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 Nakashima^{1,4}, Naoto Tomobuchi¹, Takenori Sasaki³, and Davin H. E. Setiamarga^{1,3,*}

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,

⁴Present address: Graduate School of Frontier Science, The University of Tokyo, Kashiwa-no-h 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 eDNAbased 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 vouchered 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 vouchered 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

Counts Ma	Sa sa ing		128						
Sample No.	Species	Genebank ID	BLASTn result	Homology Level	E-value	Identity (%)	Gaps%		
SS45	Batillaria multiformis	NC_047187.1	Batillaria attramentaria	Genus	0	98	0		
B308	Batillaria zonalis	HQ833855.1	Batillaria cumingi	Genus	0	98	0		
SS104	Planaxis sulcatus	HQ833854.1	Planaxis sulcatus	Species	0	99	0		
B309	Semisulcospira libertina	NC 037771.1	Semisulcospira coreana	Ĝenus	3.0E-160	93	0		
B303	Cerithidea cingulata	AB535193.1	Cerithidea djadjariensis	Genus	8.0E-126	88	3		
SS57	Nerita albicilla	LC215360.1	Nerita albicilla	Species	0	99	0		
SS124	Nerita helicinoides	MT161611.1	Nerita chamaeleon	Ĝenus	6.0E-162	94	0		
SS59	Nerita japonica	LC565707.1	Nerita japonica	Species	0	99	0		
SS126	homoiodoris japonica	MT161611.1	Nerita chamaeleon	Ĝenus	1.0E-173	96	0		
SS120	Nerita striata	KF728888.1	Nerita fulgurans	Genus	2.0E-162	94	0		
SS103	Cassidula mustelina	KJ920319.1	Cassidula nucleus	Species	2.0E-171	96	0		
SS141	Erronea errones	HQ833858.1	Erronea errones	Species	0	98	0		
SS83	Echinolittorina radiata	AJ623151.1	Echinolittorina radiata	Species	0	99	0		
SS142	Notocochlis gualteriana	MK507895.1	Notocochlis sp.	Ĝenus	0	99	0		
SS143	Canarium mutabile	MW244820.1	Tridentarius dentatus	Family	1.0E-163	94	0		
SS110	Conomurex luhuanus	KY853669.1	Conomurex luhuanus	Species	0	99	0		
B349	Mitrella burchardi	HQ833864.1	Mitrella bicincta	Ĝenus	0	99	0		
B366	Fusinus ferrugineus	NC 045906.1	Fusinus longicaudus	Genus	0	98	0		
SS115	Pleuroploca trapezium	MN322355.1	Turrilatirus turritus	Subfamily	2.0E-137	90	1		
B355	Nassarius conoidalis	NC 041310.1	Nassarius conoidalis	Species	0	99	0		
B342	Nassarius fraterculus	NC 037604.1	Nassarius fraterculus	Species	0	99	0		
SS102	Nassarius albescens	KY489008.1	Nassarius fenistratus	Ĝenus	0	99	0		
SS144	Nassarius coronatus	KY488995.1	Nassarius coronatus	Species	0	99	0		
SS44	Cantharus mollis	HQ833883.1	Cantharus cecillei	Ĝenus	0	97	0		
B344	Enzinopsis menkeana	FM999097.1	Pisania striata	Family	5.0E-158	93	0		
SS140	Ptervgia dactylus	KR087379.1	Ptervgia dactylus	Species	0	100	0		
SS123	Coralliophila neritoidea	AJ293679.1	Coralliophila neritoidea	Species	0	98	0		
SS106	Drupella cornus	FR853980.1	Drupella cornus	Species	0	99	0		
B356	Oppomorus funiculatus	HE583824.1	Morula funiculata	Species	0	99	0		
B318	Nucella lima	KJ093800.1	Nucella heyseana	Ĝenus	0	100	0		
B336	Ocenebra inornatus endermonis	NC 046052.1	Ocinebrellus falcatus	Genus	0	98	0		
SS81	Mancinella echinata	HE584089.1	Mancinella echinata	Species	0	99	0		
SS70	Mancinella siro	HE584090.1	Mancinella grossa	Ġenus	0	99	0		
SS86	Menathais tuberosa	KU747972.1	Menathais tuberosa	Species	0	99	0		
B302	Reishia bronni	NC 039165.1	Thais luteostoma	Ĝenus	0	98	0		
SS75	Reishia bronni	HQ833878.1	Thais luteostoma	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		

Table 1. (Continued).

