SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-019-9528-1

Evidence for the paternal mitochondrial DNA in the crucian carp-like fish lineage with hybrid origin

Shi Wang[1,2,3], Ni Jiao[1,2], Lu Zhao[1,2], Meiwen Zhang[1,2], Pei Zhou[1,2], Xuexue Huang[1,2], Fangzhou Hu[1,2,3], Conghui Yang[1,2], Yuqin Shu[1,2], Wuhui Li[1,2,4], Chun Zhang[1,2], Min Tao[1,2], Bo Chen[3], Ming Ma[3] & Shaojun Liu[1,2*]

[1]*State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha 410081, China;*
[2]*College of Life Sciences, Hunan Normal University, Changsha 410081, China;*
[3]*College of Chemistry and Chemical Engineering, Hunan Normal University, Changsha 410081, China;*
[4]*Key Laboratory of Tropical and Subtropical Fisheries Resource Application and Cultivation, Ministry of Agriculture, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China*

Received June 11, 2019; accepted August 11, 2019; published online November 13, 2019

In terms of taxonomic status, common carp (*Cyprinus carpio*, Cyprininae) and crucian carp (*Carassius auratus*, Cyprininae) are different species; however, in this study, a newborn homodiploid crucian carp-like fish (2*n*=100) (2*n*NCRC) lineage (F_1–F_3) was established from the interspecific hybridization of female common carp (2*n*=100)×male blunt snout bream (*Megalobrama amblycephala*, Cultrinae, 2*n*=48). The phenotypes and genotypes of 2*n*NCRC differed from those of its parents but were closely related to those of the existing diploid crucian carp. We further sequenced the whole mitochondrial (mt) genomes of the 2*n*NCRC lineage from F_1 to F_3. The paternal mtDNA fragments were stably embedded in the mt-genomes of F_1–F_3 generations of 2*n*NCRC to form chimeric DNA fragments. Along with this chimeric process, numerous base sites of F_1–F_3 generations of 2*n*NCRC underwent mutations. Most of these mutation sites were consistent with the existing diploid crucian carp. Moreover, the mtDNA organization and nucleotide composition of 2*n*NCRC were more similar to those of the existing diploid crucian carp than those of the parents. The inheritable chimeric DNA fragments and mutant loci in the mt-genomes of different generations of 2*n*NCRC provided important evidence of the mtDNA change process in the newborn lineage derived from hybridization of different species. Our findings demonstrated for the first time that the paternal mtDNA were transmitted into the mt-genomes of homodiploid lineage, which provided new insights into the existence of paternal mtDNA in the mtDNA inheritance.

hybridization, homodiploid, common carp, blunt snout bream, crucian carp, mitochondrial genome, genetic evolution

Citation: Wang, S., Jiao, N., Zhao, L., Zhang, M., Zhou, P., Huang, X., Hu, F., Yang, C., Shu, Y., Li, W., et al. (2019). Evidence for the paternal mitochondrial DNA in the crucian carp-like fish lineage with hybrid origin. Sci China Life Sci 62, https://doi.org/10.1007/s11427-019-9528-1

INTRODUCTION

Hybridization in plants and animals is more common and has more complex outcomes than previously realized (Baack and Rieseberg, 2007). Biologists generally believe that hybridization could be a catalyst for not only speciation but also major evolutionary innovations (Mallet, 2007). This process may accelerate speciation via adaptive introgression or cause near-instantaneous speciation (Abbott et al., 2013). This near-instantaneous hybrid speciation is accompanied by rapid genomic changes, including chromosomal rearrangements, genome expansion, differential gene expression, and gene silencing. These dramatic genomic alterations may lead to new beneficial phenotypes and to selection for fertility and ecological traits (Baack and Rieseberg, 2007).

