Acta Oceanol. Sin., 2024, Vol. 43, No. 8, P. 98–103
https://doi.org/10.1007/s13131-024-2291-1
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Frequent recombination in *Cynoglossus abbreviatus* (Pleuronectiformes: Cynoglossidae) ribosomal 18S rDNA

Li Gong[1, 2], Tingqi Jiang[1, 2], Bilin Hu[1, 2], Kaixin Wang[1, 2], Nannan Zhang[1, 2], Zengliang Miao[1, 2*]

[1] National Engineering Laboratory of Marine Germplasm Resources Exploration and Utilization, Zhejiang Ocean University, Zhoushan 316022, China

[2] Marine Science and Technology College, Zhejiang Ocean University, Zhoushan 316022, China

Received 22 September 2023; accepted 11 November 2023

Abstract

The conventional theory of concerted evolution has been used to explain the lack of sequence variation in ribosomal RNA (rRNA) genes across diverse eukaryotic species. However, recent investigations into rRNA genes in flatfish genome have resulted in controversial findings. This study focuses on 18S rRNA genes of the widely distributed tongue sole, *Cynoglossus abbreviatus* (Pleuronectiformes: Cynoglossidae), aiming to explore sequence polymorphism. Five distinct 18S rDNA sequence types (Type A, B, R1, R2, and R3) were identified, suggesting a departure from concerted evolution. A combination of general criteria and variations in highly conserved regions were employed to detect pseudogenes. The results pinpointed Type A sequences as potential pseudogenes due to significant sequence variations and deviations in secondary structure within highly conserved regions. Three types (Type R1, R2, and R3) were identified as recombinants between Type A and B sequences, with simple crossing over and gene conversion as the most likely recombination mechanisms. These findings not only contribute to rRNA pseudogene identification but also shed light on the evolutionary dynamics of rRNA genes in teleost genomes.

Key words: ribosomal RNA, tongue sole, non-concerted evolution, pseudogene, crossing over, gene conversion

Citation: Gong Li, Jiang Tingqi, Hu Bilin, Wang Kaixin, Zhang Nannan, Miao Zengliang. 2024. Frequent recombination in *Cynoglossus abbreviatus* (Pleuronectiformes: Cynoglossidae) ribosomal 18S rDNA. Acta Oceanologica Sinica, 43(8): 98–103, doi: 10.1007/s13131-024-2291-1

1 Introduction

Nuclear ribosomal RNA genes (rDNA) are typically organized as large tandem repeat arrays on eukaryotic chromosomes (Edger et al., 2014), consisting of three rDNAs (18S, 5.8S, and 28S) interspersed with internal transcribed spacers (ITS1 and ITS2) and an intergenic spacer region (IGS) (Hillis and Dixon, 1991). For a long time, rDNA has been considered a prime example of concerted evolution. Two prevailing mechanisms, unequal crossing over and gene conversion, have been widely accepted to explain the homogenization process of rDNA sequences (Nagylaki and Petes, 1982; Smith, 1974). They suppose that all tandem repeat units evolve collectively to preserve sequence consistency. However, recent researches have uncovered instances of rDNA sequence polymorphism that challenge this paradigm, revealing a non-concerted or incomplete concerted evolution manner (Gong et al., 2019; Guo et al., 2021; Prudkovsky et al., 2023). Some hypotheses have been proposed to explain this phenomenon, wherein new genes arise from repeated duplication events, with some duplicates retained and others mutating into pseudogenes over time due to deleterious mutations (Nei et al., 1997; Nei and Rooney, 2005).

Pseudogenes have emerged as a primary contributor to rDNA polymorphism (Gong et al., 2019; Kolarik et al., 2021; Meng et al., 2022). Due to the blurred distinction between ribosomal functional gene and pseudogene sequences, and a lack of translation products of both functional genes and pseudogenes, it remains a challenging task to identify ribosomal pseudogenes. While most ribosomal pseudogenes share common features, including shorter sequence length, lower GC content, lower minimum free energy (MFE), less stable secondary structure, and increased insertion/deletion (indel) sites, exceptions exist (Gong et al., 2016a, 2021; Harpke and Peterson, 2008). A more convincing criterion involves comparative analysis of conserved regions, as even slight variations in the highly conserved regions will possibly alter secondary structure and impact functionality. Mutations at highly conserved sites in 18S and 5.8S rDNA, as well as variations in the secondary structure 5.8S rRNA, have proven useful in pseudogene identification (Gong et al., 2021; Meng et al., 2022). Hence, combining general criteria with mutations at highly conserved sites is a prudent approach for distinguishing pseudogenes from multiple rDNA genes.

