

RESEARCH ARTICLE

Open Access



# Mechanisms of gene rearrangement in 13 bothids based on comparison with a newly completed mitogenome of the threespot flounder, *Grammatobothus polyophthalmus* (Pleuronectiformes: Bothidae)

Hairong Luo<sup>1,2</sup>, Xiaoyu Kong<sup>1\*</sup> , Shixi Chen<sup>1,2</sup> and Wei Shi<sup>1</sup>

## Abstract

**Background:** The mitogenomes of 12 teleost fish of the bothid family (order Pleuronectiformes) indicated that the genomic-scale rearrangements characterized in previous work. A novel mechanism of genomic rearrangement called the Dimer-Mitogenome and Non-Random Loss (DMNL) model was used to account for the rearrangement found in one of these bothids, *Crossorhombus azureus*.

**Results:** The 18,170 bp mitogenome of *G. polyophthalmus* contains 37 genes, two control regions (CRs), and the origin of replication of the L-strand ( $O_L$ ). This mitogenome is characterized by genomic-scale rearrangements: genes located on the L-strand are grouped in an 8-gene cluster (*Q-A-C-Y-S<sub>1</sub>-ND6-E-P*) that does not include *tRNA-N*; genes found on the H-strand are grouped together (*F-12S ... CytB-T*) except for *tRNA-D* that was translocated inside the 8-gene L-strand cluster. Compared to non-rearranged mitogenomes of teleost fishes, gene organization in the mitogenome of *G. polyophthalmus* and in that of the other 12 bothids characterized thus far is very similar. These rearrangements could be sorted into four types (Type I, II, III and IV), differing in the particular combination of the CR, *tRNA-D* gene and 8-gene cluster and the shuffling of *tRNA-V*. The DMNL model was used to account for all but one gene rearrangement found in all 13 bothid mitogenomes. Translocation of *tRNA-D* most likely occurred after the DMNL process in 10 bothid mitogenomes and could have occurred either before or after DMNL in the three other species. During the DMNL process, the *tRNA-N* gene was retained rather than the expected *tRNA-N'* gene. *tRNA-N* appears to assist in or act as  $O_L$  function when the  $O_L$  secondary structure could not be formed from intergenic sequences. A striking finding was that each of the non-transcribed genes has degenerated to a shorter intergenic spacer during the DMNL process. These findings highlight a rare phenomenon in teleost fish.

**Conclusions:** This result provides significant evidence to support the existence of dynamic dimeric mitogenomes and the DMNL model as the mechanism of gene rearrangement in bothid mitogenomes, which not only promotes the understanding of mitogenome structural diversity, but also sheds light on mechanisms of mitochondrial genome rearrangement and replication.

**Keywords:** Flatfish, Bothidae, Mitogenome, Rearrangement mechanism, Dimeric mitogenomes, Intergenic spacer

\* Correspondence: [xykong@scsio.ac.cn](mailto:xykong@scsio.ac.cn)

Please note that this article was previously reviewed at BMC Genomics, but due to a technical inconvenience it had to be resubmitted at pre-acceptance stage.

<sup>1</sup>CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, South China Sea Institute of Oceanology, Guangzhou 510301, China

Full list of author information is available at the end of the article



## Background

The mitogenome of most fish contains 37 genes, including 13 protein-coding genes, two ribosomal RNA (rRNA), and 22 transfer RNA (tRNA) genes. Most of these genes are located on the heavy strand (H-strand), only *ND6* and eight tRNA genes (*N*, *Q*, *A*, *C*, *Y*, *S<sub>I</sub>*, *E* and *P*) are located on the light strand (L-strand). Additionally, the mitogenome contains two non-coding regions; the origin of replication of the L-strand ( $O_L$ ) and the control region (CR). The CR includes the origin of replication of the H-strand ( $O_H$ ) as well as the transcription initiation site for both the L- and H-strands [1–3]. Three types of gene rearrangements have been observed in the mitogenomes of animals: shuffling, translocation, and inversion [4–7]. Before gene inversion was discovered in tongue fish [8], only gene shuffling and translocation had been reported in fishes [9–11]. Since then, an increasing number of rearranged mitogenomes of flatfishes featuring all three gene rearrangement types have been found [12–16]. One such representative case was the mitochondrial gene rearrangement in the blue flounder, *Crossorhombus azureus* [14]. In this mitogenome, genes were grouped with identical transcriptional polarities, including a cluster of eight genes on the L-strand (8-gene cluster, *Q-A-C-Y-S<sub>I</sub>-ND6-E-P*) that did not include *tRNA-N*, and a cluster of genes (*F-12S ... CytB-T-D*) on the H-strand. The order of these genes in these two clusters was maintained as in the non-rearranged mitogenome of fish, except for the novel location of *tRNA-D*. Furthermore, unlike the typical position of the CR in fish, the CR of this species is located between *tRNA-D* and *tRNA-Q*, thus separating the two gene clusters on the H-strand and L-strand.

