Gene 662 (2018) 88–96

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

# Comparative mitogenomic analysis reveals cryptic species in *Reticunassa festiva* (Neogastropoda: Nassariidae)

Yi Yang[a], Qi Li[a,b,*], Lingfeng Kong[a], Hong Yu[a]

[a] Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 266003, China
[b] Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, China

ARTICLE INFO

*Keywords:*
Mitochondrial genomes
Phylogeny
Genetic diversity
Nassariidae
*Reticunassa*

ABSTRACT

The mitochondrial genome provides sets of genome-level characteristics, which can be useful in resolving ancient relationships. The complete mitochondrial genomes of five *Reticunassa* and one *Nassarius* were sequenced, including *Reticunassa hiradoensis*, *R. fratercula*, *R. festiva* complex A, B and C and *Nassarius variciferus*, and compared with other two Nassariidae species (*Tritia reticulatus* and *T. obsolete*). Eight genomes show the same gene order as in other caenogastropods, containing 13 protein-coding genes, 2 ribosome RNA genes and 22 transfer RNA genes, except *R. fratercula*, which lacks the *trnF*. Three-nucleotide deletion in *nad6* genes, truncated major non-coding regions are detected in all *Reticunassa* mitogenomes. The reconstructed phylogenetic trees show (*Reticunassa* + (*Tritia* + *Nassarius*)) at genus level and support the relationship between *R. festiva*-C and the clade formed by *R. hiradoensis* + *R. fratercula* and *R. festiva*-A + *R. festiva*-B within *Reticunassa*. The specific mitogenomic characters and phylogenetic analyses support the validity of genus *Reticunassa*. Pairwise genetic distances of protein-coding genes between *R. festiva* complex A, B and C range from 0.063 to 0.109, which are close to that of *R. hiradoensis* and *R. fratercula* (0.065), and higher than those of several closely related sister species of mollusks. These variation in mitochondrial genomes of *Reticunassa* indicate that *R. festiva*-A, *R. festiva*-B and *R. festiva*-C should be considered as separate species.

## 1. Introduction

The gastropod family Nassariidae, which distributes worldwide from intertidal zone to a depth of about 1000 m, belongs to the order Neogastropoda. Within Nassariidae, three subfamilies (Dorsaninae, Cylleninae and Nassariinae) including over 400 extant species have been recognised (Cernohorsky, 1984; Galindo et al., 2016). The current classification of Nassariidae, which is largely established by Cossmann (1901) and modified slightly by Cernohorsky (1984), is mainly based on morphological characters, including characters of shell, operculum, radula, metapodial tentacles and egg-capsule. However, the morphology-based taxonomy appears quite problematic at the genus or subgenus level. Of all the extant nassariid species, most are attributed to genus *Nassarius* (Nassariinae), which acts as a taxonomic waste basket, and only a little more than 60 species are classified in genera other than *Nassarius* (Galindo et al., 2016). In addition, the applications of subgenera within *Nassarius* are inconsistent and dramatically subjective (Cernohorsky, 1984; Haasl, 2000; Li et al., 2010; Pu et al., 2017).

*Reticunassa* Iredale, 1936 is the genus name used for the *pauper*-complex. In the taxonomic system of Nassariidae revised by Cernohorsky (1984), 16 taxa with similar morphological characters are treated as synonyms of *Nassarius* (*Hinia*) *pauperus* and their small morphological differences are considered as intraspecific variation. Kool and Dekker (2006, 2007) described ten species included in the *Nassarius pauperus* complex with several informative morphological characters, such as the intercordal axial sculpture and intercordal spiral sculpture. Molecular phylogenetic analyses based on partial gene fragements reveal that species attributed to the *pauper*-complex grouped together in a lineage separate from the genus *Nassarius* and therefore, *Reticunassa* is elevated to full genus status to include the species formally classified to genus *Nassarius* to form the *Nassarius pauperus* complex (Galindo et al., 2016; Galindo et al., 2017). However,

*Abbreviations:* atp6 and atp8, ATPase subunit 6 and 8 genes; *cob*, cytochrome *b* gene; *cox*1–3, cytochrome coxidase subunits I-III genes; NCR, non-coding region; nad1–6 and nad4L, NADH dehydrogenase subunits 1–6 and 4 L genes; ML, maximum likelihood; BI, Bayesian inference; rRNA, ribosomal RNA; *rrnL* and *rrnS*, large and small subunits of ribosomal RNA genes; tRNA, transfer RNA; PCG, protein coding gene; mt genome, mitochondrial genome; Ka and Ks, non-synonymous and synonymous; POR, potential origin of replication; Ala-A, alanine; Arg-R, arginine; Asn-N, asparagine; Aps-D, aspartic acid; Cys-C, cysteine; Gln-Q, glutamine; Glu-E, glutamic acid; Gly-G, glycine; His-H, histidine; Ile-I, isoleucine; Leu-L, leucine; Lys-K, lysine; Met-M, methionine; Phe-F, phenylalanine; Pro-P, proline; Ser-S, serine; Thr-T, threonine; Trp-W, tryptophan; Tyr-Y, tyrosine; Val-V, valine

* Corresponding author at: Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao 266003, China.
*E-mail address:* qili66@ouc.edu.cn (Q. Li).

https://doi.org/10.1016/j.gene.2018.04.001
Received 3 January 2018; Received in revised form 24 March 2018; Accepted 3 April 2018
Available online 05 April 2018
0378-1119/ 

the validity of *Reticunassa* still needs verified with a longer sequences dataset due to the disadvantages of short gene fragements. On the other hand, the diversity within *Reticunassa* is underestimated. *Reticunassa festiva* is a common species widely distributed throughout the Indo-pacific region (Cernohorsky, 1972). The diversity of *R. festiva* has been discussed before (Pu et al., 2017; Zou et al., 2012). Three *R. festiva* clades (Clade A, B and C) with significant genetic divergence among each other are detected along the coast of the China seas in our previous study (data unpublished), revealing the cryptic diversity of *R. festiva*. However, whether these cryptic lineages are distinct species have not been determined since all these studies are based on short fragments, including the partial cytochrome oxidase subunit I (*cox1*) sequences and the *rrnL* sequences. And therefore, supplemental analyses of informative markers, such as mt genomes, will be required to solve this question.

In the present study, six complete mt genomes of Nassariidae species, including *R. festiva* complex (Clade A, Clade B and Clade C), *R. hiradoensis*, *R. fratercula*, together with *Nassarius variciferus*, a representative species of the genus *Nassarius*, were sequenced. Comparative analysis of the mitochondrial genomes were conducted with the inclusion of two available mt genomes *Tritia reticulatus* (Nassariidae) (EU827201; Cunha et al., 2009) and *T. obsolete* (DQ238598; Simison et al., 2006). The results will provide effective information to understand the specific characteristics of *Reticunassa* mt genomes. Moreover, the genetic distance analysis and phylogenetic analysis inferred in this study will help us understand the evolutionary relationship within Nassariidae and determine the taxonomic status of *R. festiva* complex A, B and C.