Common carp (*Cyprinus carpio*, Cyprininae, ♀, 2*n*=100,

*Corresponding author (email: lsj@hunnu.edu.cn)

abbreviated as 2*n*COC) and blunt snout bream (*Megalobrama amblycephala*, Cultrinae, ♂, 2*n*=48, abbreviated as 2*n*BSB) belong to different subfamilies, and hybridization between these species is considered to be distant. Common carp and crucian carp (diploid wild crucian carp, *Carassius auratus*, 2*n*=100, abbreviated as 2*n*CC) belong to the same subfamily (Cyprininae) but are classified into different genera (*Cyprinus* and *Carassius*, respectively). In terms of taxonomic status, common carp and crucian carp are the most closely related species in the Cyprininae subfamily. However, for a long time, the evolutionary relationship between these species has remained unclear (Gui and Zhou, 2010; Guo and Gui, 2008; Li and Gui, 2018; Mei and Gui, 2015; Wang et al., 2019; Wang et al., 2017; Zhang and Gui, 2018). In our previous study, we reported the spontaneous occurrence of a newborn homodiploid crucian carp-like fish (2*n*=100, abbreviated as 2*n*NCRC) that originated from 2*n*COC (♀)×2*n*BSB (♂) (Figures 1 and 2). The phenotypes and genotypes (fluorescence *in situ* hybridization and 5S rDNA) of 2*n*NCRC differed from those of its parents but were closely related to those of the existing 2*n*CC (Wang et al., 2017). To further explore the evolutionary relationship between 2*n*COC and 2*n*CC, we studied the mitochondrial (mt) DNA structures and mt-genomes of the different generations of 2*n*NCRC lineage. Determination of the genotype of the mt-genome of this lineage is very useful for understanding the process associated with the change in mtDNA accompanying changes in phenotypes and genomic DNA.

Most animal mt-genomes contain 37 genes, including 13 protein-coding genes, 2 ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs), which are necessary for translation of the proteins encoded by the mtDNA (Boore, 1999; Yue et al., 2006). High sequence divergence between species and maternal inheritance characteristics make mtDNA useful in tracing animal lineages (Guo et al., 2006). Recombination of mtDNA occurs commonly in most plants, fungal and protist species but was traditionally thought to be rare or absent in animals (Gillham, 1994; Rokas et al., 2003). In general, animal mtDNA is characterized by maternal inheritance. However, some studies have shown the occurrence of mtDNA recombination and paternal leakage in the following animal species: the nematode *Meloidogyne javanica* (Lunt and Hyman, 1997), the mussel sister-species *Mytilus galloprovincialis* (Ladoukakis and Zouros, 2001) and *Mytilus trossulus* (Burzyński et al., 2003), the flatfish *Platichthys flesus* (Hoarau et al., 2002), *Drosophila* (Dokianakis and Ladoukakis, 2014; Nunes et al., 2013), mouse (Gyllensten et al., 1991), sheep (Zhao et al., 2004), humans (Kraytsberg et al., 2004; Luo et al., 2018), and triploid crucian carp (Guo et al., 2006; Liu, 2014). However, erroneous mt sequences have been identified in most previous genetic studies as a result of sample contamination and are typically ignored. The consequences of ignoring paternal leakage and mtDNA recombination include species misidentification and failure to detect cryptic species (Morgan et al., 2013).