Recombination also contributes to intragenomic rDNA polymorphism when rDNA follows a non-concerted evolution pattern, resulting in recombinants featuring multiple sequence types. Typically, recombinants involve combinations of functional genes and pseudogenes or different pseudogene variants (Meng et al., 2022; Wu et al., 2016). In the stone flounder *Kareius bicoloratus* genome, for instance, Xu et al. (2009) revealed the

Foundation item: The Basic Scientific Research Operating Expenses of Zhejiang Provincial Universities under contract 2021JZ003; the Zhoushan Science and Technology Bureau under contract No. 2021C21007; the Natural Science Foundation of Zhejiang Province under contract Y21C190023; the National Natural Science Foundation of China under contract 31272273.
*Corresponding author, E-mail: miaozl@zjou.edu.cn

first recombination event in rDNA in flatfish genome. In their study, four ITS1 sequences were detected as recombinants between functional genes and pseudogenes. In the red-spotted tongue sole *Cynoglossus zanzibarensis* genome, two ITS2 sequences were identified as recombinants between divergent paralogues, marking the first ITS2 recombination in animal rDNA (Gong et al., 2016b). In a recent study characterizing an 18S (partial)-ITS1-5.8S-ITS2-28S (partial) rDNA sequence in the *Cynoglossus trigrammus* genome, two types of recombinants were identified between functional gene and pseudogene sequences (Meng et al., 2022).

Previous studies suggested that simple crossing over, multiple crossing over, gene conversion (Lee et al., 2016; Smith, 1976; Stadler, 1959), and even accidental "jumping PCR" events can generate recombinants (Pääbo et al., 1990). Crossing over can be subdivided into simple crossing over and multiple crossing over. The former involves one recombination site, resulting in a single exchange between paralogues. The latter captures two or more recombination sites, leading to multiple exchanges and the generation of more complex recombinants (Navarro et al., 1997). Gene conversion, unlike crossing over, generates recombinants by non-reciprocally exchanging. In other words, the "acceptor" sequence is partially replaced by the sequence from the "donor", while the sequence of the donor remains unchanged (Roman, 1985). Experimental factors, such as "jumping PCR" by chance, may also produce pseudomorphic recombinants, particularly when long tandem repeat sequences (e.g. rDNA) act as PCR templates (Alaeddini et al., 2010; Pääbo et al., 1990). In "jumping PCR", Taq polymerase terminates extension prematurely and adds one or more extra adenosines (A) to the end of the premature product. In the next cycle, this premature product will act as a primer and recombine with other tandem repeats. Thus, after repeated replication, the recombinants generated by "jumping PCR" will be featured with autoidiomorphic As or Ts at the recombination sites.

Pleuronectiformes (flatfish) serve as a vital taxon for studies on species origin and adaptive evolution due to their specialized body plans. Recent studies have revealed variations in the evolutionary patterns of rRNA genes within flatfish genomes. For example, the rRNA gene in the majority of Cynoglossidae genomes exhibits incomplete concerted evolution (Gong et al., 2018b, 2021), while those in closely related Soleidae genomes follow a concerted evolution manner (Gong et al., 2018a). While increased ITS (ITS1 and ITS2) sequence polymorphism in flatfish genomes has been attributed to their moderate amplification lengths (Gong et al., 2016b; Meng et al., 2022), studies of long ribosomal RNA coding genes (e.g. 18S rRNA) have been scarce (Gong et al., 2016a; Xu et al., 2009). This study centers on the widely distributed tongue sole, *Cynoglossus abbreviates* (Luo, 2020), to investigate 18S rDNA sequence polymorphism. Our results reveal substantial intra-individual variations in 18S rDNA sequences, allowing for the analysis of functional genes, putative pseudogenes, and their recombinants. These results not only enrich our understanding of ribosomal pseudogene identification criteria but also provide valuable insights into the evolutionary manner of rRNA genes.

2 Materials and methods

2.1 *Sampling and DNA sequencing*

Two specimens of *C. abbreviatus* were collected from Ningbo, Zhejiang Province, and Rizhao, Shandong Province, China, and immediately stored in 95% ethanol. Genomic DNA was extracted from muscle tissue samples following the marine animal genomic DNA extraction kit's instructions (Tiangen Biotech Co., China). A pair of universal primers (18SF: 5′ - TCTGGTTGATTCTGCCAG TAG - 3′ and 18SR: 5′ - ATGATCCTTCCGCAGGTTCA - 3′) (Xu et al., 2009) was employed to amplify the complete 18S rDNA sequence. PCR was conducted in 50 mL volumes containing 4.0 mol/L $MgCl_2$, 0.8 mol/L of each dNTP, 1.0 mol/L of each primer, 2.0 U of Taq polymerase (TaKaRa, Beijing, China), 5.0 ml of 10 x Taq buffer, and approximately 100 ng of DNA template. The PCR cycling conditions included an initial denaturation at 94℃ for 2 min, followed by 35 cycles at 94℃ for 1 min, an annealing temperature of 52℃ for 2.5 min, elongation at 72℃ for 2 min, and a final extension at 72℃ for 10 min. The PCR products were purified using the TaKaRa Agarose Gel DNA Purification Kit (TaKaRa, Beijing, China) and subsequently inserted into the pMD19-T vector (TaKaRa, Beijing, China). Positive clones were subjected to bidirectional sequencing using an ABI 3 730 DNA sequencer (Applied Biosystems, USA).