## 2. Material and methods

### 2.1. Sample collection and DNA extraction

Our previous study have shown that up to three highly divergent mitochondrial lineages exist in sympatry along the coast of China seas in *R. festiva* populations (data unpublished). Three individuals representing the three lineages were chosen and named as *R. festiva* Clade A, collected in Qingdao, Shandong Province (36°19′30″N; 120°42′4″E), *R. festiva* Clade B, collected in Ningbo, Zhejiang Province (30°6′59″N; 121°55′34″E), and *R. festiva* Clade C, collected in Qisha, Guangxi Province (21°33′42″N; 108° 28′49″E). *R. hiradoensis* was collected in Rongcheng, Shandong Province (37°21′29″N; 122°35′57″E), and *R. fratercula* was collected in Qingdao, Shandong Province (36°19′30″N; 120°42′4″E). For *N. variciferus*, specimen was collected in Dongying, Shandong Province (37°53′27″N; 119°6′47″E).

Specimens collected were stored in 95% ethanol. Genomic DNA was extracted from small pieces of foot tissue by the CTAB method as modified by Winnepenninckx et al. (1993) and visualized on 1.0% agarose gel.

### 2.2. PCR amplification and sequencing

Short fragements of *cox1* and *rrnL* were amplified by polymerase chain reaction (PCR) with universal primer pairs (Folmer et al., 1994) and specific primers designed by Zou et al. (2012), respectively. According to the sequences obtained above, two pairs of long PCR primers were firstly designed for *N. variciferus* mt genome (Table S1). The long primers (Table S1) of other five mt genomes were then designed based on the that of *N. variciferus*, as well as the available nassariid mt genome sequences of *T. reticulatus* and *T. obsolete*. PCR amplifications were conducted in a total volume of 25 μL, containing 19.25 μL 18.7 μL sterile distilled $H_2O$, 0.5 μL template DNA (approximately 100 ng), 2.5 μL 10 × LA-buffer ($Mg^{2+}$ plus), 0.5 μL dNTPs (10 mM), 1 μL of each primers (10 μM), 0.25 μL (1 U) LA-*Taq* DNA polymerase. Long-PCR profile included an initial denaturation step of 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 56–60 °C for 30 s, 68 °C for 1 min / Kb, and a final extension step at 72 °C for 10 min. The PCR products were confirmed by 1.5% agarose gel electrophoresis and stained with ethidium bromide, purified with EZ Spin Column PCR Product Purification Kit (Sangon). Purified products were sequenced using an ABI 3730 automatic sequencer (Applied Biosystems) at LiuHe HuaDa Biotechnology Company (Beijing, China) based on a primer-walking strategy.

### 2.3. Genome assembly, gene annotation and sequence analysis

The sequences were assembled using SeqMan (www.DNASTAR.com), and further adjusted manually in a few cases. Protein coding genes (PCGs) were determined by ORF Finder (http://www.ncbi.nlm.nih.gov/orffinder) using the invertebrate mitochondrial code. The accurate boundaries of PCGs and rRNA genes were identified by comparing with those of *T. reticulatus* and *T. obsolete*. Transfer RNA (tRNA) genes were identified using tRNA scan-SE1.21 (Lowe and Eddy, 1997) and ARWEN (Laslett and Canbäck, 2008). Codon usage of 13 PCGs and the A + T content values and nucleotide frequencies of mt genomes were estimated by MEGA 5 (Tamura et al., 2011). In addition, the base skew values for a given strand was calculated as: AT skew = (A − T) / (A + T), GC skew = (G − C) / (G + C) (Perna and Kocher, 1995). Pairwise genetic distances (p-distance) for PCGs were implemented in MEGA 5. The ratio of nonsynonymous and synonymous substitution rates (Ka/Ks) was estimated with Ka_Ks calculator 2.0 (Wang et al., 2010) in all 13 protein genes of eight nassariid species and graph plotting was conducted using GraphPad Prism 6 (GraphPad Software, La Jolla, California USA).

The gene map of nassariid mitogenome was generated with the program CGView (Grant and Stothard, 2008). The graphs of the alignments of *nad6* genes, *trnF/trnF*-like sequences, and major non-coding regions were drawn using Microsoft Visio 2010.

### 2.4. Phylogenetic analysis

A total of 11 taxa were selected for phylogenetic analysis (Table 1), including complete mt sequences of eight nassariid species. *Buccinum pemphigus* (KT962044), *Volutharpa perryi* (KT382829) and *Neptunea arthritica* (KU246047) from family Buccinidae (Neogastropoda: Buccinoidea) were used as outgroups.

The 13 PCGs were aligned by codons with Clustal W (Higgins et al., 1994) in MEGA 5 respectively, and further verified manually. GTR + I + G was selected as the best-fit nucleotide substitution model for each gene by jModelTest (Posada, 2008) based on the Akaike Information Criterion (AIC). The 3rd codons of all PCGs were discarded since a high saturation were detected on this position using Xia's test implemented in DAMBE5 (Xia, 2013). Phylogenetic analyses were conducted using maximum likelihood (ML, Felsenstein, 1981) and Bayesian inference (BI, Huelsenbeck and Ronquist, 2001). ML analyses were carried out using software RAxML v. 8.2.1 (Stamatakis, 2006) with the rapid hill-climbing algorithm and 10,000 bootstrap

**Table 1**
List of the 11 species analysed in this paper with their GenBank accession numbers.

| Species | Family | Accession no. |
|---|---|---|
| *Buccinum pemphigus* | Buccinidae | KT962044 |
| *Volutharpa perryi* | Buccinidae | KT382829 |
| *Neptunea arthritica* | Buccinidae | KU246047 |
| *Tritia obsoleta* | Nassariidae | DQ238598 |
| *Tritia reticulatus* | Nassariidae | EU827201 |
| *Nassarius variciferus* | Nassariidae | KM603509 |
| *Reticunassa hiradoensis* | Nassariidae | MG744569 |
| *R. fratercula* | Nassariidae | KT826695 |
| *R. festiva*-A | Nassariidae | KT735055 |
| *R. festiva*-B | Nassariidae | MF148855 |
| *R. festiva*-C | Nassariidae | MG744570 |