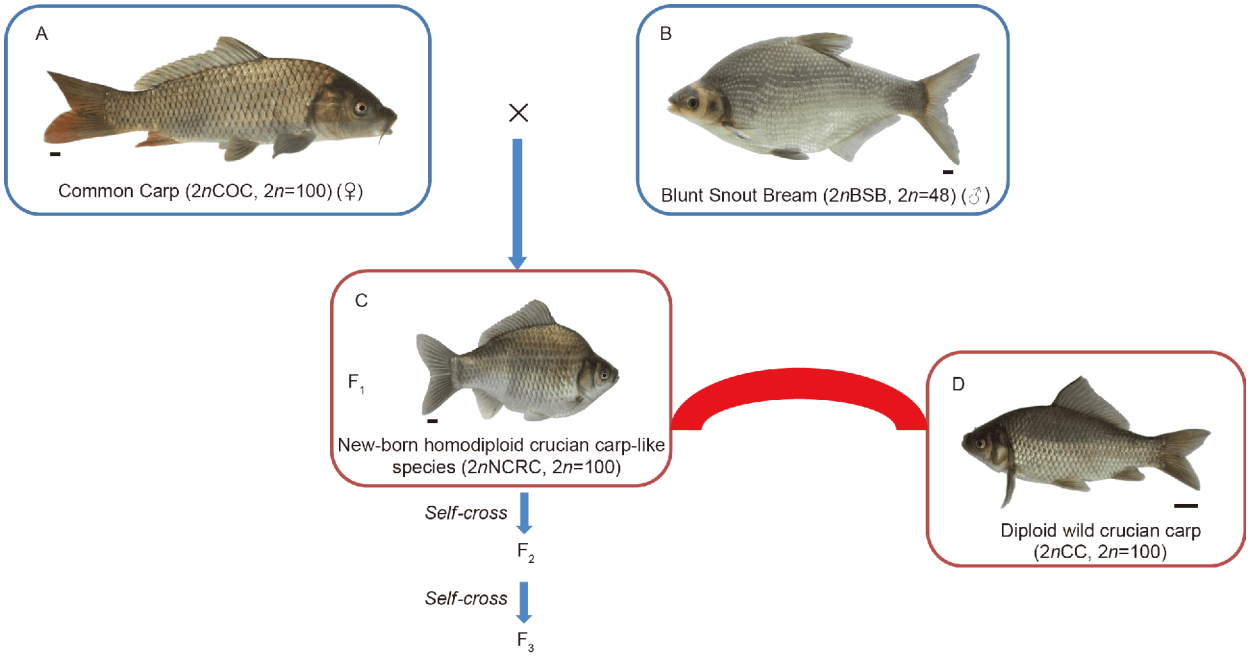

Figure 1 Crossing procedure and appearance of 2*n*COC, 2*n*BSB, 2*n*NCRC, and 2*n*CC. A, 2*n*COC. B, 2*n*BSB. C, 2*n*NCRC (F_1). D, 2*n*CC. In the first generation of common carp (*Cyprinus carpio*, 2*n*COC, ♀, Cyprininae, 2*n*=100)×blunt snout bream (*Megalobrama amblycephala*, 2*n*BSB, ♂, Cultrinae, 2*n*=48), the newborn homodiploid crucian carp-like fish (2*n*=100, 2*n*NCRC) were produced. The newborn homodiploid crucian carp-like fish of F_2 (2*n*=100, 2*n*NCRC-F_2) were produced as a second generation of 2*n*NCRC by self-crossing. The newborn homodiploid crucian carp-like fish of F_3 (2*n*=100, 2*n*NCRC-F_3) was produced as the third generation of 2*n*NCRC-F_2 by self-crossing.

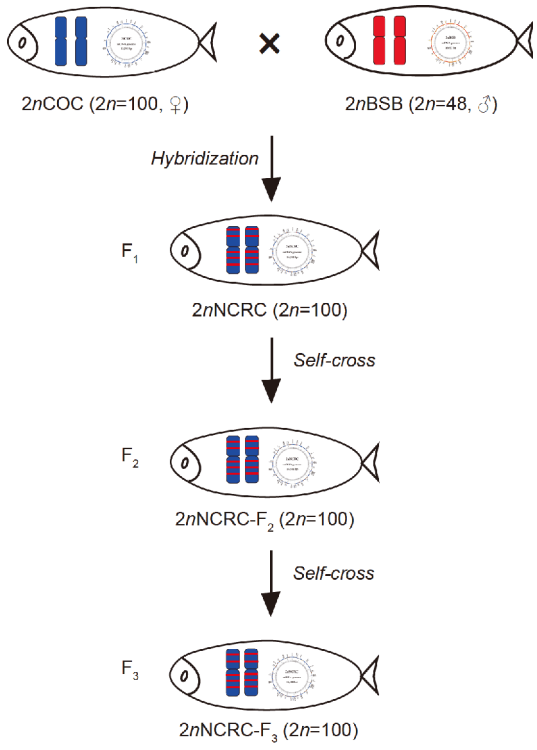

Figure 2 Pattern diagrams for the crossing procedure and the formation of different generations of 2*n*NCRC. The chromosomes (a pair of linear bars) or mitochondrial structure (circular structure) of common carp (2*n*COC) and blunt snout bream (2*n*BSB) are marked in blue and red, respectively; yellow represents mutations that differ from those in both 2*n*COC and 2*n*BSB.