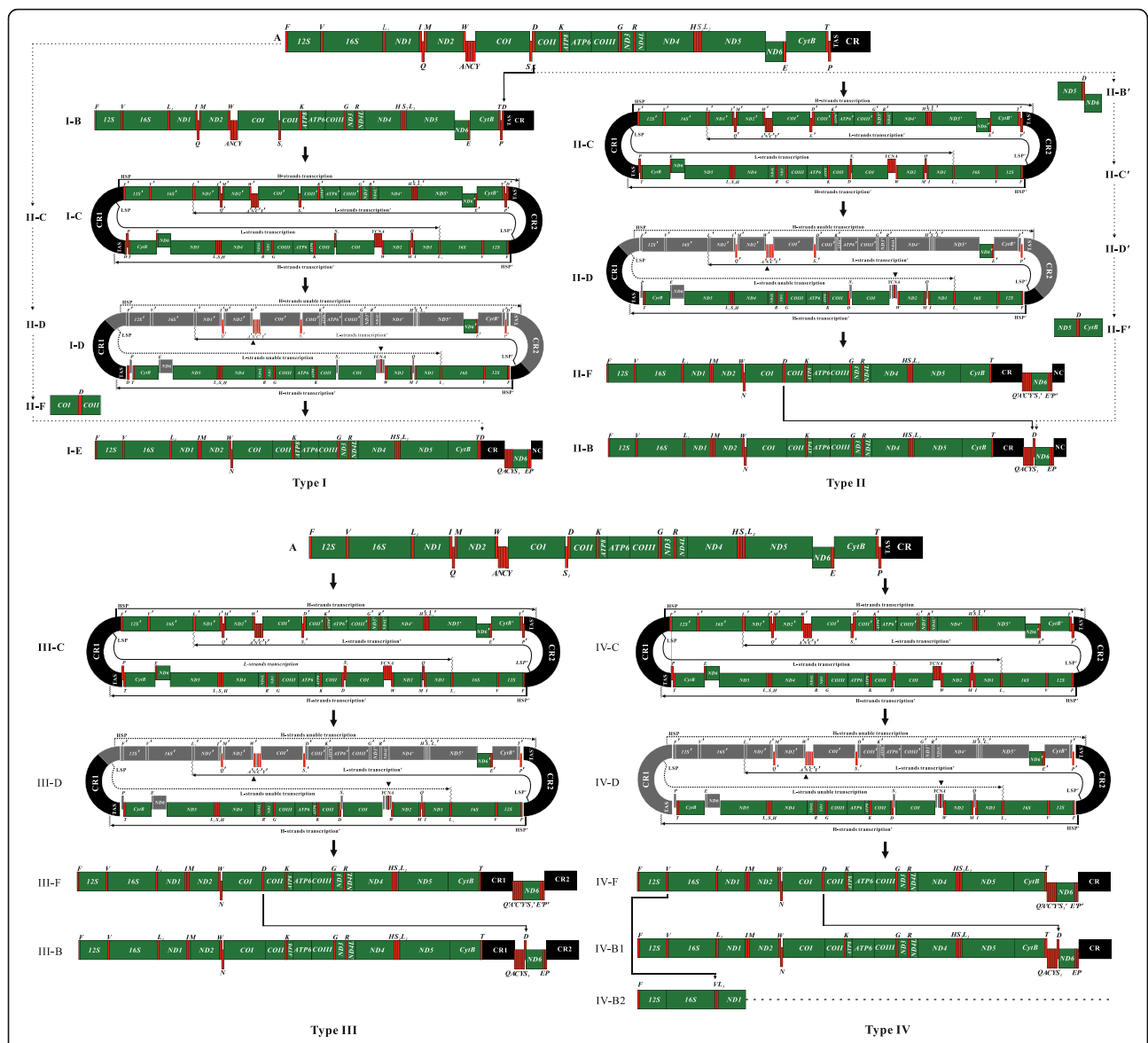
How did this particular mitogenome structure emerge? Four mechanisms have been proposed to account for mitogenomic rearrangements, including duplication-random loss [17], tRNA mis-priming model [18], intramitochondrial recombination [19], and duplication-nonrandom loss [20]. However, none of these four mechanisms can fully explain the gene rearrangements seen in the *C. azureus* mitogenome. Therefore, a novel mechanism called the Dimer-Mitogenome and Non-Random Loss (DMNL) model was proposed to account for the rearrangements found in *C. azureus* [14]. The inferred DMNL process would be as follows, starting with an ancestral mitogenome with a gene order typically seen in fish (Fig. 1a). The first step is a dimerized event of two monomer mitogenomes to form a functionally dimeric molecule linked head-to-tail (Fig. 1i-c). The dual promoter functions in one of two CRs are then lost by mutation and the genes controlled by these promoters are thus no longer transcribed (Fig. 1i-d), and may degenerate. The genes controlled by the remaining two functional promoters continue to be transcribed. Thus, the final gene order in the mitogenome of *C. azureus* is formed (Fig. 1i-e).

Retention of *tRNA-N* rather than the expected *tRNA-N'* gene occurred during the DMNL process. Shi et al. hypothesized that the exceptional retention of *tRNA-N* is related to the structure and function of  $O_L$  based on the study of Seligmann and Krishnan [14, 21].  $O_L$  is usually located between *tRNA-N* and *tRNA-C* of the so-called WANCY region, formed by a cluster tRNA genes that includes *tRNA-W*, *tRNA-A*, *tRNA-N*, *tRNA-C* and *tRNA-Y* [22]. Due to the rearrangement of *tRNA-C* and *tRNA-Y*, only a 7-bp intergenic region remained between *tRNA-N* and *COI*, an insufficient number of base pairs to allow for the necessary  $O_L$  secondary structure formation. Interestingly, a 26-bp middle portion of *tRNA-N* could form an  $O_L$ -like structure (Fig. 2a) for L-strand replication; this gene was therefore retained [14].