2.2 *Sequence alignment and polymorphism analysis*

Sequence alignment was performed using ClustalX 2.0 (Larkin et al., 2007), with manual inspection using BioEdit v7 (Hall, 1999). Base composition and polymorphic sites were determined using MEGA X (Kumar et al., 2018). Haplotype diversity (Hd) and nucleotide diversity (π) analyses were conducted using DNAsp software (Librado and Rozas, 2009). Secondary structures were predicted using the RNAfold web server, with all default parameters (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Minimum free energy (ΔG at 37℃) was estimated to compare structural stability.

2.3 *Identification of recombination and pseudogenes*

Recombination analysis was conducted using the RDP 4 program (Martin et al., 2015). Pseudogene identification followed general criteria, including shorter sequence length, lower GC content, lower minimum free energy (MFE), and less stable secondary structure (Bailey et al., 2003; Li et al., 2017; Zuriaga et al., 2015). Additionally, the presence of variable sites within highly conserved 18S sequences and deviations from the consistent secondary structure were considered for detecting putative pseudogenes.

3 Results and discussion

3.1 *18S rDNA sequence polymorphism*

A total of 25 isolate clones were obtained (Genbank accession no. OR794293-OR794317), representing complete 18S rDNA sequences from two *C. abbreviatus* specimens (12 from Ningbo, Zhejiang Province, and 13 from Rizhao, Shandong Province). Sequence length showed considerable variability, ranging from 1 869 base pairs (bp) to 1 909 bp. Sequence alignment revealed that three main 18S rDNA sequence types: Type A (11 clones) with a consistent length of 1 909 bp, Type B (3 clones) with a consistent length of 1 869 bp, and Type R (11 clones) with varying lengths between 1 871 bp to 1 903 bp. Further analysis confirmed Type R sequences as recombinants between Type A and B. Two recombination sites were identified, resulting in three types of recombinants (Type R1, R2, and R3) that contributed to length heteroplasmy among recombinants (Table 1, Fig. 1).

In addition to length variation, the main 18S rDNA sequence types differed in terms of GC content, minimum free energy, secondary structure, and the number of variable sites. Type A and R1 sequences exhibited slightly higher GC content (57.7%–57.8%)

Table 1. 18S rDNA sequence polymorphism in *Cynoglossus abbreviatus* genome

	Type A	Type B	Type R1	Type R2	Type R3	All
No. of clone	11	3	7	2	2	25
Length/bp	1 909	1 869	1 903	1 875	1 871	1 869–1 909
No. of haplotype	10	2	7	2	2	23
Haplotype diversity	0.982	0.667	1.000	1.000	1.000	0.993
Variable site	52	9	26	22	1	215
Parsimony-informative site	18	0	2	0	0	149
Nucleotide diversity (π)	0.006 4	0.003 2	0.004 1	0.011 7	0.000 5	0.032 8
GC content*/%	57.7	54.5	57.8	54.8	56.1	57.0
Minimum free energy*/(kcal·mol^{-1})	−762.7	−710.9	−764.1	−711.3	−740.1	−751.7

Note: * indicates average GC content and minimum free energy.

than other types (54.5%–56.1%), while Type B sequences had the lowest GC content at 54.5%. Correspondingly, Type A and R1 sequences had higher minimum free energy values (absolute value), while Type B sequences exhibited the lowest minimum free energy (−710.9 kcal/mol) (Table 1). Secondary structures were predicted based on the minimum free energy principle, with Type R1 and A sharing almost identical secondary structures. Type R2, R3, and B sequences displayed similar secondary structures. Notably, these secondary structures corresponded closely to sequence similarities (Fig. S1).

The rRNA gene family has traditionally been thought to undergo concerted evolution, leading to homogenization among all repetitive units. However, increasing evidence (Gong et al., 2019; Guo et al., 2021; Prudkovsky et al., 2023), including our findings in this study, has demonstrated the coexistence of multiple sequence types within a single genome. This observation suggests that the mechanisms of concerted evolution may not adequately counteract variation forces. The choice between concerted evolution or non-concerted evolution highly depends on the interplay between the homogenization processes and the mutation rate. When the rate of mutation surpasses the effectiveness of homogenization processes, new copies of genetic units emerge over time. In this context, our discovery of three distinct 18S rDNA sequence types, especially the recombination events, indicates that 18S rRNA genes are undergoing a highly dynamic evolutionary process.