RESULTS

Genome organization and gene orders

The whole-mt-genome sequences of 2*n*NCRC, 2*n*NCRC-F_2, and 2*n*NCRC-F_3 were submitted to GenBank under accession numbers MH599071, MK085972, and MK085973, respectively. The length of the whole-mt-genomes of 2*n*NCRC, 2*n*NCRC-F_2, and 2*n*NCRC-F_3 was 16,580 base pairs (bp), which was the same as that of the female parent 2*n*COC but different from that of the male parent BSB (16,623 bp). The total mtDNA identity (98.60%) between 2*n*NCRC and 2*n*NCRC-F_2 was lower than that (99.90%) between 2*n*NCRC-F_2 and 2*n*NCRC-F_3. The complete mt-genome sequence similarity was high among different generations of 2*n*NCRC, so the following analysis of mt-genome organization and gene orders is presented for only the first generation of 2*n*NCRC (abbreviated as 2*n*NCRC). The arrangement of all the mtDNA genes of 2*n*NCRC was identical to that of the parents 2*n*COC and 2*n*BSB (Figure 3, Table 1). In 2*n*NCRC, 2*n*COC, and 2*n*BSB, all the protein-coding genes were identical in length except for NADH 5; there were differences in the lengths of the two rRNA genes (*12S rRNA* and *16S rRNA*); and most of the tRNA genes (17/22) were of the same length (Table 1). The D-loop regions (CRs) in 2*n*NCRC, 2*n*COC, and 2*n*BSB were located between the tRNA-Pro and tRNA-Phe and were 923, 927, and 937 bp in length, respectively.

The reading frames of two pairs of genes, namely ATPase 8-ATPase 6 and NADH 4L-NADH 4, each overlapping by seven nucleotides, and one pair of genes, namely NADH 5-NADH 6, overlapping by four nucleotides, appeared to be common in most vertebrate mt-genomes, and the sizes of these reading frames in fish range from 7–10 bp (Broughton et al., 2001). A number of other genes share one, two, or three nucleotides in common with adjacent tRNA genes (Table 1). Similarly, 12–14 intervals between contiguous genes were observed in 2*n*NCRC, 2*n*COC, and 2*n*BSB (Table 1).

As shown in Figure 3 and Table 1, the analysis results showed that the mtDNA organization of 2*n*NCRC was more similar to that of 2*n*CC than to that of the parents. However, there were many similarities in the mtDNA organization of 2*n*NCRC and 2*n*COC, except in the control region, two rRNA genes, and four tRNA genes. These structural differences were further narrowed by minor variations in the intergenic regions and tRNA genes, so the mtDNA organization of 2*n*NCRC and 2*n*COC was roughly the same. Our results indicated that the mtDNA structures of 2*n*NCRC and the maternal parent 2*n*COC were roughly the same, although there were some minor differences between these structures.

Nucleotide and codon composition analysis of the protein-coding genes

Nucleotide composition can accurately reflect the basic characteristics of genetic variations among mtDNA sequences. As shown in Figure 2, the mtDNA nucleotide composition of 2*n*NCRC was more similar to that of 2*n*CC than to that of the parents. Compared to the parents, the T in the overall nucleotide composition of the 13 protein-coding genes of 2*n*NCRC was mainly derived from C in 2*n*COC; especially at the third codon positions, where this effect was obvious (Table 2). The mt protein-coding gene sequences of 2*n*NCRC had 1,401 sites that were mutated relative to the female parent 2*n*COC. Most of these mutations (1,161, 82.87%) occurred at the third codon position. Table 3 shows the type and proportion of mutations at the third codon position of the mt protein-coding gene sequence of 2*n*NCRC (with 2*n*COC as a reference sequence). Among these mutations, 93.11% were synonymous mutations, and only 6.89% were non-synonymous mutations. Among the types of mutations at the third codon position, 53.49% were mutations from C to T or T to C. These two base substitutions correspond exactly to the methylation and demethylation of cytosine; cytosine methylation is a universal epigenetic phenomenon (Bird, 1992; Richards and Elgin, 2002). To a