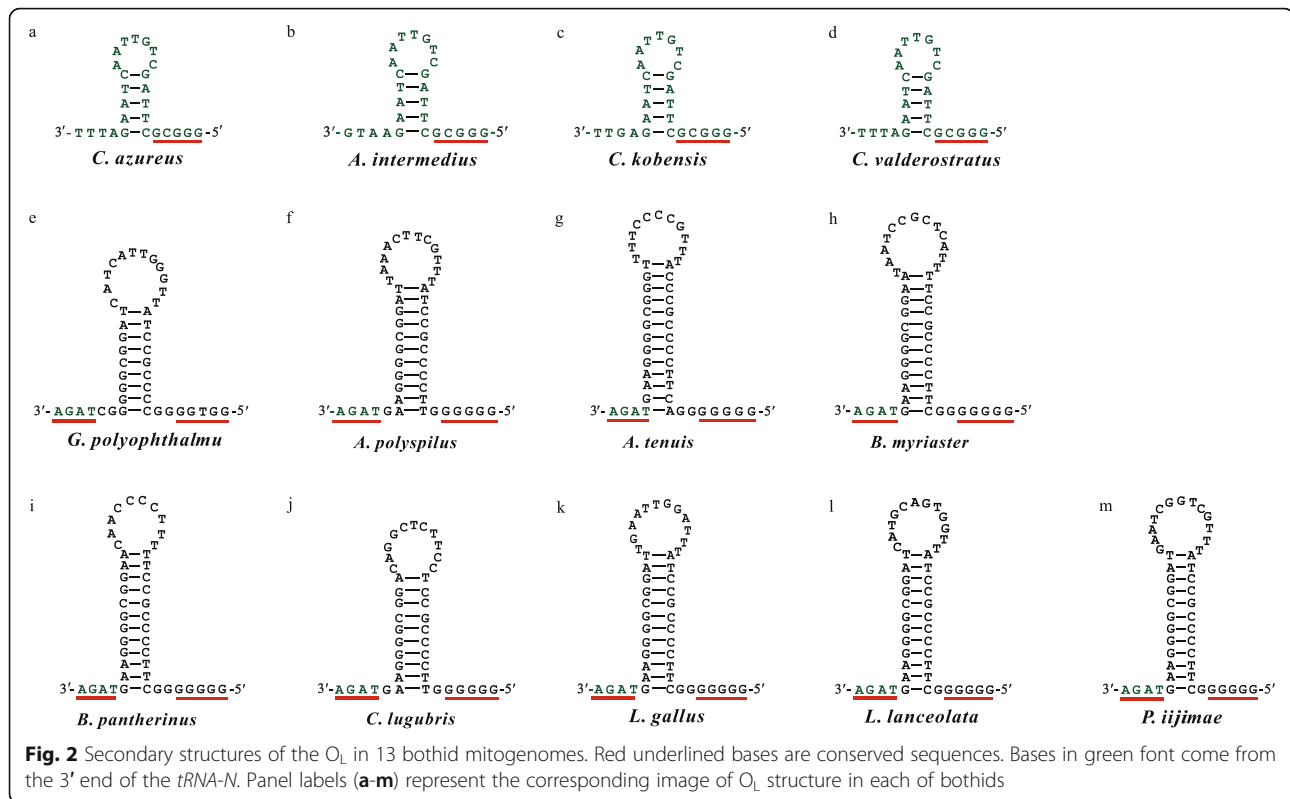
One question remains: when was *tRNA-D* translocated from an original site between *tRNA-S<sub>I</sub>* and *COII* to a region between *tRNA-T* and the CR in the *C. azureus* mitogenome? Shi et al. speculated that translocation of *tRNA-D* could have occurred either before (Fig. 1i-b) or after (Fig. 1i-e) the DMNL process [14]. Careful examination and comparison of available mitogenomes of 12 bothid species from eight genera revealed interesting patterns of gene arrangements. Gene organization in these 12 genomes is very similar, except for the shuffling *tRNA-V* (Fig. 1iV-B2), and the existence of one of the following four arrangements of CRs, *tRNA-D* and the 8-gene cluster (*Q-A-C-Y-S<sub>I</sub>-ND6-E-P*): *D*-CR-8-gene cluster, CR-5-gene cluster (*Q-A-C-Y-S<sub>I</sub>*)-*D*-3-gene cluster (*ND6-E-P*), CR1-5-gene cluster-*D*-3-gene cluster-CR2, and 5-gene cluster-*D*-3-gene cluster-CR.

The DMNL model was also used by Gong et al. [12] to explain the production of the mitogenome of *Bothus myriaster*, the only bothid species other than *C. azureus* for which a mechanism of gene rearrangement has been provided (Fig. 1a, II-c, d, f and b). One difference was noted between the gene structure of the *B. myriaster* and *C. azureus* mitogenomes. In *C. azureus*, *tRNA-D* was translocated outside of the 8-gene cluster (Fig. 1i-e), while in *B. myriaster*, *tRNA-D* was found inside this gene cluster (Fig. 1II-B). Whether *tRNA-D* translocation occurred before or after the DMNL process was not determined by Gong et al [12].

To better understand the mitochondrial gene structure and gene rearrangement mechanisms of bothids, the mitogenome of the threespot flounder, *Grammatobothus polyophthalmus*, was sequenced and characterized. This species is one of few bothids featuring one lateral line on both sides of the body. We wondered, what are the mitogenomic characteristics of this species? Did rearrangement occur in this mitogenome, and if so, what is the rearrangement type? Our results reveal additional mitochondrial diversity in Bothidae, and provide a foundation for further research on mitochondrial gene rearrangement of fish.



**Fig. 1** Four mitogenomic rearrangement variants in 13 bothdids using the DMNL model. The letters **a** to **f** after the variant number I-IV represent the steps of DMNL process. **a**: Ancestral gene order in typical fish. The rRNA, protein-coding genes and CRs are indicated by boxes; the tRNA genes are indicated by columns, genes labeled above the columns are located on the H-strand, those below the columns on the L-strand. TAS is an acronym for the terminal associated sequences. **b**: Translocated position of *tRNA-D* prior to DMNL gene rearrangement and the finally mitogenomic structure in Type I, and that of *tRNA-D* posterior to DMNL gene rearrangement and the finally mitogenomic structure in Type II, III, and IV. **c**: The dimeric molecule formed by two monomers linked head-to-tail. The HSP, HSP', LSP, and LSP' indicate the H- and L- strand promoters of transcription. The TAS, TAS', *tRNA-L<sub>1</sub>*, and *tRNA-L<sub>1</sub>'* indicate the H- and L- strand terminations of transcription. The direction of transcription is shown by an arrow. **d**: Functional loss of LSP and HSP. Grey boxes indicate the degenerated genes controlled by non-functional HSP and LSP. The triangle marks the retention of *tRNA-N* rather than the expected *tRNA-N'* gene. **e**: The translocation of *tRNA-D* posterior to DMNL gene rearrangement and the finally mitogenomic structure in Type I. **f**: The mitogenomic structure generated after degeneration of non-transcribed genes in the dimeric molecule. The path from a, II-C, II-D, II-F to I-E in Type I indicates the translocation of *tRNA-D* occurred after the DMNL process; the path from a, II-B', II-C', II-D', II-E' to II-B in Type II indicates the translocation of *tRNA-D* occurred before DMNL process; the difference between the steps of B', C', D' and F' with that of B, C, D and F is the location of *tRNA-D*. The broken line in IV-B2 indicates the omitted genes that are the same as those shown in IV-B1