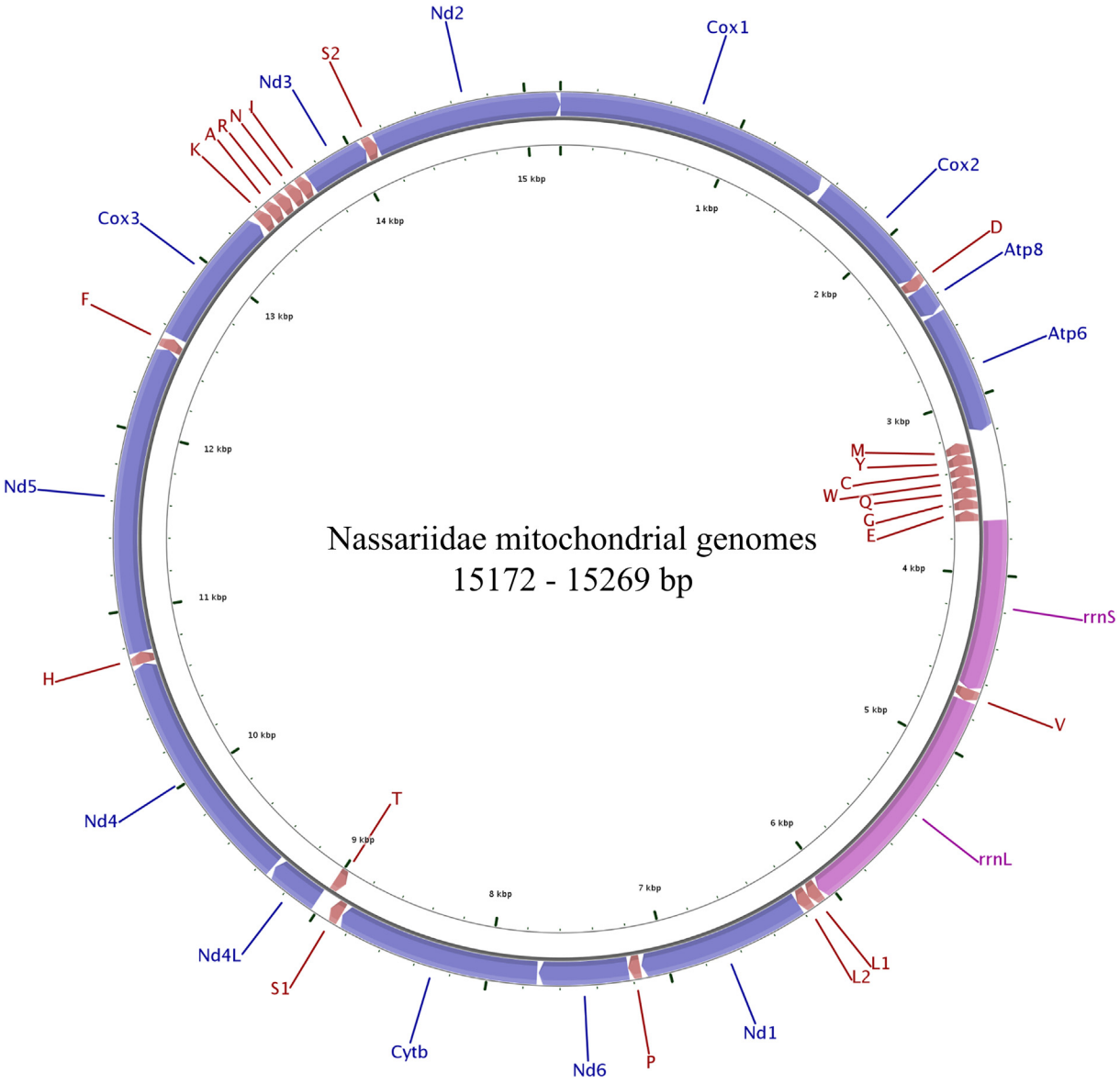


**Fig. 1.** Gene map of the mt genomes of six nassariids.

pseudoreplicates (BP). BI analyses were performed with MrBayes v. 3.1.2 (Ronquist et al., 2012), running four simultaneous Monte Carlo Markov chains (MCMC) for 10,000,000 generations, sampling every 1000 generations, and discarding the first 25% generations as burn-in. Parameter convergence was achieved within ten million generations and the standard deviation of split frequencies was less than 0.01. All parameters were checked with Tracer v. 1.6 (Drummond and Rambaut, 2007), and the ESS was more than 200. The resulting phylogenetic trees were visualized in FigTree v1.4.2.

## 3. Results and discussion

### 3.1. Genome structure, organization and composition

The six mt genomes are completely sequenced and submitted to Genbank and accession numbers are shown in Table 1. The sequences are found to be from 15,174 bp to 15,269 bp in size, possessing the typical Neogastropoda mt gene content of 13 PCGs, 22 tRNA genes, 2 rRNA genes as well as a potential origin of replication (Cunha et al., 2009), with the exception of *R. fratercula*, in which a *trnF* was missing (Table S2). All of the 13 PCGs, 14 tRNAs (*trnD*, *trnV*, *trnL1*, *trnL2*, *trnP*, *trnS1*, *trnS2*, *trnH*, *trnF*, *trnK*, *trnA*, *trnR*, *trnN* and *trnI*) and 2 rRNAs (*rrnL* and *rrnS*) are encoded in the major strand, while the other 8 tRNAs are encoded in the minor strand (Fig. 1).

The nucleotide compositions and AT contents of complete mt genomes of six nassariids have a similar pattern (Table 2). The AT contents of six mt genomes range from 69% to 71%, similar to other molluscan taxa reported (Arquez et al., 2014; Grande et al., 2008; Sun et al., 2015; Xu et al., 2012). The AT skew and GC skew values were calculated with the addition of *T. reticulatus* and *T. obsolete* mt genomes. The nucleotide compositions are slightly skewed away from C in favor of G (the GC-skews are from 0.032 to 0.067) and from A in favor of T (the AT-skews are from −0.130 to −0.114) (Table 3). The negative AT skew and positive GC skew are common in other Gastropoda species, e.g., Neritimorpha (Arquez et al., 2014), Nudibranch (Sevigny et al., 2015) and Neogastropoda (Cunha et al., 2009).

### 3.2. Protein-coding genes

Negative AT skew and positive GC skew are also detected in the 13 PCGs (Table 3). The AT skews show little difference within the eight nassariids, while larger differences are found in the GC skews,

**Table 2**
List of total size, AT content, AT- and GC-skew, for mitochondrial genes of *Reticunassa hiradoensis* (*Rhi*), *R. fratercula* (*Rfr*), R. festiva-A (*Rfe*A), R. festiva-B (*Rfe*B), R. festiva-C (*Rfe*C), *Nassarius variciferus* (Nva), *Tritia reticulatus* (*Tre*) and *T. obsolets* (*Tob*) with lengths of genes and potential origin of replication (POR), and initiation/termination codon of protein-coding genes.

| | *Rhi* | *Rfr* | *Rfe*A | *Rfe*B | *Rfe*C | Nva | *Tre* | *Tob* |
|---|---|---|---|---|---|---|---|---|
| Total size | 15,194 | 15,174 | 15,195 | 15,194 | 15,172 | 15,269 | 15,271 | 15,263 |
| %A + T | 0.7 | 0.71 | 0.69 | 0.69 | 0.69 | 0.7 | 0.67 | 0.69 |
| *rrnS* | 958 | 964 | 956 | 956 | 955 | 958 | 960 | 886 |
| *rrnL* | 1346 | 1346 | 1345 | 1342 | 1346 | 1351 | 1353 | 1492 |
| POR | 14 | 15 | 17 | 15 | 15 | 56 | 57 | 55 |
| *atp6* | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAG) | 696(ATG/TAA) |
| *atp8* | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) | 159(ATG/TAA) |
| *cob* | 1140(ATG/TAA) | 1140(ATG/TAA) | 1140(ATG/TAA) | 1140(ATG/TAA) | 1140(ATG/TAA) | 1140(ATG/TAG) | 1140(ATG/TAA) | 1140(ATG/TAA) |
| *cox1* | 1536(ATG/TAA) | 1536(ATG/TAA) | 1536(ATG/TAA) | 1536(ATG/TAA) | 1536(ATG/TAA) | 1536(ATG/TAG) | 1536(ATG/TAA) | 1536(ATG/TAA) |
| *cox2* | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) | 687(ATG/TAA) |
| *cox3* | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) | 780(ATG/TAA) |
| *nad1* | 942(ATG/TAG) | 942(ATG/TAA) | 942(ATG/TAA) | 942(ATG/TAA) | 942(ATG/TAA) | 942(ATG/TAA) | 942(ATG/TAA) | 942(ATG/TAA) |
| *nad2* | 1055(ATG/TA) | 1055(ATG/TA) | 1055(ATG/TA) | 1055(ATG/TA) | 1055(ATG/TA) | 1055(ATG/TA) | 1056(ATG/TAA) | 1054(ATG/T) |
| *nad3* | 354(ATG/TAA) | 354(ATG/TAA) | 354(ATG/TAA) | 354(ATG/TAG) | 354(ATG/TAA) | 354(ATG/TAA) | 354(ATG/TAA) | 354(ATG/TAA) |
| *nad4* | 1536(TTA/TAA) | 1536(TTA/TAA) | 1536(TTA/TAA) | 1536(TTG/TAA) | 1536(TTA/TAA) | 1536(ATA/TAG) | 1536(ATA/TAG) | 1536(ATA/TAG) |
| *nad4l* | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) | 297(ATG/TAG) |
| *nad5* | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) | 1722(ATG/TAG) |
| *nad6* | 495(ATG/TAA) | 495(ATG/TAA) | 495(ATG/TAG) | 495(ATG/TAG) | 495(ATG/TAA) | 498(ATG/TAG) | 498(ATG/TAA) | 498(ATG/TAA) |