Samula No	Supprise	125					
Sample No.	Species	Genebank ID	BLASTn result	Homology Level	E-value	Identity (%)	Gaps%
SS116	Vasum turbinellum	HQ833909.1	Vasum turbinellus	Species	0	99	0
B363	Homoiodoris japonica	KP635442.1	Homoiodoris japonica	Species	0	99	0
B323	Acmaea pallida	MT370382.1	Niveotectura pallida	Species	5.0E-132	93	0
B253	Collisella dorsuosa	AB106454.1	Lottia dorsuosa	Species	1.0E-152	99	0
B334	Tectura emvdia	MT370375.1	Lottia instabilis	Ĝenus	0	99	0
B357	Lepeta kuragiensis	AB238235.1	Cryptobranchia kuragiensis	Species	1.0E-149	99	0
SS95	Lottia luchuana	AB106453.1	Lottia luchuana	Species	2.0E-172	99	0
B396	Nipponacmea boninensis	KU316594.1	Nipponacmea radula	Ĝenus	0	99	0
SS14	Nipponacmea radula	KU316566.1	Nipponacmea fuscoviridis	Genus	0	100	0
S391	Nipponacmea schrenkii	KU316582.1	Nipponacmea schrenckii	Species	0	96	0
SS17	Patelloida pygmaea	AB106436.1	Patelloida pygmaea	Species	7.0E-161	99	0
SS91	Patelloida rvukvuensis	AB238280.1	Patelloida ryukyuensis	Species	2.0E-166	100	0
B256	Patelloida saccharina lanx	LC142820.1	Nipponacmea boninensis	Family	1.0E-163	99	0
SS51	Patelloida saccharina lanx	AB106439.1	Patelloida saccharina lanx	Species	8.0E-171	100	0
SS2	Cellana grata	GQ455887.1	Cellana toreuma	Ĝenus	0	100	0
SS209	Cellana grata	AB106427.1	Cellana grata	Species	3.0E-134	99	0
SS18	Cellana nigrolineata	LC600801.1	Cellana nigrolineata	Species	2.0E-141	99	0
SS174	Cellana radiata	AB106430.1	Cellana radiata orientalis	Species	2.0E-152	99	0
SS84	Cellana testudinaria	AB106431.1	Cellana testudinaria	Species	1.0E-153	100	0
SS3	Cellana toreuma	GQ455887.1	Cellana toreuma	Species	0	99	0
SS4	Cellana toreuma	GQ455860.1	Cellana mazatlandica	Ĝenus	0	99	0
B247	Scutellastra flexuosa	AF058183.1	Scutellastra flexuosa	Species	6.0E-117	92	1
B257	Siphonaria sirius	KF001136.1	Siphonaria sp.	Ġenus	1.0E-172	100	0
B359	Siphonaria sp.	KF001069.1	Siphonaria sp.	Genus	8.0E-81	84	3
SS131	Têctus pyramis	MF138911.1	Tectus pyramis	Species	0	97	0
B268	Chlorostoma argyrostoma lischkei	HE800683.1	Chlorostoma lischkei	Ĝenus	5.0E-167	100	0
B285	Chlorostoma lischkei	HE800683.1	Chlorostoma lischkei	Species	0	99	0
SS72	Chlorostoma turbinatum	HE800683.1	Chlorostoma lischkei	Ĝenus	0	99	0
SS73	Tegula pfeifferi pfeifferi	NC 056356.1	Omphalius rusticus	Genus	0	98	0
SS67	Omphalius nigerrimus	NC 031862.1	Omphalius nigerrimus	Species	0	99	0
SS69	Omphalius rusticus	AF080631.1	Tegula rusticus	Species	0	99	0
B254	Cantharidus japonicus	AB505369.1	Cantharidus jessoensis	Ĝenus	0	93	1
B345	Cantharidus jessoensis	AB505369.1	Cantharidus jessoensis	Species	0	99	0
B332	Lunella coreensis	MN604179.1	Lunella correensis	Species	0	99	0
SS54	Lunella coreensis	MN604179.1	Lunella correensis	Species	0	99	0
SS101	Lunella coronata	KX298890.1	Lunella granulata	Ĝenus	0	99	0
B316	Turbo stenogyrus	FR695555.1	Turbo kenwilliamsi	Genus	0	94	1

Table 1. (Continued).