3.2 *Recombinants between divergent 18S rDNA paralogues*

In this study, three primary 18S rDNA sequence types (Type A, B, and R) were identified, with one of them (Type R) classified as a recombinant. Further analysis revealed that Type R sequences consisted of three recombinant types (Type R1, R2, and R3). Type R1 and R2 sequences shared the same recombination site (site 1 792) between Type A and B sequences. Type R1 sequences were composed of Type A sequence in the upstream region (1 bp–1 792 bp) and Type B sequence in the downstream region (1 793 bp–1 913 bp), while Type R2 sequences exhibited the opposite nucleotide composition, featuring Type B sequence in the forward part (1 bp–1 792 bp) and Type A sequence in the remaining part (1 793 bp–1 913 bp). Type R3 had an additional recombination site (site 1 181) and consisted of three segments from Type A and B sequences. The upstream (1 bp–1 181 bp) and downstream sequence (1 793 bp–1 913 bp) originated from Type B, with the intermediate sequence (1 182 bp–1 792 bp) derived from Type A. However, the corresponding recombinant type (Type R4) was not detected (Fig. 1b).

As mentioned in the introduction, simple or multiple crossing over, gene conversion and experimental "jumping PCR" can result in recombination. In this study, only one recombination site was detected in Type R1 and Type R2. Interestingly, these two types of recombinants exactly form the complementary counterpart between Type A and Type B, separated by the same recombination site (site 1 792) (Fig. 2b). Consequently, the generation of Type R1 and Type R2 can easily be explained by simple crossing over through reciprocally exchanged process.

Regarding the Type R3 sequence, simple crossing over is not appropriate to explain the recombination event, because it possesses two recombination sites (sites 1 181 and 1 792). Then, it is naturally to consider the multiple crossing over hypothesis. We assume that the sequence breaks concurrently at sites 1 181 and 1 792; theoretically, it will produce two types of recombinants through reciprocally exchanged processes after rejointing. Type R3 is one type of the recombinants after a multiple crossing over process; however, its reciprocal counterpart (Type R4) has not been detected (Fig. 2c). Gene conversion can generate recombinants by non-reciprocally exchanging between the "acceptor" and "donor" sequences. In such way, only one type of recombinant will be generated. Here, Type A sequence serves as the "donor", and Type R3 sequence serves as the "acceptor", which can be explained by gene conversion perfectly (Fig. 2d). However, we cannot guarantee whether Type R4 was naturally absent or just missed artificially because of the limited clones (25 clones) selected randomly.

Of course, experimental factors, such as "jumping PCR" by chance, may also produce pseudomorphic recombinants, especially when long tandem repeat sequences act as PCR templates (Alaeddini et al., 2010; Pääbo et al., 1990). Here, the tandem repeat 18S rDNA works as a template, providing a prerequisite for jumping PCR. According to the possible processes of jumping PCR, the recombinants will be featured with autoidiomorphic As or Ts at the recombination sites. In this study, a triplet guanine (G) follows the first recombination site (site 1 181) in the recombinants. In the second recombination site (site 1 792), a double cytimidine (C) exists in Type R1 and R3; and a double guanine (G) exists in Type R2. Neither As nor Ts were present in the recombination sites, which did not align with the characteristics of "jumping PCR". Hence, "jumping PCR" is not a suitable explanation for the recombinants in this study.

3.3 *18S rRNA pseudogene identification*

Identifying rRNA pseudogenes presents a complex and essential challenge. Typically, pseudogenes exhibit characteristics, such as shorter sequence length, lower GC content, reduced minimum free energy (MFE), and less stable secondary structure (Meng et al., 2022; Mighell et al., 2000; Zuriaga et al., 2015). However, it is worth noting that exceptions to this general criterion have been documented, as discussed in the introduction (Gong et al., 2016a, 2021; Gong et al.,Harpke and Peterson, 2008). For example, two types of 18S rDNA sequences coexist in *Cynoglossus melampetalus* genome, with Type A sequences being