## Results

### Organization and gene rearrangement of the *G. polyophthalmus* mitogenome

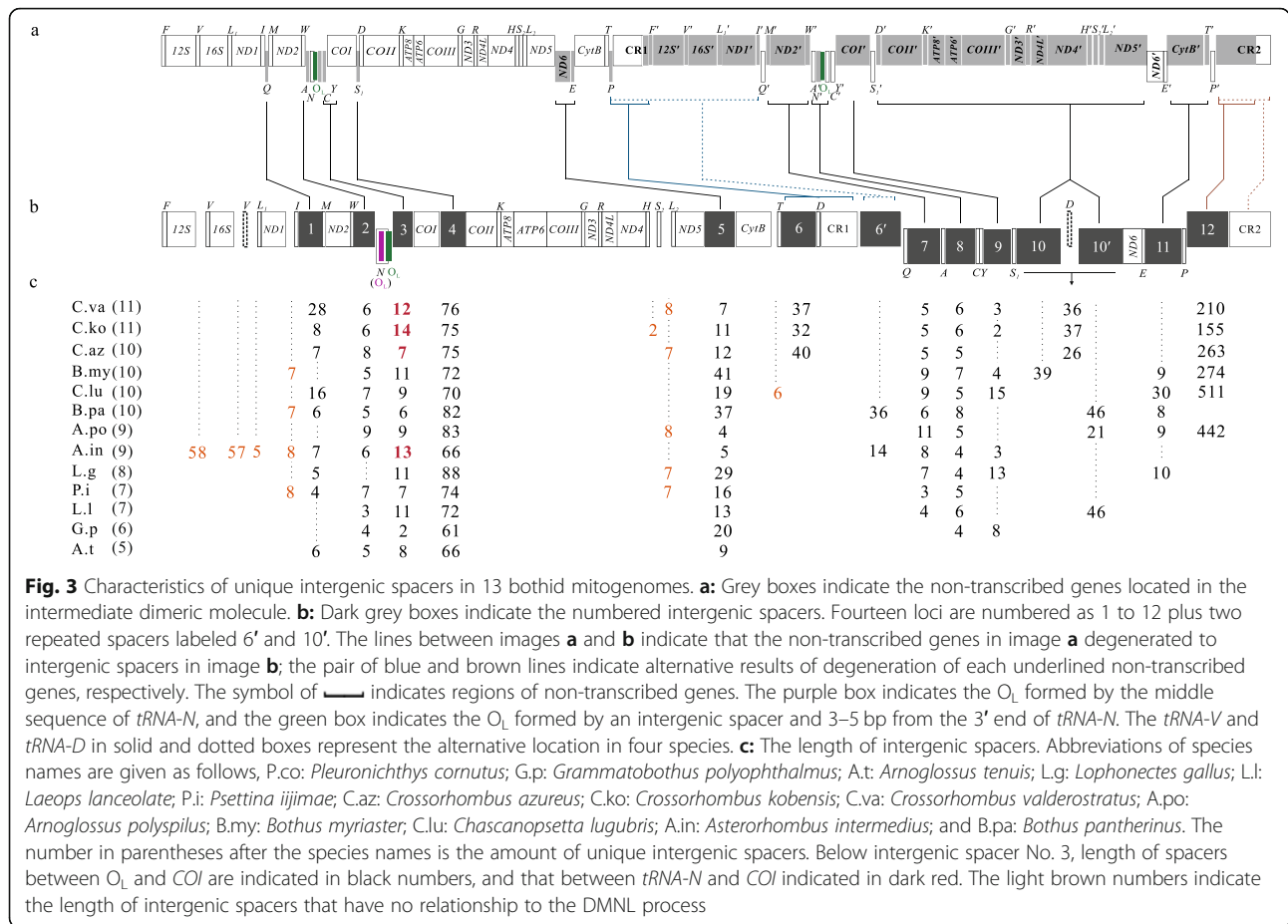
A total of 18,170 bp of *G. polyophthalmus* mitogenome contained 37 genes, including 13 protein-coding genes, two rRNA genes, and 22 tRNA genes. Among these genes, 28 genes are located on the H-strand, while *ND6* and eight tRNA genes (*N*, *Q*, *A*, *C*, *Y*, *S*<sub>1</sub>, *E* and *P*) are located on the L-strand (Fig. 1III-B, Additional file 1: Table S1 and Additional file 2: Figure S1). In this mitogenome, after *tRNA-C* and *tRNA-Y* were rearranged, a 40-bp intergenic spacer remained between *tRNA-N* and *COI*. The secondary structure of  $O_L$  was formed with 38 bp of this intergenic spacer and 4 bp of the 3' end of *tRNA-N* (Figs. 2e and 3b). The  $O_L$  had the same 5'-TAGA-3' sequence motif on the 3' end as that of eight other bothid species (Fig. 2f-m); the sequence 5'-GGTGG-3' on the 5' end was slightly different from either of the motifs 5'-GGGGG-3' seen in most bothids (Fig. 2f-m), or 5'-GCCGG-3' seen in most of the 17 flatfishes from seven families [23].

Two large non-coding (NC) regions were also discovered, NC1 is composed of 773 bp located between *tRNA-T* and *tRNA-Q*, while NC2 is comprised of 1611 bp between *tRNA-P* and *tRNA-F*. We compared NC1 and NC2 with conserved CR structures existed in 12 other bothid species and the ridged-eye flounder *Pleuronichthys cornutus*, and found similar conserved structures, including the terminal associated sequences (TAS) with the core

sequence ACAT-cTGTA; the conserved sequence blocks (CSB) of central conserved domain, CSB-F, E, D, C and B; the Pyrimidine T-tract; and the other two conserved sequence blocks, CSB-1 and -2 (Fig. 4). Additionally, NC2 had tandem repeated sequences (35 copies of a 22 bp motif) at the 3' end, as did CR or CR2 in the other species (Additional file 3: Figure S2). Based on sequence conservation of the NCs and the tandem repeats in NC2 of *G. polyophthalmus*, we conclude that both NC1 and NC2 are control regions, thus named CR1 and CR2, respectively (Fig. 1III-B and Fig. 4).