Samula No	Supprise		COI					
Sample No.	Species	Genebank ID	BLASTn result	Homology Level	E-value	Identity (%)	Gaps%	
SS45	Batillaria multiformis	AB845820.1	Batillaria multiformis	Species	0	99	0	
B308	Batillaria zonalis	MN389045.1	Batillaria cumingii	Genus	0	99	0	
SS104	Planaxis sulcatus	MT620956.1	Planaxis sulcatus	Species	0	99	0	
B309	Semisulcospira libertina	KM031760.1	Semisulcospira libertina	Species	0	98	0	
B303	Cerithidea cingulata	HE680370.1	Cerithideopsilla cingulata	Species	0	100	0	
SS57	Nerita albicilla	EU253356.1	Nerita albicilla	Species	0	97	0	
SS124	Nerita helicinoides	EU732252.1	Nerita helicinoides	Species	0	99	0	
SS59	Nerita iaponica	EU732262.1	Nerita iaponica	Species	0	98	0	
SS126	homoiodoris iaponica	MW277894.1	Nerita polita	Species	0	99	0	
SS120	Nerita striata	EU732335.1	Nerita undata	Genus	Ó	100	Ó	
SS103	Cassidula mustelina	MN389193.1	Cassidula nucleus	Species	õ	99	õ	
SS141	Erronea errones	MK 507895.1	Notocochlis sp		õ	96	õ	
SS83	Echinolittorina radiata	HM560004.1	Echinolittorina radiata	Species	õ	99	õ	
SS142	Notocochlis gualteriana	MK 507895.1	Notocochlis sp	Genus	õ	95	õ	
SS143	Canarium mutahile	DO525218.1	Strombus mutabilis	Species	õ	96	õ	
SS110	Conomurex luhuanus	KY8536691	Conomurey lubuanus	Species	ŏ	96	2	
B349	Mitrella burchardi	JN052989.1	Mitrella hicincta	Genus	ŏ	100	õ	
B366	Fusinus ferrugineus	HM180585.1	Fusinus longicaudus	Genus	õ	100	õ	
SS115	Pleuronloca tranezium	KT753962.1	Pleuronloca tranezium	Species	õ	97	õ	
B355	55 Nassarius conoidalis		Nassarius conoidalis	Species	õ	99	õ	
B342	Nassarius fraterculus	KX069666.1	Nassarius fraterculus	Species	õ	100	õ	
SS102	Nassarius albescens	KY4997271	Nassarius albescens	Species	ŏ	99	ŏ	
SS144	Nassarius coronatus	KY4512871	Nassarius coronatus	Species	ŏ	97	ŏ	
SS44	Cantharus mollis	IN053007.1	Cantharus cecillei	Genus	3 0E-89	88	ĩ	
B344	Enzinonsis menkeana	KX5195141	Engina lanceolata	Genus	0	93	ò	
\$\$140	Ptervaja dactylus	KR087291.1	Ptervaia dactylus	Species	ő	96	ő	
\$\$123	Corallionhila neritoidea	MG917504.1	Corallionhila violacea	Species	ő	99	ő	
\$\$106	Drunella corrus	FR853843 1	Drupella corrus	Species	ő	99	ő	
B356	Oppomorus funiculatus	HE584045.1	Morula funiculata	Species	Ő	00	ő	
B318	Nucalla lima	K 1003701 1	Nucalla haysaana	Genus	0	00	0	
B336	Ocanabra inormatus andarmonis	HM180657.1	I unalla coroansis	Genus	0	00	0	
\$\$81	Mancinella echinata	KC466605.1	Mancinalla achinata	Species	0	97	2	
SS70	Mancinella siro	HE584344.1	Mancinella ectinala Mancinella siro	Species	0	97	2	
5586	Manathais tubarosa	KU747072 1	Manathais tubarosa	Species	0	00	0	
B302	Paishia hronni	HM180825.1	Thais lutaostoma	Genus	0	00	0	
\$\$75	Reishia bronni	NC 0301651	Thais huteostoma	Genus	0	00	0	
B203	Thais claviaara	HM180810 1	Paishia claviaara	Spacias	0	99	1	
5524	Thais clavigera	MU/00216.1	Poishia elavigera	Species	0	20	0	
3334	inais ciuvigera	MI1400310.1	Keisnia ciavigera	species	U	79	U	