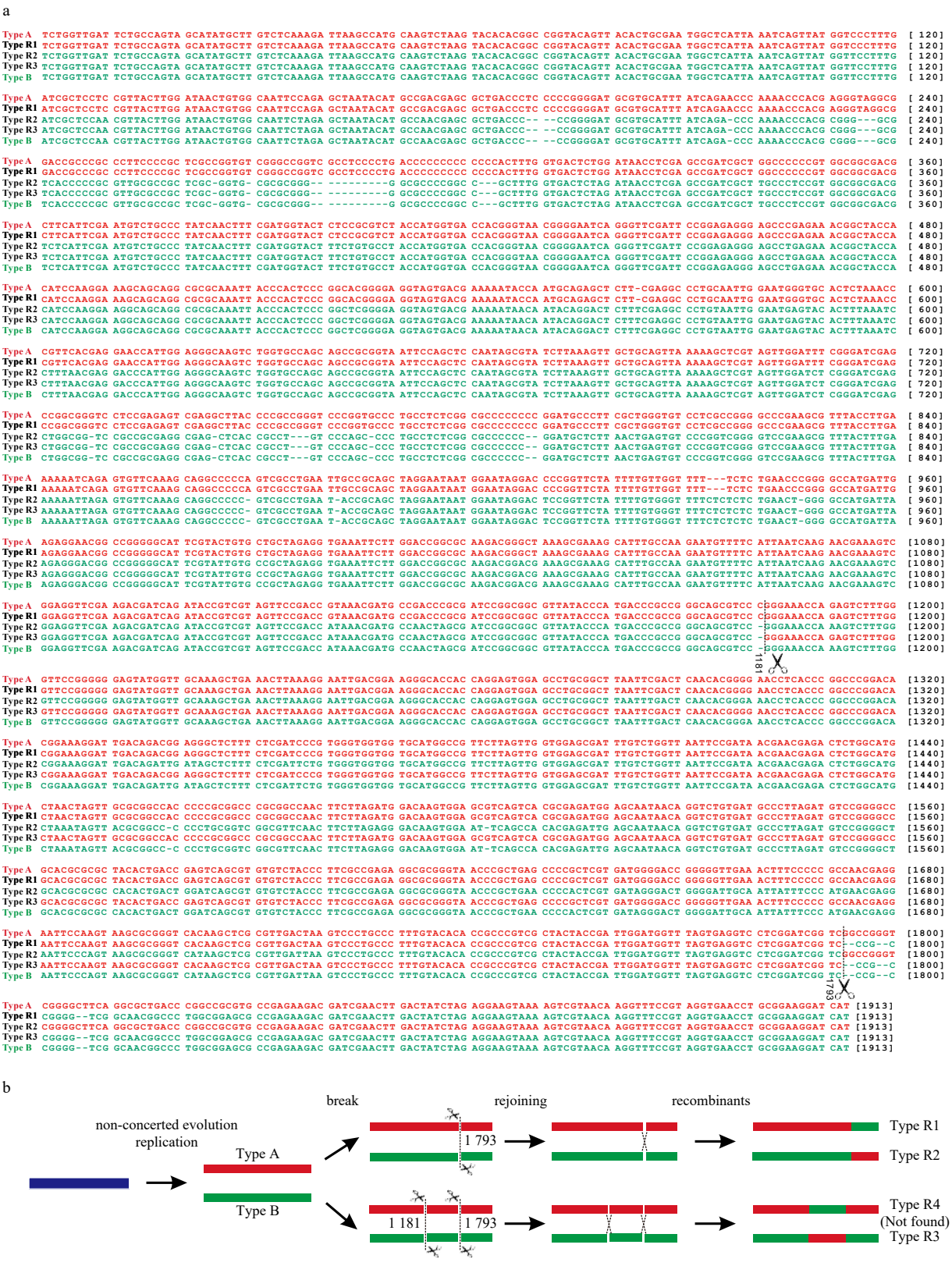

Fig. 1. Sequence alignment of distinct 18S rDNA sequences within the *Cynoglossus abbreviatus* genome (a) and a schematic representation illustrating the plausible recombinant generation process (b). Dashes (–) indicate deletions or insertions, with colors used to distinguish different types of sequences.

slightly shorter (4 bp) and having slightly lower GC content (0.2%) compared to Type B sequences. However, the secondary structure of Type A sequences is more stable, with a higher MFE (absolute value), which contradicts the expectations based solely on length and GC content. To resolve this, the authors compared the conserved regions of these sequences with those of 21 closely

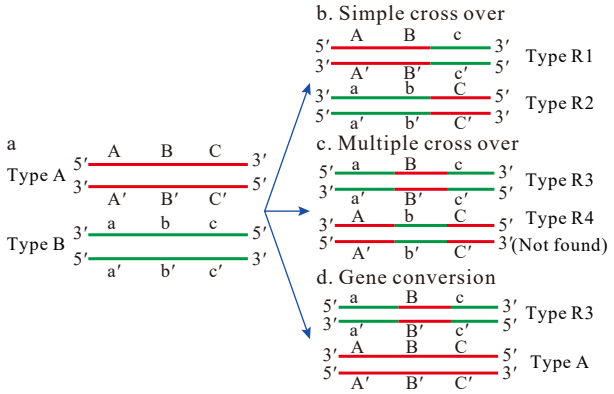

Fig. 2. A schematic illustration depicting potential mechanisms for the formation of recombinants.

related species and found that Type B sequences had specific insertions in highly conserved regions. Consequently, Type B sequences were classified as pseudogenes. Similar cases have occurred in the 18S rDNA of *C. lineolatus* (Gong et al., 2016a) and the 5.8S rDNA of *Mammillaria* (Harpke and Peterson, 2008). Therefore, a combination of general criteria and deviation in highly conserved regions may be a reliable approach for identifying pseudogenes.