When the previously sequenced mitogenomes of 12 bothid species were compared with that of *G. polyophthalmus*, the gene order of this species was found to be identical to that of four other species, namely, *Arnoglossus tenuis*, *Lophonectes gallus*, *Laeops lanceolata*, and *Psettina iijimae*. This finding raises a question, how were genomic-scale rearrangements and two CRs generated in these five species? Although the mitogenomes of the above four bothids have been reported for years, the mechanism of their gene rearrangement remains unaddressed. Here, using the *G. polyophthalmus* mitogenome as a representative (Fig. 1 Type III), the process of gene rearrangement was found to be consistent with the DMNL model used to explain the evolution of the *C. azureus* and *B. myriaster* mitogenomes.

The process of gene rearrangement in these fishes can be reconstructed as follows, starting with the typical



mitogenome of fish (Fig. 1a). First, a dimerized event occurred to form a functionally dimeric mitochondrial DNA (mtDNA) (Fig. 1III-C). At this stage, transcription of genes from the dimeric mtDNA could be initiated normally using the H-strand promoters (HSP and HSP') and the L-strand promoters (LSP and LSP') located in the two CRs. Transcription from the H-strand would terminate at TAS and TAS' in the CRs, while on the L-strand *tRNA-L<sub>1</sub>* and *tRNA-L<sub>1</sub>'* would act as transcription terminators. Second, the function of the promoters (assumed to be LSP and HSP) in one of CRs was lost by mutation, thus the genes controlled by the disabled promoters could not be transcribed and then degenerated or disappeared entirely due to compact feature of mitochondrion (Fig. 1III-D). Consequently, the genes transcribed from LSP' were clustered together forming the 8-gene cluster "Q'-A'-C'-Y'-S<sub>1</sub>'-ND6'-E'-P'", and the other group of genes (*F*, *12S*, *V* ... *ND5*, *CytB* and *T*) were transcribed from HSP' (Fig. 1III-F). With the exception of *tRNA-N*, genes with identical transcriptional polarity were placed together but were separated by two CRs via the DMNL process. Finally, the gene *tRNA-D* was translocated from the site between *COI* and *COII* to

between *tRNA-S<sub>1</sub>* and *ND6* to form the 9-gene cluster "Q-A-C-Y-S<sub>1</sub>-D-ND6-E-P" (Fig. 1III-B).

### Gene-rearrangement mechanism of 13 bothid mitogenomes

In summary, three variants of the DMNL model can be used to account for the gene rearrangements seen in 11 bothid mitogenomes: Type I explains rearrangements seen in *C. azureus*, *Crossorhombus kobensis*, and *Crossorhombus valderostratus*, Type II explains those of *B. myriaster*, *Arnoglossus polyspilus*, and *Chascanopsetta lugubris*, and Type III describes the rearrangement mechanism used in the other five species (Fig. 1 Type I, II and III).

A fourth variant of the DMNL model can also be adopted to explain the rearrangement process that created the *Bothus pantherinus* and *Asterorhombus intermedius* mitogenomes (Fig. 1 Type IV). These two species appear to share the same rearrangement process (Fig. 1a, IV-c, d, and f) except for a difference in *tRNA-V* shuffling (Fig. 1IV-B1, B2); in *A. intermedius*, *tRNA-V* is moved from the location between *12S* and *16S* (*12S-V-16S-L<sub>1</sub>*) to between *16S* and *tRNA-L<sub>1</sub>* (*12S-16S-V-L<sub>1</sub>*). In Type IV, step A, C, and F of the DMNL processes are

		TAS (TAS-CTAS box)										
P.co	ATATTATAG	TAAACATTTT	ATACAGTGTG	TCGGGACATT	T-ATGTAATA	TCAGCTTATA	CAGTTTTTTA	AACAACAATG	CTCGGTAATT	CACACTAAGG	TAGACAAAAA	CCA-----ACG
G.p1	GATCCGAC	C.ATGA.A.A	G..GAC...T	..A..T.A.	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
A.t1	TTGTCTAT	GTTA.C.T.A	FRG.ATAG	CT-A-----	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
A.t2	-----	AC.TTAAAG	CT-A-----	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
L.g1	CTCC.CCACA	C.AA.A.A.G.	-GAG-S	A...T.A.	CA...C.T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
L.g2	CTCC.CCACA	C.AA.A.A.G.	-GAG-S	A...T.A.	CA...C.T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
L.l1	G.CAC.T	CA.AA.G	AT.AAAT	AC.A.C	CA.A.C	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
L.l2	G.CAC.T	CA.AA.G	AT.AAAT	AC.A.C	CA.A.C	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
P.l1	G.C.C.CACA	CCAAATGA	G-GAG-G	A...T.A.	CA...C.T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
P.l2	CT.CBTA-A	GTC.GCA	AG-G	AT.T.A	CA...C.T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
C.az	-----	CA.AA.C	GAG-G	A...T.A.	CA...C.T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
C.ko	TAT...GAT	C.GAAG.G	GGAATG	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
C.va	GGG.CC	CA.C.C.C	GC-GAG-S	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
A.po	CTT.CCCCTA	C-CAA.A.A.G	..-AA-AT	ATT.L.C	CT..T..T	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
B.my	CCC.ACCCT	TGGGGGGGG	GGT-----	GGGG	GGGG.AGTGG	T.TAA.AC	T..GCCGA	GG.CT---	TATAA..GC	ATT-GGTTT	ATTTT...C	ATT-----TA
C.lu	CCCCCCCG	AGTGGGGGG	GGCTCAC	CC.CCCG	C.C.GG	GGG.ACCCT	C.C.CCAGT	T.C.A.C	T.G.CT---	TATTA.C.CC	ATA---GT	A.TTT...C
A.in	TGCGTAT	C.C.C.G.A	CAATATG	CTATCA	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..	..-...T..
B.pa	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Fig. 4 Aligned sequences of control regions in 13 bothods and *Pleuronichthys cornutus*. The shaded blocks represent the conserved sequences. TAS is an acronym for the terminal associated sequence. CSB is an acronym for the conserved sequence block. Species names are abbreviated as in Fig. 3. The number 1 and 2 after the names indicate CR1 and CR2

similar to those in Type I, II and III; unique to step D of Type IV variant is the degeneration of the whole CR1 (Fig. 1IV-D), leaving only one CR in both the *B. pantherinus* and the *A. intermedius* mitogenomes (Fig. 1IV-B1 and B2).