**Table 3**
Skew values for mt genomes, PCGs and rRNAs of eight nassariids.

| Species | A + T skew (mt genome) | A + T skew (PCGs) | A + T skew (rRNAs) | G + C skew (mt genome) | G + C skew (PCGs) | G + C skew (rRNAs) |
|---|---|---|---|---|---|---|
| *R. hiradoensis* | −0.114 | −0.180 | 0.043 | 0.067 | 0.019 | 0.156 |
| *R. fratercula* | −0.127 | −0.181 | 0.053 | 0.034 | 0.017 | 0.145 |
| *R. festiva − A* | −0.130 | −0.176 | 0.052 | 0.032 | 0.007 | 0.167 |
| *R. festiva − B* | −0.130 | −0.181 | 0.051 | 0.032 | 0.028 | 0.172 |
| *R. festiva − C* | −0.130 | −0.171 | 0.053 | 0.032 | 0.016 | 0.158 |
| *N. variciferus* | −0.114 | −0.186 | 0.062 | 0.067 | 0.044 | 0.150 |
| *T. reticulatus* | −0.116 | −0.173 | 0.058 | 0.039 | 0.017 | 0.145 |
| *T. obsolete* | −0.127 | −0.185 | 0.058 | 0.042 | 0.019 | 0.146 |

especially those between *N. variciferus* and the remaining species. The initiation and termination codons of the 13 PCGs encoded by the eight Nassariidae genomes are shown in Table 2. All genes of the analysed mt genomes begin with the standard ATG codon except *nad4*, which used other start codons (ATA in *N. variciferus*, *T. reticulatus* and *T. obsolets*; TTG in *R. festiva*-B; TTA in *R. hiradoensis*, *R. fratercula*, *R. festiva*-A and *R. festiva*-C). TTG has been reported as a start codon in quite a few mollusk mt genomes (Cunha et al., 2009; Grande et al., 2008; Uribe et al., 2016; Yuan et al., 2012), but the usage of TTA initiation codon has only been reported in the *cob* gene of oyster *Crassostrea virginica* (Wu et al., 2010). All 13 PCGs of eight mt genomes end in full termination codon (TAG, $N = 33$; TAA,$N = 64$), except for *nad2* genes of *R. hiradoensis*, *R. fratercula*, *R. festiva-A*, *R. festiva-B*, *R. festiva-C*, and *N. variciferus*, ending with the incomplete stop codon TA and *nad2* gene of *T. obsolets*, ending with incomplete stop codon T. The presence of incomplete stop codons is very common in invertebrate mt genes (Xin et al., 2017) and the incomplete TA or T codons may be modified to the TAA termini via posttranscriptional polyadenylation (Ojala et al., 1981).

Identical initiation codons are shared by all the PCGs except *nad4* gene of the eight Nassariidae species, but different termination codons are found in eight of the PCGs between any pair of the eight species. On the other hand, five PCGs with different termination codons are discovered within genus *Reticunassa* (Table 2).

When the *nad6* genes of the eight nassariids are aligned, a deletion of three continuous nucleotides are discovered in all the five *Reticunassa* mt genomes (Fig. 2). A translated amino acid alignment confirms the locations. As a result, the three nucleotides deletion only lead to the delection of one amino acid and does not bring other changes in the deduced amino acid sequences. Mitochondrial gene deletion among related species has been found in *cox1* gene of *Melibe* (Gastropoda) taxa, and the deletions are explained as a derived trait, which may reflect unusual constraints on the protein in those taxa (Sevigny et al., 2015). The *nad6* deletion may be a conservative trait of *Reticunassa* mt genomes compared with those of other two genus *Nassarius* and *Tritia* from the same family.

### 3.3. Nonsynonymous and synonymous substitutions

The estimation of non-synonymous (Ka) and synonymous (Ks) substitution rates is important to understand the evolutionary dynamics of PCGs across closely related species (Fay and Wu, 2003; Ohta, 1995). Pairwise values of Ka/Ks of the 13 PCGs were calculated among the eight nassariids. The ratio varied from 0 (*R. festiva-A - R. festiva-B*, *N. variciferus - T. obsoleta*, *atp8*; *R. festiva-A - R. festiva-C*, *cox1*; *R. hiradoensis - R. festiva-B*, *R. hiradoensis - R.festiva-C*, *cox2*; *R. festiva-A - R. festiva-B*, *nad1*; *R. festiva-A - R. festiva-B*, *T. reticulatus - T. obsoleta nad4l*;) to 0.153 (*R. fratercula - N. variciferus*, *nad6*), indicating the existence of different selection pressures among genes (Fig. 3). In the present study, most higher Ka/Ks values are found in *nad6*, *nad2* and *nad3* genes, suggesting that these NADH complex genes are suffering a more relaxed purifying selection compared with conservative genes like *cox1*, *cox2* and *cox3* (Ren et al., 2010). Within Nassariidae, *cox1* gene has been widely used in phylogenetic analyses (Galindo et al., 2016; Pu et al., 2017; Zou et al., 2012). In addition to *cox1*, *cox2* could also be informative as a phylogenetic marker within Nassariidae due to its low Ka/Ks ratio (Arquez et al., 2014).

### 3.4. Transfer and ribosomal RNA genes

Most metazoan mt genomes contain a set of 22 tRNA genes in total, including two copies of *trnL* and two of *trnS*, and other 18 tRNA genes with only one copy (Podsiadlowski et al., 2008). In the present study, all the nassariid mt genomes follow the rule, except for that of *R. fratercula*, which lacks the *trnF*. By comparison with *trnF* gene sequences of nassariids, a *trnF*-like sequence is also found in the *R. fratercula* mt genome in the same locus (Fig. 4A). In arthropod mt genomes, truncated and lost tRNAs has been described, and compensatory mechanisms to a functional tRNA during the translation and the interaction with the ribosome seem to be an acceptable hypothesis (Amaral et al., 2016; Domes et al., 2008; Masta and Boore, 2008).

