | 12S_species 19 | COI_species 19 |
| 20 | SS102      | Cassidula mustelina | 12S_species 20 | COI_species 20 |
| 21 | SS144      | Nassarius albecens | 12S_species 21 | COI_species 21 |
| 22 | SS144      | Nassarius diriatus | 12S_species 22 | COI_species 22 |
| 23 | SS144      | Canarium mollis | 12S_species 23 | COI_species 23 |
| 24 | SS144      | Echinolittorina radiata | 12S_species 24 | COI_species 24 |
| 25 | SS144      | Pleuroloca trapezium | 12S_species 25 | COI_species 25 |
| 26 | SS144      | Conomurex lahaanuus | 12S_species 26 | COI_species 26 |
| 27 | SS144      | Mitrella burchardi | 12S_species 27 | COI_species 27 |
| 28 | SS144      | Nassarius fraterculi | 12S_species 28 | COI_species 28 |
| 29 | SS144      | Nassarius fraterculi | 12S_species 29 | COI_species 29 |
| 30 | SS144      | Nassarius fraterculi | 12S_species 30 | COI_species 30 |
| 31 | SS144      | Nassarius fraterculi | 12S_species 31 | COI_species 31 |
| 32 | SS144      | Nassarius fraterculi | 12S_species 32 | COI_species 32 |
| 33 | SS144      | Nassarius fraterculi | 12S_species 33 | COI_species 33 |
| 34 | SS144      | Nassarius fraterculi | 12S_species 34 | COI_species 34 |
| 35 | SS144      | Nassarius fraterculi | 12S_species 35 | COI_species 35 |
| 36 | SS144      | Nassarius fraterculi | 12S_species 36 | COI_species 36 |
| 37 | SS144      | Nassarius fraterculi | 12S_species 37 | COI_species 37 |
Table 3. (Continued).

| #  | Sample No. | Morphological | 12S       | COI    |
|----|-------------|---------------|-----------|--------|
| 41 | B323        | Acmaea pallida| 12S species 33 | COI species 38 |
| 42 | B253        | Collisella dorsuosa | 12S species 34 | COI species 39 |
| 43 | B334        | Tectura emylia | 12S species 35 | COI species 40 |
| 44 | B357        | Lepeta kuragensis | 12S species 36 | COI species 41 |
| 45 | SS95        | Lottia luchuana | 12S species 37 | COI species 42 |
| 46 | B396        | Nipponacmea boninensis | 12S species 38 | COI species 43 |
| 47 | SS14        | Nipponacmea radula | 12S species 39 | COI species 44 |
| 48 | SS17        | Patelloida pygmaea | 12S species 40 | COI species 45 |
| 49 | SS91        | Patelloida ryukyuensis | 12S species 41 | COI species 46 |
| 50 | B256        | Patelloida saccharina | 12S species 42 | COI species 47 |
| 51 | SS51        | Cellana grata | 12S species 43 | COI species 48 |
| 52 | SS2         | Cellana tereuma | 12S species 44 | COI species 49 |
| 53 | SS3         | Cellana tereuma | 12S species 45 | COI species 50 |
| 54 | SS209       | Cellana grata | 12S species 46 | COI species 51 |
| 55 | SS5         | Cellana tereuma | 12S species 47 | COI species 52 |
| 56 | SS18        | Cellana nigrolinesata | 12S species 48 | COI species 53 |
| 57 | SS174       | Cellana radiata | 12S species 49 | COI species 54 |
| 58 | SS84        | Cellana testudinaria | 12S species 50 | COI species 55 |
| 59 | B247        | Scutellastra fuscuna | 12S species 51 | COI species 56 |
| 60 | B257        | Siphonaria siri | 12S species 52 | COI species 57 |
| 61 | B359        | Siphonaria sp. | 12S species 53 | COI species 58 |
| 62 | SS131       | Tectus pyramis | 12S species 54 | COI species 59 |
| 63 | B268        | Chlorostoma argyrostroma lischkei | 12S species 55 | COI species 60 |
| 64 | SS72        | Chlorostoma turbinatum | 12S species 56 | COI species 61 |
| 65 | B285        | Chlorostoma lischkei | 12S species 57 | COI species 62 |
| 66 | SS73        | Tegula pfeifferi pfeifferi | 12S species 58 | COI species 63 |
| 67 | SS69        | Omphalina rusticus | 12S species 59 | COI species 64 |
| 68 | SS67        | Omphalina nigerrimus | 12S species 60 | COI species 65 |
| 69 | B254        | Cantharidus japonicus | 12S species 61 | COI species 66 |
| 70 | B345        | Cantharidus jessensis | 12S species 62 | COI species 67 |
| 71 | B391        | Nipponacmea schrenkii | 12S species 63 | COI species 68 |
| 72 | B332        | Lumella coreensis | 12S species 64 | COI species 69 |
| 73 | SS54        | Lumella coronata | 12S species 65 | COI species 70 |