In this study, Type B sequences were considered putative pseudogenes based on general criteria, including their shorter sequence length, lower GC content, and lower MFE. However, due to the limited number of clones, other features, such as haplotype diversity and variable sites, were not considered. In order to verify this inference, we compared the 18S rDNA sequences with those of 15 other Soleoidei species. Interestingly, Type B sequences were highly conserved across these species. In contrast, Type A sequence exhibited notable variations, including 25 specific singleton sites and 16 specific insertions (Fig. 3). Additionally, we compared the secondary structures of different 18S rDNA types, using the functional sequences of *Paraplagusia japonica* (Cynoglossidae) as reference. This analysis revealed that Type B sequences had a highly conserved secondary structure similar to that of the functional 18S rRNA of *P. japonica*, including the minimum free energy (Fig. S1). Conversely, Type A and other recombinant sequences exhibited greatly divergent structures. Pseudogenes are expected to accumulate mutations even in conserved regions, and the substantial sequence variations and secondary structure deviation in Type A sequences strongly indicate that they are putative pseudogenes. Therefore, we recommend using a combination of general criteria and variations at highly conserved sites as a robust approach for pseudogene identification. In many cases, the latter criterion carries more weight.

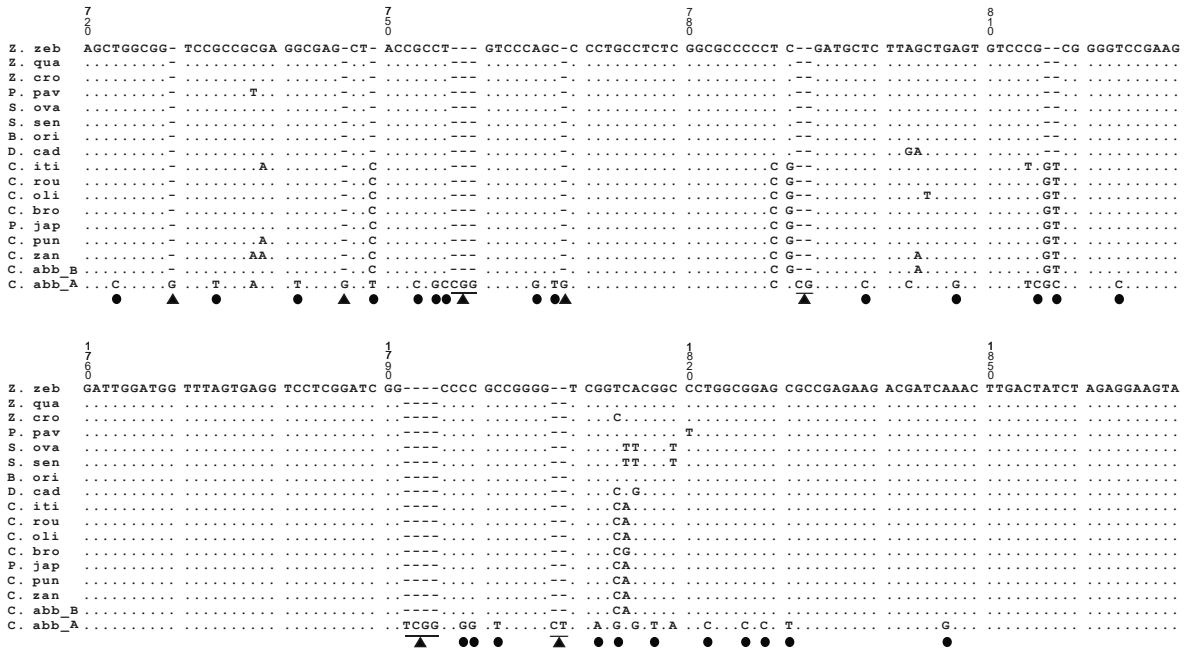

Fig. 3. Sequence alignment of 18S rDNA in Soleoidei species (partial). The specific singleton sites and insertions in C. pur_A are labelled with solid circles and triangles; respectively. Dots (.) denote the conserved sites, dashes (–) denote the deletion or insertion. Abbreviations in 16 Soleoidei species are as follows. Z. zeb: *Zebrias zebrinus*; Z. qua: *Zebrias quagga*; Z. cro: *Zebrias crossolepis*; P. pav: *Pardachirus pavoninus*; S. ova: *Solea ovata*; S. sen: *Solea senegalensis*; B. ori: *Brachirus orientalis*; D. cad: *Dagetichthys cadenati*; C. iti: *Cynoglossus itinus*; C. rou: *Cynoglossus roulei* C. oli: *Cynoglossus oligolepis*; C. bro: *Cynoglossus browni*; P. jap: *Paraplagusia japonica*; C. pun: *Cynoglossus puncticeps*; C. zan: *Cynoglossus zanzibarensis*; C.abb_B: *Cynoglossus abbreviatus* Type B; C.abb_A: *Cynoglossus abbreviatus* Type A.

Acknowledgements

We are grateful to Xiaoyu Kong (South China Sea Institute of Oceanology, Chinese Academy of Sciences) in assisting sample collection and identification. We appreciate the linguistic assistance provided by TopEdit (www.topeditsci.com) during the preparation of this manuscript.