## Discussion

### *tRNA-D* translocation

The *tRNA-D* gene is found in one of two sites in the 13 bothid mitogenomes studied here: outside (Fig. 1i-e) or inside (Fig. 1II-B, III-B, IV-B1 and IV-B2) the 8-gene cluster. In the Type I variant, translocation of *tRNA-D* to a locus outside of the cluster could occur either before (Fig. 1a, I-B) or after (Fig. 1II-F, I-E) the DMNL process, because either translocation times requires the same number of gene rearrangement steps. In variants Type II, III and IV, translocation of *tRNA-D* occurs to the same locus inside the cluster. Using the Type II variant as an example, if the *tRNA-D* gene was firstly translocated to a region between *ND5* and *ND6* (Fig. 1II-B'), then after the DMNL process (Fig. 1II-C', D') *tRNA-D* would be located outside the cluster between *ND5* and *CytB* (Fig. 1II-F'). To become located inside the cluster, this gene would require one more translocation step (Fig. 1II-B). In contrast, if translocation of *tRNA-D* occurred after the DMNL process (Fig. 1II-F), one translocation step would be sufficient to create the current gene structure. Therefore, it is more parsimonious to assume that *tRNA-D* translocation occurred after, rather than before, the DMNL process in variants Type II-IV.

### The fate of *tRNA-N* and *tRNA-N'* during DMNL

The *tRNA-N* gene is retained in mitogenome of *C. azureus* [14], and Shi et al. hypothesized that this gene acts as a functional  $O_L$  (Fig. 2a) because the intergenic spacer remaining between *tRNA-N* and *COI* is too small to form the necessary  $O_L$  secondary structure. In the mitogenomes of 12 other bothids, both shorter (12–14 bp) and longer (47–55 bp) intergenic spacers were found. As in *C. azureus* mitogenome, the shorter intergenic spacer in *A. intermedius*, *C. kobensis* and *C. valderostratus* (Fig. 2b-d) could also not form the  $O_L$  secondary structure directly, but the middle sequence of *tRNA-N* could form an  $O_L$ -like structure. The longer intergenic spacer in the other nine species (Fig. 2e-m), when including only 3–5 bp from the 3' end of *tRNA-N*, could also form the  $O_L$  structure, therefore, *tRNA-N* appears to assist in  $O_L$  function. This finding supports the correlation between the *tRNA-N* and the  $O_L$ , which lays a foundation for further studying mitochondrial gene rearrangement and replication.

### Evidence for the DMNL model

Compared with intergenic spacers in non-rearranged mitogenomes of four representative flatfishes, *Psettodes*

*erumei*, *Platichthys stellatus*, *Peltorhamphus novaezeelandiae* and *Pelotretis flavilatus*, six unique intergenic spacers (2–61 bp), either at unique locations or of longer length, were found in the *G. polyopthalmus* mitogenome (Fig. 3). The discovery of such intergenic spacer diversity led us to ask whether such spacers also occurred in the other 12 bothid mitogenomes. We found 121 unique spacers in 12 bothids at fourteen loci (number 1 to 12 plus repeated spacers labeled 6' and 10'); the length of 115 of these spacers ranged from 2 to 88 bp while the length of another six spacers (at the location numbered 12) ranged from 155 to 511 bp (Fig. 3b). What is the origin of these unique intergenic spacers? And what is the significance of these spacers in rearranged mitogenome? As we traced the DMNL process in each of the 13 mitogenomes, a striking finding was that each non-transcribed gene degenerated to a shorter intergenic spacer (Fig. 3a and b). This finding supports the DMNL model as the mechanism of mitochondrial gene rearrangement in these species, as well as supports the existing of dimeric mitogenome in mitochondrion.

Further analyses showed that the intergenic spacers are evolutionarily diverse. For example, in *C. valderostratus* and *A. tenuis*, the numbers of unique intergenic spacers are 5 and 11, respectively. This result indicates that the corresponding non-transcribed genes are progressively degenerating or have completely disappeared in these 13 bothid mitogenomes. The length of each spacer also varies in different species, for example, spacers numbered 2 and 5 range in length from 3 to 9 bp and 4–41 bp, respectively. This result suggests that each of non-transcribed genes degenerated at different rates in different species.

Seven intergenic regions appear to have no relationship to the DMNL process. Three intergenic spacers of the *12S-16S-V-L<sub>1</sub>* region were found only in *A. intermedius* and were generated by the shuffling of *tRNA-V*. The double-A spacer between *tRNA-H* and *tRNA-S<sub>2</sub>* shared the same base as seen at the 3' end of *tRNA-H*. The other three intergenic regions were also found in the non-rearranged mitogenomes of flatfishes and other teleosts.

## Conclusions

In summary, the newly completed mitogenome of *G. polyopthalmus* and the sequenced mitogenomes of 12 other bothids all possessed genomic-scale rearrangements. These rearrangements could be sorted into four types (Type I, II, III and IV), differing in the particular combination of CR, *tRNA-D* and 8-gene cluster and the shuffling of *tRNA-V*. The DMNL model can be used to account for all the gene rearrangements in all 13 bothid mitogenomes, except for the translocation of *tRNA-D* which appears to have occurred after the DMNL process in 10 of these mitogenomes, and either before or after in three others. During the DMNL process, *tRNA-N* was

retained rather than the expected *tRNA-N'* gene, and *tRNA-N* appears to assist in or act as  $O_L$  function when the  $O_L$  secondary structure could not be formed directly.