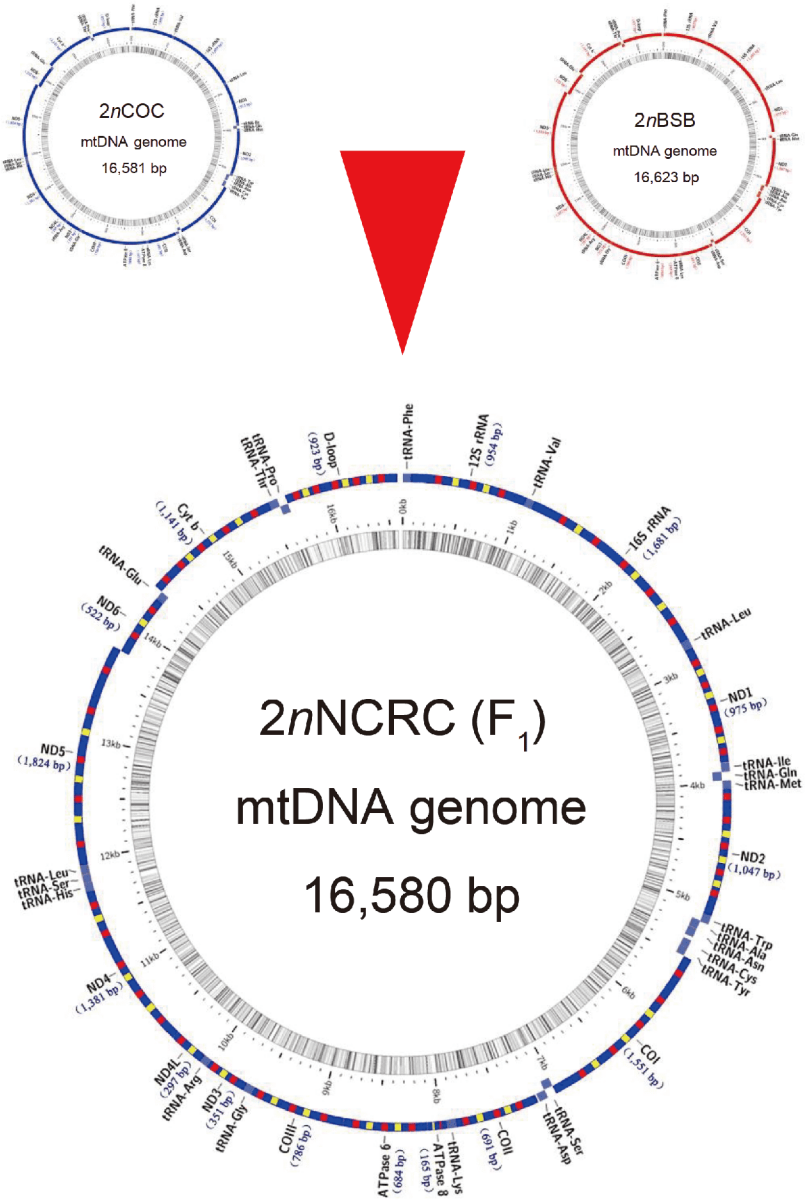

Figure 3 Mitochondrial structures of 2*n*NCRC (F_1) and its parents. The mitochondrial structures (circular structure) of common carp (2*n*COC) and blunt snout bream (2*n*BSB) are marked in blue and red, respectively; yellow represents mutations that differ from those in both 2*n*COC and 2*n*BSB.

certain extent, epigenetic changes contribute to the survival and reproduction of nascent species and may play an important role in the evolution of species. In addition, 12.14% of the mutations were from A to G, and 10.59% of the mutations were from G to A.

The codon usage (number of codons) bias among 2*n*NCRC, 2*n*COC, 2*n*BSB, 2*n*CC, and zebrafish is shown in Table 4. The analysis results showed that the mtDNA codon usage of 2*n*NCRC was more similar to that of 2*n*CC than to that of the parents. Excluding stop codons, there were 3,794 codons in both 2*n*NCRC and its female parent 2*n*COC, and its male parent 2*n*BSB had 3,798 codons. For amino acids with fourfold degenerate third position, codons ending in A were consistently the most frequent in 2*n*NCRC, 2*n*COC, 2*n*BSB, 2*n*CC, and zebrafish. Codons ending in A were followed in frequency by codons ending in T or C. Among twofold degenerate codons, C appeared to be used somewhat more frequently than T. Consistent with the overall bias against G, G was the least common third-position nucleotide in all the codons except for those encoding glycine (Broughton et al., 2001). Notably, 2*n*NCRC had almost the same number of codons for the same amino acid as its parents, but the replacement of C with T was more frequent in 2*n*NCRC, and these substitutions were almost synonymous mutations; therefore, the encoded amino acid did not change (Tables 2–4). Cytosine methylation is a universal epigenetic