Within the eight nassariids mt genomes, the size of *rrnS*, flanked by

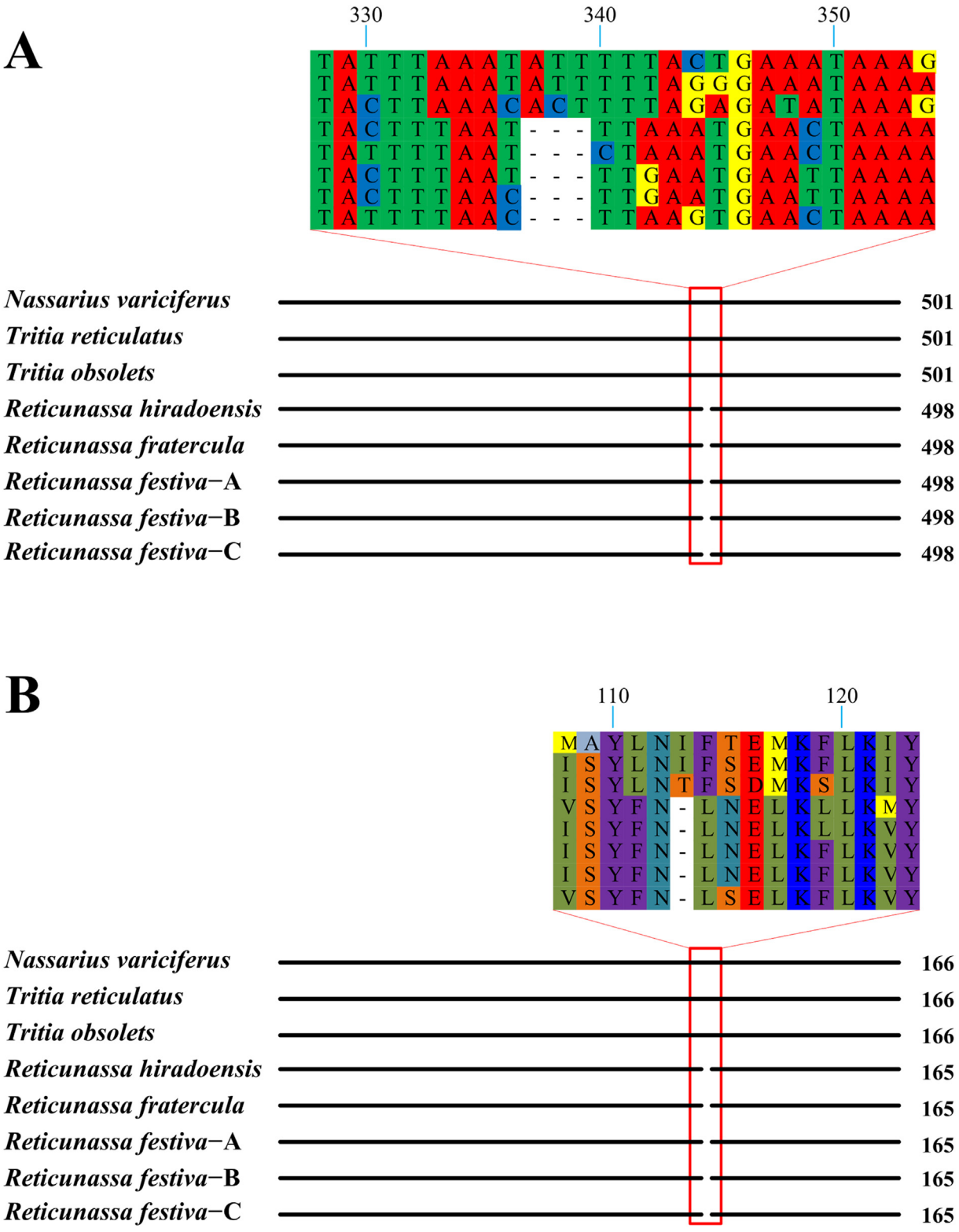


**Fig. 2.** *Nad6* sequence differences in eight nassariidas. Nucleotide (A) and amino acid (B) alignments of a portion of nad6 gene for *Reticunassa* and other members of Nassariidae indicate that *Reticunassa* lack the three nucleotide deletion present in *Nassarius* and *Tritia*.

*trnE* and *trnV*, ranges from 886 bp (*T. obsoleta*) to 964 bp (*R. fratercula*). The length of *rrnL* gene, located between *trnV* and *trnL1* (CUA), varies from 1342 bp (*R. festiva*-B) to 1492 bp (*T. obsoleta*) (Table 2). The A + T content of the two rRNA genes range from 70.3% (*T. reticulatus*) to 72.2% (*R.hiradoensis*). The AT and GC skews of two rRNA genes are all positive (Table 3).

### 3.5. Non-coding region

The nassariid mt genomes also feature several non-coding regions, among which the intergenic region between *trnF* and *cox3*, has been assumed as candidate to contain the control region in other gastropod mt genomes (Cunha et al., 2009; Kurabayashi and Ueshima, 2000; Osca

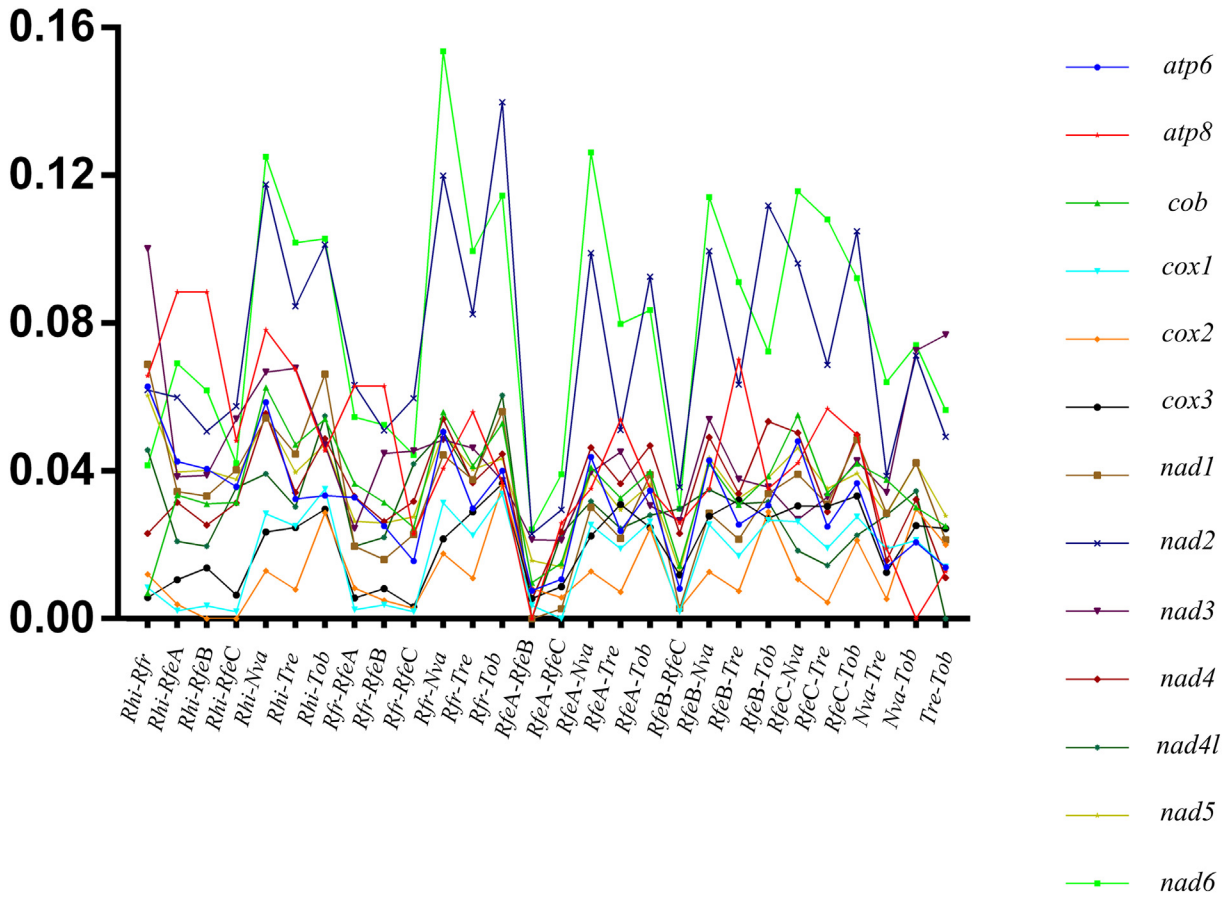


**Fig. 3.** The ratio of nonsynonymous and synonymous substitutions (Ka/Ks) estimated in all 13 protein genes of eight nassariids. *Nva*, *Nassarius variciferus*; *Tre*, *Tritia reticulatus*; *Tob*, *T. obsolets*; *Rhi*, *Reticunassa hiradoensis*; *Rfr*, *R. fratercula*; *Rfe*A, *R. festiva*-A; *Rfe*B, *R. festiva*-B; *Rfe*C, *R. festiva*-C.