Table 1. (Continued).

Samula No.	Supping.	COI						
Sample No.	Species	Genebank ID	BLASTn result	Homology Level	E-value	Identity (%)	Gaps%	
SS116	Vasum turbinellum	HQ834084.1	Vasum turbinellus	Species	0	99	0	
B363	Homoiodoris japonica	KP635442.1	Homoiodoris japonica	Species	0	99	0	
B323	Acmaea pallida	LC416617.1	Niveotectura pallida	Species	0	97	0	
B253	Collisella dorsuosa	KM221108.1	Lottia dorsuosa	Species	0	91	0	
B334	Tectura emydia	MT814212.1	Lottia instabilis	Ĝenus	0	99	0	
B357	Lepeta kuragiensis	AB543974.1	Cryptobranchia kuragiensis	Species	0	99	0	
SS95	Lottia luchuana	AB238471.1	Lottia luchuana	Species	0	99	0	
B396	Nipponacmea boninensis	LC383956.1	Japeuthria ferrea	· -	0	99	0	
SS14	Nipponacmea radula	KC844158.1	Nipponacmea fuscoviridis	Genus	0	99	0	
S391	Nipponacmea schrenkii	FR693994.1	Lunella coreensis	-	0	100	0	
SS17	Patelloida pygmaea	AB238519.1	Patelloida pygmaea	Species	0	98	0	
SS91	Patelloida ryukyuensis	AB238520.1	Patelloida ryukyuensis	Species	0	99	0	
B256	Patelloida saccharina lanx	HM180776.1	Patelloida saccharina lanx	Species	0	99	0	
SS51	Patelloida saccharina lanx	HM180776.1	Patelloida saccharina lanx	Species	0	99	0	
SS2	Cellana grata	HM180722.1	Notoacmea schrenckii	· -	0	100	0	
SS209	Cellana grata	KM221072.1	Cellana grata	Species	0	98	0	
SS18	Cellana nigrolineata	LC600801.1	Cellana nigrolineata	Species	0	100	0	
SS174	Cellana radiata	AB238554.1	Cellana radiata orientalis	Species	0	99	0	
SS84	Cellana testudinaria	AB238563.1	Cellana testudinaria	Species	0	99	0	
SS3	Cellana toreuma	HM180724.1	Notoacmea schrenckii	· -	0	100	0	
SS4	Cellana toreuma	KM221072.1	Cellana grata	Genus	0	99	0	
B247	Scutellastra flexuosa	KT149318.1	Scutellastra flexuosa	Species	0	87	0	
B257	Siphonaria sirius	KF000832.1	Siphonaria sp.	Ĝenus	0	99	0	
B359	Sîphonaria sp.	MF652008.1	Siphonaria fuegiensis	Genus	1.0E-175	84	0	
SS131	Têctus pyramis	MN388983.1	Tectus pyramis	Species	6.0E-177	95	0	
B268	Chlorostoma argyrostoma lischkei	EU530144.1	Chlorostoma lischkei	Ĝenus	0	99	0	
B285	Chlorostoma lischkei	EU530145.1	Chlorostoma lischkei	Species	0	99	0	
SS72	Chlorostoma turbinatum	LC413975.1	Chlorostoma turbinatum	Species	0	99	0	
SS73	Tegula pfeifferi pfeifferi	HM180731.1	Omphalius pfeifferi carpenteri	Species	0	99	0	
SS67	Omphalius nigerrimus	HE800629.1	Tegula aff. Argyrostoma	Ĝenus	0	98	0	
SS69	Omphalius rusticus	NC 056356.1	Omphalius rusticus	Species	0	99	0	
B254	Cantharidus japonicus	EU530120.1	Cantharidus callichroa	Ĝenus	0	90	0	
B345	Cantharidus jessoensis	AB505280.1	Cantharidus jessoensis	Species	0	100	0	
B332	Lunella coreensis	MN604179.1	Lunella correensis	Species	0	99	0	
SS54	Lunella coreensis	MN604179.1	Lunella correensis	Species	0	99	0	
SS101	Lunella coronata	KX298890.1	Lunella granulata	Ĝenus	0	99	0	
B316	Turbo stenogyrus	AM403915.1	Turbo stenogyrus	Species	0	99	0	