A striking finding was that each of the non-transcribed genes has degenerated to an intergenic spacer during the DMNL process. This result provided significant evidence to support the existence of dynamic dimeric mitogenomes and the DMNL model as the mechanism of gene rearrangement in bothid mitogenomes. These findings highlight a rare phenomenon in teleost fish, which not only promotes the understanding of mitogenome structural diversity, but also sheds light on mechanisms of mitochondrial genome rearrangement and replication.

## Methods

### Sampling, DNA extraction, PCR and sequencing

A single specimen of *G. polyopthalmus* was collected from fishing activity in Fangcheng Port, China, conducted by artisanal bottom-trawling, therefore no specific information is available on depth of capture, substrata, or longitude and latitude of the catch. Fresh specimen was stored in crushed ice immediately after collection, and then frozen at  $-20\text{ }^\circ\text{C}$  in the lab until further processing. Total genomic DNA was extracted from the epaxial musculature of the specimen using the SQ Tissue DNA Kit (OMEGA) following the manufacturer's protocol. Nine primer pairs were designed for amplification of the mtDNA genome (Additional file 4: Table S2), based on the known complete mitochondrial sequences of bothids. More than 70 bp of overlaps between sequenced fragments ensured correct assembly of the complete mtDNA genome.

PCR was performed in a 25  $\mu\text{l}$  reaction volume containing approximate 50 ng DNA template, 0.5  $\mu\text{M}$  of each primer, 1.0 U Taq polymerase (Takara, China), 2.0 mM  $\text{MgCl}_2$ , 0.4 mM of each dNTP, and 2.5  $\mu\text{l}$  of 10 $\times$  Taq buffer. The PCR cycling protocol included one cycle at  $95\text{ }^\circ\text{C}$  for 3 min (initial denaturation), followed by 35 cycles at  $95\text{ }^\circ\text{C}$  for 30s (denaturation),  $45\text{--}55\text{ }^\circ\text{C}$  for 50s (annealing) and  $68\text{--}72\text{ }^\circ\text{C}$  for 1.5–5 min (extension), followed by a final extension at  $68\text{--}72\text{ }^\circ\text{C}$  for 10 min. The PCR products were purified by the Takara Agarose Gel DNA Purification Kit (Takara, China) and then were used directly as templates for cycle sequencing reactions. Sequence specific primers were designed for walking across each fragment with an ABI 3730 DNA sequencer (Applied Biosystems, USA). Sequenced fragments were assembled to a complete mitochondrial genome using CodonCode Aligner v3 and BioEdit v7 [24]. For large fragments and walking sequences, manual examinations ensured reliable assembly of the mitogenome. The complete sequence of *G. polyopthalmus* mitogenome was submitted to GenBank under the accession number MK770643.

### Sequence analysis

Identification and annotation of protein coding genes and rRNA genes were performed using NCBI-BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The tRNA genes and their secondary structures were determined using tRNAscan-SE 1.21 [25], setting the cut-off values to 1 when necessary. The secondary structure of  $O_L$  was identified by using the mfold web server (<http://unafold.rna.albany.edu/?q=mfold>). The gene map of *G. polyopthalmus* mitogenome was generated by using CGView [26]. Mitogenomes of eight out of 12 bothid fishes used in this study were determined in our lab, including *A. intermedius*, *A. tenuis*, *B. myriaster*, *C. azureus*, *C. lugubris*, *C. valderostratus*, *L. gallus*, and *P. ijijimae*, available under the accession numbers MK256952, KP134337, KJ433563, JQ639068, KJ433561, KJ433566, KJ433567 and KP134336, respectively. Mitogenome sequences of four other bothids were retrieved from GenBank, including *A. polyspilus*, *B. pantherinus*, *C. kobensis*, and *L. lanceolata*, using accession numbers NC\_024946, NC\_024947, NC\_024949, and AP014591, respectively.

### Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-6128-9>.

**Additional file 1: Table S1.** Organization of the *G. polyopthalmus* mitogenome. (DOCX 20 kb)

**Additional file 2: Figure S1.** Gene map of the *G. polyopthalmus* mitogenome.

**Additional file 3: Figure S2.** Aligned sequences of complete control region in 13 bothids and *Pleuromichthys cornutus*.

**Additional file 4: Table S2.** Primers used for fragment amplification of the *G. polyopthalmus* mitogenome.

### Abbreviations

3-gene cluster: Three genes (*ND6-E-P*) on L-strand grouped to a cluster; 5-gene cluster: Five genes (*Q-A-C-Y-S<sub>1</sub>*) on L-strand grouped to a cluster; 8-gene cluster: Eight genes (*Q-A-C-Y-S<sub>1</sub>-ND6-E-P*) on L-strand grouped to a cluster; CR: Control region; DMNL: Dimer-Mitogenome and Non-Random Loss mechanism of mitochondrial rearrangement; HSP: H-strand promoters; H-strand: Heavy strand; LSP: L-strand promoters; L-strand: Light strand; Mitogenome: Mitochondrial genome; mtDNA: Mitochondrial DNA; NC: Non-coding region;  $O_H$ : Origin for H-strand replication;  $O_L$ : Origin for L-strand replication; rRNA: Ribosomal ribonucleic acid; TAS: Terminal associated sequences; CSB: Conserved sequence blocks; tRNA: Transfer ribonucleic acid; WANCY: tRNA cluster of *tRNA-W*, *tRNA-A*, *tRNA-N*, *tRNA-C* and *tRNA-Y*

### Acknowledgments

The authors are very grateful to Prof. Elizabeth De Stasio for valuable comments, and also thank Zhuocheng Zhou and Xiaolu Yu for providing assistance in sample collection.

### Authors' contributions

HRL sampled specimen, collected the datasets and drafted the manuscript. XYK directed the entire research project and revised the manuscript. SXC carried out the experiments. WS revised the analysis of datasets. All authors read and approved the final manuscript.