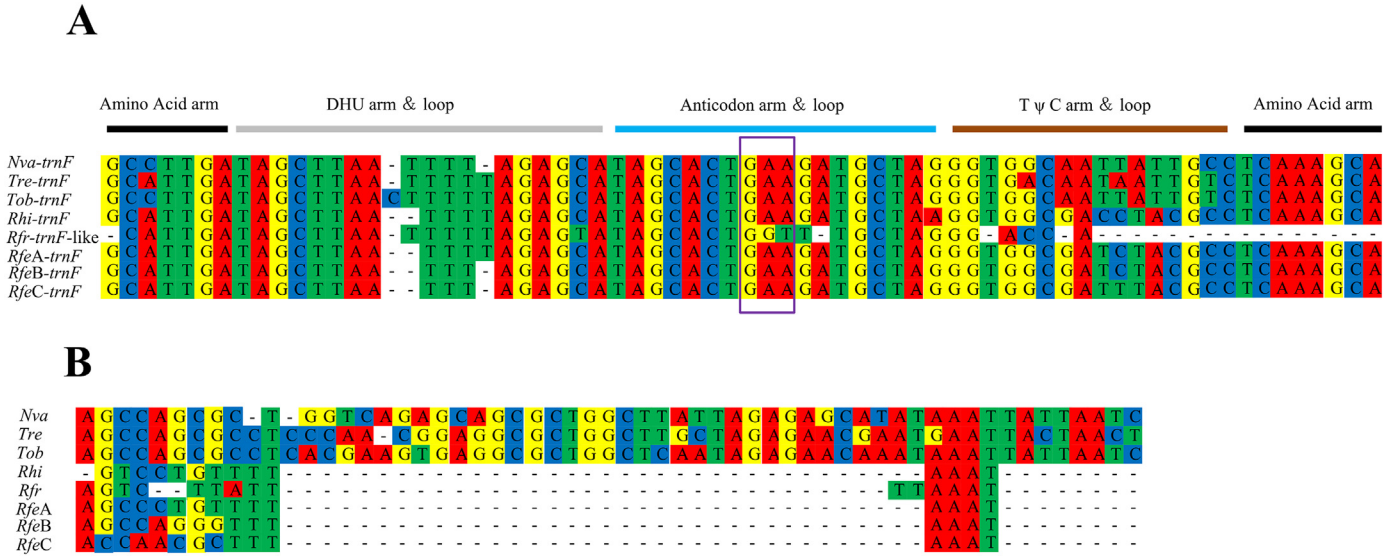


**Fig. 4.** Alignment of *trnF* and *trnF*-like sequences (A), and that of the major non-coding region (B) in mt genomes of eight nassariids. TRNA secondary structure is displayed above the alignment, and the position of the anticodon is highlighted with rectangular frame. *Nva*, *Nassarius variciferus*; *Tre*, *Tritia reticulatus*; *Tob*, *T. obsolets*; *Rhi*, *Reticunassa hiradoensis*; *Rfr*, *R. fratercula*; *Rfe*A, *R. festiva*-A; *Rfe*B, *R. festiva*-B; *Rfe*C, *R. festiva*-C.

et al., 2015). Secondary structures have been detected in this potential origin of replication in some gastropod mt genomes (Zhao et al., 2010). In nassariid mt genomes, this putative control region of *Reticunassa* is significantly truncated compared with that of *N. variciferus*, *T. reticulatus* and *T. obsoleta* (Fig. 4B). Such non-coding region deletions may result from the slippage events, which occur in region with high contents of A, T or AT/TA repetition (Amaral et al., 2016). This finding is another important trait of the *Reticunassa* mt genomes. In addition to the noncoding region upstream to the 5′ end of the *cox3* gene, another major intergenic region, which is the longest non-coding region in

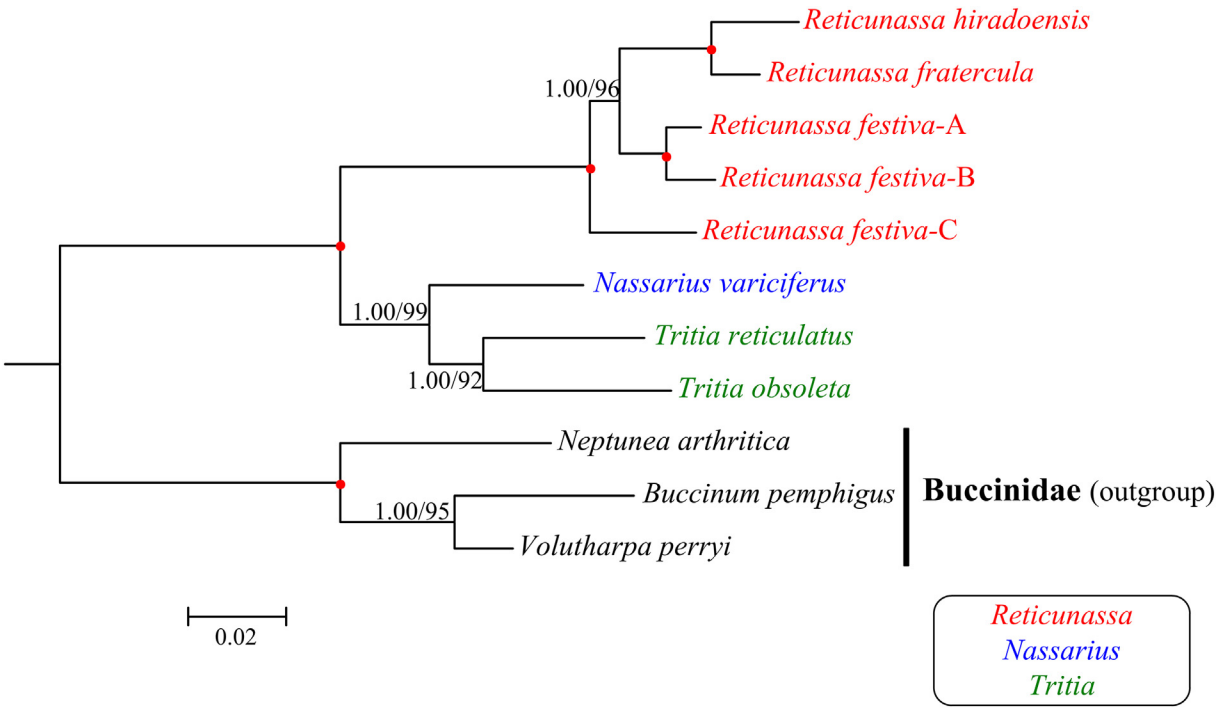


**Fig. 5.** Phylogenetic trees derived from ML and BI analyses based on nucleotide sequences of 13 mitochondrial PCGs. The first number at each node is Bayesian posterior probability and the second number is bootstrap probability of ML analyses. Solid red circles represent nodes with posterior probabilities = 1.00 and bootstrap proportions = 100. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 4**
Pairwise genetic distance of 13 PCGs of *R. hiradoensis* (Rhi), *R. fratercula* (Rfr), *R. festiva*-A (RfeA), *R. festiva*-B (RfeB), *R. festiva*-C (RfeC), *Nassarius variciferus* (Nva), *Tritia reticulatus* (Tre) and *T. obsolets* (Tob).