Table 1 Organization of the 2*n*NCRC, 2*n*COC, 2*n*BSB, and 2*n*CC mitochondrial genomes

Name of genes	Location				Size (bp) 2*n*NCRC/ 2*n*COC/2*n*BSB/ 2*n*CC	Start codons 2*n*NCRC/ 2*n*COC/2*n*BSB/ 2*n*CC	Stop codons 2*n*NCRC/ 2*n*COC/2*n*BSB/ 2*n*CC	Intergenic (bp) 2*n*NCRC/ 2*n*COC/2*n*BSB/ 2*n*CC
	2*n*NCRC	2*n*COC	2*n*BSB	2*n*CC				
The full length	16,580 bp	16,581 bp	16,623 bp	16,580 bp				
Control region	1–923	1–927	1–937	1–923	923/927/937/923			0
tRNA-Phe	924–992	928–996	938–1,006	924–992	69/69/69/69			0
12S rRNA	993–1,946	997–1,951	1,007–1,968	993–1,946	954/955/962/954			0
tRNA-Val	1,947–2,018	1,952–2,023	1,969–2,040	1,947–2,018	72/72/72/72			0
16S rRNA	2,019–3,699	2,024–3,702	2,041–3,732	2,019–3,699	1,681/1,679/ 1,692/1,681			0/0/+1/0
tRNA-Leu	3,700–3,777	3,703–3,778	3,734–3,809	3,700–3,777	78/76/76/78			0/+1/+1/0
NADH 1	3,778–4,752	3,780–4,754	3,811–4,785	3,778–4,752	975/975/975/975	ATG/ATG/ATG/ ATG	TAA/TAA/TAA/ TAA	+3/+4/+4/+3
tRNA-Ile	4,756–4,829	4,759–4,830	4,790–4,861	4,756–4,829	74/72/72/74			–3/–2/–2/–3
tRNA-Gln (L)	4,827–4,897	4,829–4,899	4,860–4,930	4,827–4,897	71/71/71/71			+1/+2/+1/+1
tRNA-Met	4,899–4,967	4,902–4,970	4,932–5,000	4,899–4,967	69/69/69/69			0
NADH 2	4,968–6,014	4,971–6,017	5,001–6,047	4,968–6,014	1,047/1,047/ 1,047/1,047	ATG/ATG/ATG/ ATG	TAG/TAG/TAG/ TAG	–2/–2/–2/–2
tRNA-Trp	6,013–6,083	6,016–6,086	6,046–6,116	6,013–6,083	71/71/71/71			+2/+2/+1/+2
tRNA-Ala (L)	6,086–6,154	6,089–6,157	6,118–6,186	6,086–6,154	69/69/69/69			+1/+1/+1/+1
tRNA-Asn (L)	6,156–6,228	6,159–6,231	6,188–6,260	6,156–6,228	73/73/73/73			+32/+33/+32/+32
tRNA-Cys (L)	6,261–6,331	6,265–6,331	6,293–6,361	6,261–6,331	71/67/69/71			–2/–1/+1/–2
tRNA-Tyr (L)	6,330–6,400	6,331–6,401	6,363–6,433	6,330–6,400	71/71/71/71			+1/+1/+1/+1
COI	6,402–7,952	6,403–7,953	6,435–7,985	6,402–7,952	1,551/1,551/ 1,551/1,551	GTG/GTG/GTG/ GTG	TAA/TAA/TAA/ TAA	0
tRNA-Ser (L)	7,953–8,023	7,954–8,024	7,986–8,056	7,953–8,023	71/71/71/71			+3/+3/+2/+3
tRNA-Asp	8,027–8,098	8,028–8,099	8,059–8,132	8,027–8,098	72/72/74/72			+12/+13/+13/+12
COII	8,111–8,801	8,113–8,803	8,146–8,836	8,111–8,801	691/691/691/691	ATG/ATG/ATG/ ATG	T–/T–/T–/T–	0