### Funding

This work was supported by the National Nature Science Foundation of China (grant numbers 31872570 and 31471979). The funding body plays no



role in the design of the study, collection, analysis and interpretation of data or in writing the manuscript.

#### Availability of data and materials

The sequence of complete mtDNA of *G. polyopthalmus* in this study was submitted to GenBank under the accession number MK770643. Mitogenomes of other 12 bothids fishes (*A. intermedius*, *A. polyspilus*, *A. tenuis*, *B. myriaster*, *B. pantherinus*, *C. azureus*, *C. kobensis*, *C. lugubris*, *C. valderostratus*, *L. gallus*, *L. lanceolata* and *P. iijimae*) were under the GenBank accession numbers MK256952, NC\_024946, KP134337, KJ433563, NC\_024947, JQ639068, NC\_024949, KJ433561, KJ433566, KJ433567, AP014591 and KP134336, respectively.

#### Ethics approval and consent to participate

The specimen of *G. polyopthalmus* was collected from the commercial fishery. This species was not involved in the International Union for Conservation of Nature Red List of Threatened Species. Specimen collection and maintenance were performed in strict accordance with both of the recommendations of Animal Care Quality Assurance in Chinese Mainland and the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, South China Sea Institute of Oceanology, Guangzhou 510301, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China.

Received: 24 August 2019 Accepted: 23 September 2019

Published online: 30 October 2019

#### References

- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290(5806):457.
- Clayton DA. Replication and transcription of vertebrate mitochondrial DNA. *Annu Rev Cell Biol*. 1991;7(1):453–78.
- Fernández-Silva P, Enriquez JA, Montoya J. Replication and transcription of mammalian mitochondrial DNA. *Exp Physiol*. 2003;88(1):41–56.
- Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Mol Biol Evol*. 1997;14(1):91–104.
- Thyagarajan B, Padua RA, Campbell C. Mammalian mitochondria possess homologous DNA recombination activity. *J Biol Chem*. 1996;271(44):27536.
- Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E. Widespread recombination in published animal mtDNA sequences. *Mol Biol Evol*. 2005;22(4):925.
- Zhuang X, Cheng CH. ND6 gene “lost” and found: evolution of mitochondrial gene rearrangement in Antarctic notothenioids. *Mol Biol Evol*. 2010;27(6):1391.
- Kong XY, Dong XL, Zhang YC, Shi W, Wang ZM, Yu ZN. A novel rearrangement in the mitochondrial genome of tongue sole, *Cynoglossus semilaevis*: control region translocation and a tRNA gene inversion. *Genome*. 2009;52(12):975–84.
- Miya M, Nishida M. Organization of the Mitochondrial Genome of a Deep-Sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. *Mar Biotechnol*. 1999;1(5):416–26.
- Ki JS, Jung SO, Hwang DS, Lee YM, Lee JS. Unusual mitochondrial genome structure of the freshwater goby *Odontobutis platycephala*: rearrangement of tRNAs and an additional non-coding region. *J Fish Biol*. 2008;73(2):414–28.
- Satoh TP, Miya M, Endo H, Nishida M. Round and pointed-head grenadier fishes (Actinopterygii: Gadiformes) represent a single sister group: evidence from the complete mitochondrial genome sequences. *Mol Phylogenet Evol*. 2006;40(1):129–38.
- Gong L, Shi W, Yang M, Li DH, Kong XY. Novel gene arrangement in the mitochondrial genome of *Bothus myriaster* (Pleuronectiformes: Bothidae): evidence for the dimer-Mitogenome and non-random loss model. *Mitochondrial DNA Part A*. 2016;27(5):3089–92.
- Dong JX, Shi W, Kong XY, Chen SX. Characteristic analysis of the complete mitogenome of *Plagiopsetta glossa* and a possible mechanism for gene rearrangement. *J Trop Oceanography*. 2018;37(1):1–11 (in Chinese).
- Shi W, Dong XL, Wang ZM, Miao XG, Wang SY, Kong XY. Complete mitogenome sequences of four flatfishes (Pleuronectiformes) reveal a novel gene arrangement of L-strand coding genes. *BMC Evol Biol*. 2013;13(1):173.
- Shi W, Gong L, Wang SY, Miao XG, Kong XY. Tandem duplication and random loss for mitogenome rearrangement in *Symphurus* (teleost: Pleuronectiformes). *BMC Genomics*. 2015;16(1):1–9.
- Shi W, Miao XG, Kong XY. A novel model of double replications and random loss accounts for rearrangements in the Mitogenome of *Samariscus latus* (Teleostei: Pleuronectiformes). *BMC Genomics*. 2014;15(1):352.
- Moritz C, Dowling TE, Brown WM. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu Rev Ecol Syst*. 1987;18(1):269–92.
- Cantatore P, Gadaleta MN, Roberti M, Saccone C, Wilson AC. Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. *Nature*. 1987;329(6142):853–5.
- Poulton J, Deadman ME, Bindoff L, Morten K, Land J, Brown G. Families of mtDNA re-arrangements can be detected in patients with mtDNA deletions: duplications may be a transient intermediate form. *Hum Mol Genet*. 1993;2(1):23–30.
- Lavrov DV, Boore JL, Brown WM. Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Mol Biol Evol*. 2002;19(2):163–9.
- Seligmann H, Krishnan NM, Rao BJ. Possible multiple origins of replication in primate mitochondria: alternative role of tRNA sequences. *J Theor Biol*. 2006;241(2):321–32.
- Hurst CD, Bartlett SE, Davidson WS, Bruce IJ. The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar*. *Gene*. 1999;239(2):237–42.
- Wang SY, Shi W, Miao XG, Kong XY. Complete mitochondrial genome sequences of three rhombosoleid fishes and comparative analyses with other flatfishes (Pleuronectiformes). *Zool Stud*. 2014;53(1):80.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. In: Hall T, Hall TA, Hall D, editors. *Nucleic acids symposium series*: 1999. London: Information Retrieval Ltd.; 1999. p. 95–8.
- Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25(5):955–64.
- Stothard P, Wishart D. Circular genome visualization and exploration using CGView. *Bioinformatics*. 2005;21(4):537–9.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