Gene	Primer	Sequence(5'-3')	Reference
12S	12Sma (F)	CTGGGATTAGATACCCTGTTAT	[19]
	12S97L (F)	AACYCAAAGRACTTGGCGGT	[20]
	12Smb (R)	CAGAGAGTGACGGGCGATTTGT	[19]
COI	LCO1490 (F)	GGTCAACAAATCATAAAGATATTGG	[22]
	LCOmod (F)	TCTACTAATCATAAGGAYATYGGNAC	[24]
	HCO2198 (R)	TAAACTTCAGGGTGACCAAAAAATCA	[22]
	H7005 (R)	CCGGATCCACNACRTARTANGTRTCRTG	[23]
	HCOmod (R)	ACTTCTGGGTGTCCRAARAAYCARAA	[24]

Table 2. List the	12S-rRNA	and COI 🤉	gene primers	used in this study.
I WOIC I. DISC HIC	120 11011		Serie primero	abea m mb braay.

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 (*Nasarius*: Buccinoidea; *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 *Ocenebra inornatus endermonis*, were placed in Buccinoidea, while *Notocochlis gualteriana* (Naticoidea) was placed in Muricoidea; the Cypraeoid *Erronea errones* was placed in Buccinoidea; some members of Littorinimorpha, *Erronea errones*, *Notocochlis gualteriana*, *Canarium mutabile*, *Conomurex luhuanus*, were instead put inside Neogastropoda, and *Cassidula mustelina* (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.



Fig. 1. Maximum likelihood phylogenetic tree of our samples based on 12S-rRNA sequences obtained in this study. Numbers on a branch denote bootstrap value of an adjacent node. Bootstrap values lower than 41% are not written, while fully supported nodes are denoted by ■.

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 data base [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].

#	Sample No.	Morphological	12S	COI
1	SS45	Batillaria multiformis	128	COI species 1
2	B308	Batillaria zonalis	125_species 1	COI species 2
3	SS104	Planaxis sulcatus	12S species 2	COI species 3
4	B309	Semisulcospira libertina	12S species 3	COI species 4
5	B303	Cerithidea cingulata	12S species 4	COI species 5
6	SS57	Nerita albicilla	12S species 5	COI species 6
7	SS124	Nerita helicinoides	129	COI species 7
8	SS120	Nerita striata	125_species o	COI species 8
9	SS59	Nerita japonica	12S species 7	COI species 9
10	SS126	Homoiodoris japonica	12S species 8	COI species 10
11	SS103	Cassidula mustelina	12S species 9	COI species 11
12	SS141	Erronea errones	12S species 10	COL amorian 12
13	SS142	Notocochlis gualteriana	12S species 11	COI_species 12
14	SS83	Echinolittorina radiata	12S species 12	COI species 13
15	SS143	Canarium mutabile	12S species 13	COI species 14
16	SS110	Conomurex luhuanus	12S species 14	COI species 15
17	B349	Mitrella burchardi	12S species 15	COI species 16
18	B366	Fusinus ferrugineus	12S species 16	COI species 17
19	SS115	Pleuroploca trapezium	12S_species 17	COI_species 18
20	B342	Nassarius fraterculus	12S species 18	COI species 19
21	B355	Nassarius conoidalis		COI species 20
22	SS102	Nassarius albescens	12S_species 19	COI species 21
23	SS144	Nassarius coronatus		COI_species 22
24	SS44	Cantharus mollis	12S species 20	COI species 23
25	B344	Enzinopsis menkeana	12S species 21	COI species 24
26	SS140	Pterygia dactylus	12S species 22	COI species 25
27	SS123	Coralliophila neritoidea	12S species 23	COI species 26
28	SS106	Drupella cornus	12S species 24	COI species 27
29	B356	Oppomorus funiculatus	12S species 25	COI species 28
30	B318	Nucella lima	128 spacias 26	COI species 29
31	B336	Ocenebra inornatus endermonis	125_species 20	COI species 30
32	SS81	Mancinella echinata	12S species 27	COI species 31
33	SS70	Mancinella siro	12S species 28	COI species 32
34	SS86	Menathais tuberosa	12S_species 29	COI_species 33
35 36	B302 SS75	Reishia bronni	125 spacias 30	COI_species 34
37 38	B293 SS34	Thais clavigera	125_species 50	COI_species 35
39	SS116	Vasum turbinellum	12S_species 31	COI_species 36
40	B363	Homoiodoris japonica	12S species 32	COI species 37