References

Alaeddini, R. , S. J. Walsh & A. Abbas, 2010. Forensic implications of genetic analyses from degraded DNA—a review. Forensic science international: genetics 4(3): 148–157.

Bailey, C. D. , T. G. Carr, S. A. Harris, et al. , 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol Phylogenet Evol 29(3): 435–455.

Edger, P. P. , M. Tang, K. A. Bird, et al., 2014. Secondary structure analyses of the nuclear rRNA internal transcribed spacers and assessment ofits phylogenetic utility across the Brassicaceae (Mustards). Plos One 9(7), doi: 10.1371/journal.pone.0101341

Gong, L. , H. Luo, W. Shi, et al. , 2019. Intra-individual variation and transcribed pseudogenes in the ribosomal ITS1-5.8S-ITS2 rDNA of *Paraplagusia japonica* (Pleuronectiformes: Cynoglossidae). Biochem Bioph Res Co 513(3): 726–731, doi: 10.1016/j.bbrc.2019.04.064

Gong, L. , W. Shi, M. Yang, et al. , 2018a. Characterization of 18S-ITS1-5.8S rDNA in eleven species in Soleidae: implications for phylogenetic analysis. Hydrobiologia 819(1): 161–175, doi: 10.1007/s10750-018-3634-8

Gong, L. , W. Shi, M. Yang, et al. , 2018b. Marked intra-genomic variation and pseudogenes in the ITS1-5.8S-ITS2 rDNA of *Symphurus plagiusa* (Pleuronectiformes: Cynoglossidae). Animal Biology 68(4): 353–365, doi: 10.1163/15707563-17000134

Gong, L. , W. Shi, M. Yang, et al. , 2021. Variations in the conserved 18S and 5.8 S reveal the putative pseudogenes in 18S-ITS1-5.8 S rDNA of *Cynoglossus melampetalus* (Pleuronectiformes: Cynoglossidae). Biochem Bioph Res Co 534: 233–239.

Gong, L. , W. Shi, M. Yang, et al. , 2016a. Long duplication of 18S ribosomal DNA in *Cynoglossus lineolatus* (Pleuronectiformes: Cynoglossidae): novel molecular evidence for unequal crossing over model. Acta Oceanologica Sinica 35(12): 38–50.

Gong, L. , W. Shi, M. Yang, et al. , 2016b. Non-concerted evolution in ribosomal ITS2 sequence in *Cynoglossus zanzibarensis* (Pleuronectiformes: Cynoglossidae). Biochem Syst Ecol 66: 181–187, doi: 10.1016/j.bse.2016.04.002

Guo, Z. S. , Z. Wang & X. G. Hou, 2021. Comparative Analysis of the nrDNA Repeat Unit of Manila Clam *Ruditapes philippinarum* and Quahog *Mercenaria mercenaria*. FISHES 6(3), doi: 10.3390/fishes6030042

Hall, T. A. , 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41(41): 95–98.

Harpke, D. & A. Peterson, 2008. Extensive 5.8 S nrDNA polymorphism in *Mammillaria* (Cactaceae) with special reference to the identification of pseudogenic internal transcribed spacer regions. J Plant Res 121(3): 261–270.

Hillis, D. M. & M. T. Dixon, 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Q Rev Biol 66(4): 411–453.