| | Rhi | Rfr | RfeA | RfeB | RfeC | Nva | Tre | Tob |
|---|---|---|---|---|---|---|---|---|
| Rhi | – | | | | | | | |
| Rfr | 0.065 | | | | | | | |
| RfeA | 0.106 | 0.100 | | | | | | |
| RfeB | 0.109 | 0.103 | 0.063 | | | | | |
| RfeC | 0.113 | 0.108 | 0.109 | 0.109 | | | | |
| Nva | 0.171 | 0.169 | 0.173 | 0.171 | 0.167 | | | |
| Tre | 0.181 | 0.178 | 0.183 | 0.182 | 0.178 | 0.149 | | |
| Tob | 0.174 | 0.169 | 0.177 | 0.174 | 0.170 | 0.144 | 0.148 | – |

**Table 5**
Comparison of pairwise genetic distance of 13 PCGs of 12 taxa in Gastropoda.

| Genus | Species | Accession Number | Species | Accession Number | Genetic distance |
|---|---|---|---|---|---|
| *Conus* | *C. ventricosus* | KX263251 | *C. borgesi* | EU827198 | 0.092 |
| | *C. hybridus* | KX263252 | *C. ventricosus* | KX263251 | 0.100 |
| | *C. hybridus* | KX263252 | *C. borgesi* | EU827198 | 0.103 |
| | *C. gloriamaris* | KU996360 | *C. textile* | DQ862058 | 0.117 |
| | *C. striatus* | KX156937 | *C. consors* | KF887950 | 0.130 |
| | *C. venulatus* | KX263250 | *C. borgesi* | EU827198 | 0.133 |
| | *C. ventricosus* | KX263251 | *C. venulatus* | KX263250 | 0.136 |
| | *C. hybridus* | KX263252 | *C. venulatus* | KX263250 | 0.137 |
| *Babylonia* | *B. lutosa* | KF897830 | *B. areolata* | HQ416443 | 0.125 |
| *Nerita* | *N. tessellate* | KF728889 | *N. fulgurans* | KF728888 | 0.091 |

*Reticunassa* mt genomes, was detected downstream the 3′ end of the *atp6* gene. It is still unclear which of the non-coding regions may act as the control region.

### 3.6. Interspecific genetic distance and phylogenetic relationship

Phylogenetic trees based on 13 PCGs of the 11 buccinoid mitochondrial genomes were reconstructed (Fig. 5). The ML and BI analyses arrived similar tree topologies. Within Nassariidae, the five *Reticunassa* form a good tight group, with *Nassarius* + *Tritia* being the sister taxon. This result is consistent with previous founding revealed by short fragments (Galindo et al., 2016) and support the validity of genus *Reticunassa* at mitogenomic level. The phylogenetic trees show that the clade clustered by *R. festiva*-A + *R. festiva*-B and *R. fratercula* + *R. hiradoensis* groups with *R. festiva*-C and together form the genus *Reticunassa*. This evidence indicate that the four *Reticunassa* species except *R. festiva*-C, share a more colse relationship and *R. festiva*-C should be treated as a distinct species.

Pairwise genetic distance of 13 PCGs between species in genus *Reticunassa* (0.063–0.113) is lower than that in genus *Tritia* (0.148) (Table 4), which also indicate a closer relationship within members in *Reticunassa*. To have a better understanding of the genetic divergence of closely related taxa in gastropods, the genetic distance of 13 PCGs between several congeneric species are also conducted in this study (Table 5). The genetic distance between *R. festiva*-C and any other *Reticunassa* species is quite similar (0.108–0.113), and consists with that between congeneric gastropods which are at species level. This is another evidence that support *R. festiva*-C should be considered as a distinct species instead of a cryptic lineage of *R. festiva* complex. Geographically, *R. festiva*-C is restricted to the Beibu Gulf, indicating that it may be an endemic species in the South China Sea. Pairwise genetic distance value between *R. festiva*-A and *R. festiva*-B is 0.063 (Table 4), close to that of *R. fratercula* and *R. hiradoensis* (0.065), and it is greater than several values between congeneric species in mollusks, especially Bivalvia, such as those of *Mytilus* (*M. trossulus* and *M. galloprovincialis*, 0.043; *M. trossulus* and *M. edulis*, 0.060), *Meretrix* (*M. petechialis* and *M. lusoria*, 0.027; *M. meretrix* and *M. lusoria*, 0.026) (Shen et al., 2014). According to the description of Zou et al. (2012), a slightly ambiguous but significant morphological difference between *R. festiva*-A and *R. festiva*-B is that the verrucous protuberances on the shell of *R. festiva*-B is bigger than that of *R. festiva*-A. Genetic analyses, combined with the morphological traits, show that *R. festiva*-A and *R. festiva*-B can be considered as separate species. Although sympatric, the distribution of

*R. festiva*-A and *R. festiva*-B seems to follow certain laws. *R. festiva*-A is mainly distributed in the north of the Yangtze River, while *R. festiva*-B is in majority in the south of the Yangtze River. In the previous study, the different distribution patterns between *R. festiva*-A and *R. festiva*-B were explained as the preferences of different environmental factors (e.g. temperature, salinity) of their living conditions.

Genetic evidence reveals five lineages in genus *Reticunassa*, which suggests that the diversity of *Reticunassa* is still underestimated and more morphological characters should be found to define these closely related species.

## 4. Conclusion

This study presents mt genomes of five *Reticunassa* species and one *Nassarius* species. The mt genomes arrangements in different nassariids are similar, except for *R. fratercula*, which does not show the *trnF* gene. Compared with *Nassarius* and *Tritia*, three major characters of *Reticunassa* mt genomes are detected: a deletion of three nucleotides in *nad6* gene, a truncated major non-coding region and relatively lower genetic distances. Comparative mitogenomic analyses, combined with phylogenetic analyses, support the validity of genus *Reticunassa*. Within *Reticunassa*, considerable variations have been found between the genetic distances of five mt genomes. The genetic distances among different congeneric molluscan species and comparative analyses of the five *Reticunassa* mt genomes, together with phylogenetic analyses, show that the differences between *R. festiva*-A, *R. festiva*-B and *R. festiva*-C are similar to or even greater than several closely related species, indicating that they should be considered as three distinct species.

## Acknowledgments

This study was supported by research grants from Fundamental Research Funds for the Central Universities (201762014), Shandong Province (2016ZDJS06A06), and major project for Tianjin Seed Technology (15ZXZYNC00050).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2018.04.001.