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

Table 3.	(Continu	ed).
----------	----------	------

#	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 emydia	12S species 35	COI species 40
44	B357	Lepeta kuragiensis	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	Batelleida anoshanina lamu	12S species 42	COL amonios 47
51	SS51	Patenoiaa saccharina ianx	12S species 43	COI_species 47
52	SS2	Cellana grata	128 species 11	COL spagios 48
53	SS3	Cellana toreuma	125_species 44	COI_species 48
54	SS209	Cellana grata		COL species 40
55	SS4	Cellana toreuma	12S_species 45	COI_species 49
56	SS18	Cellana nigrolineata		COI species 50
57	SS174	Cellana radiata	12S species 46	COI species 51
58	SS84	Cellana testudinaria	12S species 47	COI species 52
59	B247	Scutellastra flexuosa	12S_species 48	COI_species 53
60	B257	Siphonaria sirius	12S species 49	COI species 54
61	B359	Siphonaria sp.	12S species 50	COI species 55
62	SS131	Tectus pyramis	12S species 51	COI species 56
63	B268	Chlorostoma argyrostoma lischkei	128 spacias 52	
64	SS72	Chlorostoma turbinatum	125_species 52	COI_species 57
65	B285	Chlorostoma lischkei	12S species 53	
66	SS73	Tegula pfeifferi pfeifferi	128 spacias 54	COI species 58
67	SS69	Omphalius rusticus	125_species 54	COI species 59
68	SS67	Omphalius nigerrimus	12S species 55	COI species 60
69	B254	Cantharidus japonicus	128 spacias 56	COI species 61
70	B345	Cantharidus jessoensis	125_species 50	COI species 62
71	S391	Nipponacmea schrenkii	12S species 57	
72	B332	Lunella corgensis	12S species 58	COI_species 63
73	SS54	Lunciu corcensis	12S species 59	
74	SS101	Lunella coronata	12S_species 60	COI_species 64
75	B316	Turbo stenogyrus	12S species 61	COI species 65
		69	61	65

References

- W. Appeltan, 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)
- J.C. Orr, V.J. Fabry, O. Aumont, L. Bopp, S.C. Doney, R.A. Feely, et al., Nature 437, 681–686 (2005)
- 4. R. Rodolfo-Metalpa, F. Houlbaèque, É. Tambutté, 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)
- 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)
- 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)
- 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)
- 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)
- D.H.E. Setiamarga, R. Shiba, Y. Kamito, M. Yamamoto, N.N.B. Razali, M. Arai, et al, Zoosymposia 15, 129-140 (2019)
- 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)
- 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)
- F. De Mattia, R. Gentili, I. Bruni, A. Galimberti, S. Sgorbati, M. Casiraghi, et al, Bot. J. Linn. Soc. 169, 3 (2012)