Kolarik, M. , I. C. Wei, S. Y. Hsieh, et al. , 2021. Nucleotide composition bias of rDNA sequences as a source of phylogenetic artifacts in Basidiomycota-a case of a new lineage of a uredinicolous Ramularia-like anamorph with affinities to Ustilaginomycotina. Mycological Progress 20(12): 1553–1571, doi: 10.1007/s11557-021-01749-x

Kumar, S. , G. Stecher, M. Li, et al. , 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6): 1547–1549, doi: 10.1093/molbev/msy096

Larkin, M. A. , G. Blackshields, N. P. Brown, et al. , 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23(21): 2947–2948, doi: 10.1093/bioinformatics/btm404

Lee, M. O. , S. Bornelov, L. Andersson, et al. , 2016. Duplication of chicken defensin7 gene generated by gene conversion and homologous recombination. P Natl Acad Sci Usa 113(48): 13815–13820, doi: 10.1073/pnas.1616948113

Li, Y. , R. H. Yang, L. Jiang, et al. , 2017. rRNA Pseudogenes in Filamentous Ascomycetes as Revealed by Genome Data. G3-Genes Genom Genet 7(8): 2695–2703, doi: 10.1534/g3.117.044016

Librado, P. & J. Rozas, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452.

Luo, H. , 2020. Study on Validity in Species of Cynoglossidae and Mechanisms of Gene Rearrangement in Mitogenomes of Bothidae Living in Coastal Waters off China (Teleostei: Pleuronectiformes). University of Chinese Academy of Sciences.

Martin, D. P. , B. Murrell, M. Golden, et al. , 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus evolution 1(1): vev003.

Meng, L. , Y. Gao & L. Gong, 2022. Recombination and incomplete concerted evolution of the ribosomal 18S (partial) -ITS1-5.8S-ITS2-28S (partial) rDNA in *Cynoglossus trigrammus* genome. Biochem Syst Ecol 105: 104513, doi: 10.1016/j.bse.2022.104513

Mighell, A. , N. Smith, P. Robinson, et al. , 2000. Vertebrate pseudogenes. Febs Lett 468(2): 109–114.

Nagylaki, T. & T. D. Petes, 1982. Intrachromosomal Gene Conversion and the Maintenance of Sequence Homogeneity among Repeated Genes. Genetics 100(2): 315–337.

Navarro, A. , E. Betrán, A. Barbadilla, et al. , 1997. Recombination and Gene Flux Caused by Gene Conversion and Crossing Over in Inversion Heterokaryotypes. Genetics 146(2): 695–709, doi: 10.1093/genetics/146.2.695

Nei, M. , X. Gu & T. Sitnikova, 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. P Natl Acad Sci 94(15): 7799–7806.

Nei, M. & A. P. Rooney, 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39: 121–152, doi: 10.1146/annurev.genet.39.073003.112240

Pääbo, S. , D. M. Irwin & A. C. Wilson, 1990. DNA damage promotes jumping between templates during enzymatic amplification. J Biol Chem 265(8): 4718-4721.

Prudkovsky, A. , A. Vetrova & S. Kremnyov, 2023. What Does "ITS" Say about Hybridization in Lineages of Sarsia (Corynidae, Hydrozoa) from the White Sea? Diversity 15(5): 675.

Roman, H. , 1985. Gene conversion and crossing-over. Environmental Mutagenesis 7(6): 923–932, doi: 10.1002/em.2860070614

Smith, G. P. , 1974. Unequal crossover and the evolution of multigene families. Cold Spring Harb Sym 38: 507–513.

Smith, G. P. , 1976. Evolution of repeated DNA sequences by unequal crossover. Science 191(4227): 528–535, doi: 10.1126/science.1251186

Stadler, D. R. , 1959. The relationship of gene conversion to crossing over in *Neurospora*. P Natl Acad Sci 45(11): 1625–1629, doi: 10.1073/pnas.45.11.1625

Wu, Z. W. , Q. M. Wang, X. Z. Liu, et al. , 2016. Intragenomic polymorphism and intergenomic recombination in the ribosomal RNA genes of strains belonging to a yeast species Pichia membranifaciens. Mycology 7(3): 102–111, doi: 10.1080/21501203.2016.1204369

Xu, J. , Q. Zhang & X. Xu, 2009. Intragenomic variability and pseudogenes of ribosomal DNA in Stone flounder *Kareius bicoloratus*. Mol Phylogenet Evol 52(1): 157–166.

Zuriaga, M. A., S. Mas-Coma & M. D. Bargues. 2015. A nuclear ribosomal DNA pseudogene in triatomines opens a new research field of fundamental and applied implications in Chagas disease. Mem I Oswaldo Cruz, 110(3): 353–362

Supplementary information:

Figure S1. Inferred secondary structures of different type sequences in 18S rRNA.

The supplementary information is available online at https://doi.org/10.1007/s13131-024-2291-1 and http://www.aosocean.com. The supplementary information is published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.