## References

Amaral, D.T., Mitani, Y., Ohmiya, Y., Viviani, V.R., 2016. Organization and comparative analysis of the mitochondrial genomes of bioluminescent Elateroidea (Coleoptera: Polyphaga). Gene 586, 254–262.
Arquez, M., Colgan, D., Castro, L.R., 2014. Sequence and comparison of mitochondrial genomes in the genus *Nerita* (Gastropoda: Neritimorpha: Neritidae) and phylogenetic considerations among gastropods. Mar. Genomics 15, 45–54.
Cernohorsky, W.O., 1972. Indo-Pacific Nassariidae (Mollusca: Gastropoda). Rec. Auckl. Inst. Mus. 9, 125–194.
Cernohorsky, W.O., 1984. Systematics of the family Nassariidae (Mollusca:Gastropoda). Bull. Auckl. Inst. Mus. 14, 1–356.
Cossmann, M., 1901. Essais de Paléoconchologie Comparée, Volume 4. Self-published, Paris. 293.
Cunha, R.L., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evol. Biol. 9, 210.
Domes, K., Maraun, M., Scheu, S., Cameron, S.L., 2008. The complete mitochondrial genome of the sexual oribatid mite *Steganacarus magnus*: genome rearrangements and loss of tRNAs. BMC Genomics 9, 13.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Fay, J.C., Wu, C.I., 2003. Sequence divergence, functional constraint, and selection in protein evolution. Annu. Rev. Genomics Hum. Genet. 4, 213–235.
Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplication of mitochondrial cytpchrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.
Galindo, L.A., Puillandre, N., Utge, J., Lozouet, P., Bouchet, P., 2016. The phylogeny and systematics of the Nassariidae revisited (Gastropoda, Buccinoidea). Mol. Phylogenet. Evol. 99, 337–353.
Galindo, L.A., Kool, H.H., Dekker, H., 2017. Review of the *Nassarius pauperus* (Gould, 1850) complex (Nassariidae): part 3, reinstatement of the genus *Reticunassa*, with the description of six new species. Eur. J. Taxon. (275), 1–43.
Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evol. Biol. 8, 61.
Grant, J.R., Stothard, P., 2008. The CGView server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 36, 181–184.
Haasl, D.M., 2000. Phylogenetic relationships among nassariid gastropods. J. Paleontol. 74, 839–852.
Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Huelsenbeck, J., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Kool, H.H., Dekker, H., 2006. Review of the *Nassarius pauper* (Gould, 1850) complex (Gastropoda, Nassariidae). Part 1, with the description of four new species from the Indo-West Pacific. Visaya 1, 54–75.
Kool, H.H., Dekker, H., 2007. Review of the *Nassarius pauper* (Gould, 1850) complex (Gastropoda, Nassariidae). Part 2, the western Indian Ocean species, with the description of two new species and introducing a nomen novum. Visaya 2, 62–77.
Kurabayashi, A., Ueshima, R., 2000. Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: systematic implication of the genome organization. Mol. Biol. Evol. 17, 266–277.
Laslett, D., Canbäck, B., 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24, 172–175.
Li, H., Lin, D., Fang, H., Zhu, A., Gao, Y., 2010. Species identification and phylogenetic analysis of genus *Nassarius* (Nassariidae) based on mitochondrial genes. Chin. J. Oceanol. Limnol. 28, 565–572.
Lowe, T.M., Eddy, S.R., 1997. TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Masta, S.E., Boore, J.L., 2008. Parallel evolution of truncated transfer RNA genes in arachnid mitochondrial genomes. Mol. Biol. Evol. 25, 949–959.
Ohta, T., 1995. Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. J. Mol. Evol. 40, 56–63.
Ojala, D., Montoya, J., Attardi, G., 1981. TRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.
Osca, D., Templado, J., Zardoya, R., 2015. Caenogastropod mitogenomics. Mol. Biol. Evol. 93, 118–128.
Perna, N.T., Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41, 353–358.
Podsiadlowski, L., Braband, A., Mayer, G., 2008. The complete mitochondrial genome of the onychophoran *Epiperipatus biolleyi* reveals a unique transfer RNA set and provides further support for the ecdysozoa hypothesis. Mol. Biol. Evol. 25, 42–51.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Pu, C., Li, H., Zhu, A., Chen, Y., Zhao, Y., Zhan, A., 2017. Phylogeography in *Nassarius* mud snails: complex patterns in congeneric species. PLoS One 12, e0180728.
Ren, J., Liu, X., Jiang, F., Guo, X., Liu, B., 2010. Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. BMC Evol. Biol. 10, 394.
Ronquist, F., Teslenko, M., Mark, P.V.D., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Sevigny, J.L., Kirouac, L.E., Thomas, W.K., Ramsdell, J.S., Lawlor, K.E., Sharifi, O., Grewal, S., Baysdorfer, C., Curr, k., Naimie, A.A., Okamoto, K., Murray, J.A., Newcomb, J.M., 2015. The mitochondrial genomes of the nudibranch mollusks, *Melibe leonina* and *Tritonia diomedea*, and their impact on gastropod phylogeny. PLoS One 10, e0127519.
Shen, X., Meng, X.P., Chu, K.H., Zhao, N.N., Tian, M., Liang, M., Hao, J., 2014. Comparative mitogenomic analysis reveals cryptic species: a case study in Mactridae (Mollusca: Bivalvia). Comp. Biochem. Physiol. Part D Genomics Proteomics 12, 1–9.
Simison, W.B., Lindberg, D.R., Boore, J.L., 2006. Rolling circle amplification of metazoan mitochondrial genomes. Mol. Phylogenet. Evol. 39, 562–567.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Sun, S.E., Kong, L.F., Yu, H., Li, Q., 2015. The complete mitochondrial DNA of *Tegillarca granosa* and comparative mitogenomic analyses of three Arcidae species. Gene 557, 61–70.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45, 145–159.
Wang, D., Zhang, Y., Zhang, Z., Zhu, J., Yu, J., 2010. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics 8, 77–80.
Winnepenninckx, B., Backeljau, T., Dewachter, R., 1993. Extraction of high molecular weight DNA from molluscs. Trends Genet. 9, 407.
Wu, X., Xu, X., Yu, Z., Wei, Z., Xia, J., 2010. Comparison of seven *Crassostrea* mitogenomes and phylogenetic analyses. Mol. Phylogenet. Evol. 57, 448–454.
Xia, X., 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol. Biol. Evol. 30, 1720–1728.
Xin, Z.Z., Zhang, D.Z., Wang, Z.F., Zhang, H.B., Tang, B.P., Zhou, C.L., Chai, X.Y., Liu, Q.N., 2017. Mitochondrial genome of *Helice tientsinensis* (Brachyura: Grapsoidea:

Varunidae): gene rearrangements and higher-level phylogeny of the Brachyura. Gene 627, 307–314.
Xu, X., Wu, X., Yu, Z., 2012. Comparative studies of the complete mitochondrial genomes of four *Paphia* clams and reconsideration of subgenus *Neotapes* (Bivalvia: Veneridae). Gene 494, 17–23.
Yuan, Y., Li, Q., Yu, H., Kong, L.F., 2012. The complete mitochondrial genomes of six heterodont bivalves (Tellinoidea and Solenoidea): variable gene arrangements and phylogenetic implications. PLoS One 7, e32353.
Zhao, Q.P., Zhang, S.H., Deng, Z.R., Jiang, M.S., Nie, P., 2010. Conservation and variation in mitochondrial genomes of gastropods *Oncomelania hupensis* and *Tricula hortensis*, intermediate host snails of *Schistosoma* in China. Mol. Phylogenet. Evol. 57, 215–226.
Zou, S.M., Li, Q., Kong, L.F., 2012. Monophyly, distance and character–based multigene barcoding reveal extraordinary cryptic diversity in *Nassarius*: a complex and dangerous community. PLoS One 7, e47276.