| 74 | SS101       | Lumella coronata | 12S species 66 | COI species 71 |
| 75 | B316        | Turbo stenogenus | 12S species 67 | COI species 72 |
## References

1. W. Appelton, S.T. Ahyong, G. Anderson, M.V. Angel, T. Artois, N. Bailly, et al., Curr. Biol. 22, 23 (2012)
2. W.F. Ponder, D.R. Lindberg, J.M. Ponder, *Biology and evolution of the Mollusca: volume two* (CRC Press, Florida, 2019)
3. J.C. Orr, V.J. Fabry, O. Aumont, L. Bopp, S.C. Doney, R.A. Feely, et al., Nature 437, 601–686 (2005)
4. R. Rodolfo-Metalpa, F. Houbaëque, É. Tambahutte, F. Boisson, C. Gabbini, F.P. Patti, et al, Nature Clim. Change 1, 308–312 (2011)
5. P.D.N. Hebert, A. Cywinska, S.L. Ball, J.R. deWaard, Royal Soc. B. 270, 1512 (2003)
6. P.F. Thomsen, E. Willerslev, Biol. Conserv. 183, 4–18 (2015)
7. K. Deiner, H.M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, et al, Mol. Ecol. 26, 21 (2017)
8. M.E. Cristescu, P.D.N. Hebert, Annu. Rev. Ecol. Evol. Syst. 49, 209-230 (2018)
9. J. Troudet, P. Grandcolas, A. Blin, R. Vignes-Lebbe, F. Legendre, Sci. Rep. 7, 9132 (2017)
10. R.J. Machida, Y. Hashiguchi, M. Nishida, S. Nishida, BMC Genom. 10, 438 (2009)
11. R.D.M. Page, Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371, 1702 (2016)
12. O. Folmer, M. Black, W. Hoah, R. Lutz, R. Vrijenhoek, Mar. Biotechnol. 3, 294–299 (1994)
13. Y. Kano, Zool. Scr. 37, 1–20 (2008)
14. LMS. Borges, C. Hollatz, J. Lobo, A.M. Cunha, A.P. Vilela, G. Calado, et al, Sci. Rep. 6, 20226 (2016)
15. S. Zou, Q. Li, L. Kong, H. Yu, X. Zheng, PLoS ONE 6, 10, (2011)
16. N. Nakaji, S. Iwamoto, S. Teruya, M. Kusube, A. Kosaka, T. Sasaki, et al, SEE 2018, 28–44 (2018)
17. D.H.E. Setiamarga, N. Nakaji, S. Iwamoto, S. Teruya, T. Sasaki, Int. J. GEOMATE, 17, 62 (2019)
18. D.J. Colgan, A. McLauchlan, G.D.F. Wilison, S.P. Livingston, G.D. Edgecombe, J. Macaranas, et al., Aust. J. Zool. 46, 5 (1998)
19. V. Koufopanou, D.G. Reid, S.A. Ridgway, R.H. Thomas, Mol. Biol. Evol. 70, 1 (1999)
20. T. Nakano, T. Ozawa, J. Molluscan Stud. 70, 1 (2004)
21. T. Okutani, Marine mollusks in Japan (in Japanese) (The Association of Japanese University Press, Tokyo, 2017)
22. O. Folmer, M. Black, W. Hoah, R. Lutz, R Vrijenhoek, Mar. Biotechnol. 3, 5 (1994)
23. M.S. Hafner, P. D. Sudman, F.X. Villablanca, T.A. Spradling, J.W. Demastes, S.A. Nadler, Science 265, 5175 (1994)
24. Y. Kano, Zool. Scr. 37, 1 (2008)
25. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. 215, 3 (1990)
26. W. Maddison, D. Maddison, Evolution 11, 5 (2009)
27. W.P. Maddison, D.R. Maddison, *Mesquite: a modular system for evolutionary analysis*, Version 3.61 http://www.mesquiteproject.org (2018)
28. M.W. Davis, *ApE-A plasmid editor*, https://jorgensen.biology.utah.edu/wayned/ape/
29. K. Katoh, J. Rozewicki, K.D. Yamada, Brief. Bioinf. 20, 4 (2019)
30. J. Castresana, Mol. Biol. Evol. 17, 4 (2002)
31. D. Silvestro, I. Michalak, Org. Divers. Evol. 12, 4 (2012)
32. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mol. Biol. Evol. 35, 6 (2018)
33. N. Puillandre, A. Lambert, S. Brouillet, G.J.M.E. Achaz, Mol. Ecol. 21, 8 (2012)
34. D.H.E. Setiamarga, R. Shiba, Y. Kamito, M. Yamamoto, N.N.B. Razali, M. Arai, et al, Zoosymposia 15, 129-140 (2019)
35. P.D.N. Hebert, M.Y. Stoeckle, T.S. Zemlak, C.M. Francis, C. Godfray, PLOS Biol. 2, 10 (2004)
36. T. Sasaki, *Malacology* (In Japanese) (The University of Tokyo Press, Tokyo, 2010)
37. T.J. Cunha, G. Giribet, Proc. Royal Soc. B, 286, 1898 (2019)
38. S. Nakashima, N. Nakaji, D.V. Tu, T. Sasaki, D.H.E. Setiamarga, *Proc. 28th OES*, 1-6 (2020)
39. A. Som, Brief. Bioinf. 16, 3 (2015)
40. F. Lutzoni, P. Wagner, V. Reeb, S. Zoller, Syst. Biol. 49, 4 (2000)
41. H. Song, J.E. Buhay, M.F. Whiting, K.A. Crandall, *Proc. Natl. Acad. Sci.*, 105 (2008)
42. M. Miya, R.O. Gotoh, T. Sado, Fish. Sci. 86, 1-32 (2020)
43. E. Valsecchi, J. Bylemans, S.J. Goodman, R. Lombardi, I. Carr, L. Castellano, et al., Environ. DNA 2, 4 (2020)
44. F. De Mattia, R. Gentili, I. Bruni, A. Galimberti, S. Sgorbati, M. Casiraghi, et al, Bot. J. Linn. Soc. 169, 3 (2012)