tRNA-Lys	8,802–8,877	8,804–8,879	8,837–8,912	8,802–8,877	76/76/76/76			+1/+1/+1/+1
ATPase8	8,879–9,043	8,881–9,045	8,914–9,078	8,879–9,043	165/165/165/165	ATG/ATG/ATG/ ATG	TAG/TAG/TAA/ TAG	–7/–7/–7/–7
ATPase6	9,037–9,720	9,039–9,722	9,072–9,755	9,037–9,720	684/684/684/684	ATG/ATG/ATG/ ATG	TAA/TAA/TAA/ TAA	–1/–1/–1/–1
COIII	9,720–10,505	9,722–10,507	9,755–10,540	9,720–10,505	786/786/786/786	ATG/ATG/ATG/ ATG	TAA/TAA/TAA/ TAA	–1/–1/–1/–1
tRNA-Gly	10,505–10,576	10,507–10,578	10,540–10,611	10,505–10,576	72/72/72/72			0
NADH 3	10,577–10,927	10,579–10,929	10,612–10,962	10,577–10,927	351/351/351/351	ATG/ATG/ATG/ ATG	TAG/TAG/TAG/ TAG	–2/–2/–2/–2
tRNA-Arg	10,926–10,995	10,928–10,997	10,961–11,030	10,926–10,995	70/70/70/70			0
NADH 4L	10,996–11,292	10,998–11,294	11,031–11,327	10,996–11,292	297/297/297/297	ATG/ATG/ATG/ ATG	TAA/TAA/TAA/ TAA	–7/–7/–7/–7
NADH 4	11,286–12,666	11,288–12,668	11,321–12,702	11,286–12,666	1381/1381/1382/ 1381	ATG/ATG/ATG/ ATG	T–/T–/TA-/T–	0
tRNA-His	12,667–12,735	12,669–12,737	12,703–12,771	12,667–12,735	69/69/69/69			0
tRNA-Ser	12,736–12,804	12,738–12,806	12,772–12,840	12,736–12,804	69/69/69/69			+1/+1/+1/+1
tRNA-Leu	12,806–12,878	12,808–12,880	12,842–12,914	12,806–12,878	73/73/73/73			+3/+3/0/+3
NADH 5	12,882–14,705	12,884–14,707	12,915–14,750	12,882–14,705	1824/1824/1836/ 1824	ATG/ATG/ATG/ ATG	TAA/TAA/TAA/ TAA	–4/–4/–4/–4
NADH 6 (L)	14,702–15,223	14,704–15,225	14,747–15,268	14,702–15,223	522/522/522/522	ATG/ATG/ATG/ ATG	TAG/TAA/TAA/ TAG	0
tRNA-Glu (L)	15,224–15,292	15,226–15,294	15,269–15,337	15,224–15,292	69/69/69/69			+5/+5/+4/+5
Cytb	15,298–16,438	15,300–16,440	15,342–16,482	15,298–16,438	1141/1141/1141/ 1141	ATG/ATG/ATG/ ATG	T–/T–/T–/T–	0
tRNA-Thr	16,439–16,510	16,441–16,512	16,483–16,554	16,439–16,510	72/72/72/72			–2/–1/–1/–2
tRNA-Pro (L)	16,509–16,580	16,512–16,581	16,554–16,623	16,509–16,580	72/70/70/72			

Table 2 Nucleotide frequencies and proportions in 2*n*NCRC, 2*n*COC, 2*n*BSB, 2*n*CC, and zebrafish protein-coding genes

		2*n*COC	2*n*NCRC	2*n*BSB	2*n*CC	Zebrafish
Overall Codons	A	3,412 (29.9%)	3,407 (29.9%)	3,325 (29.1%)	3,389 (29.7%)	3,399 (29.8%)
	C	3,228 (28.3%)	3,062 (26.8%)	3,271 (28.6%)	3,055 (26.8%)	2,717 (23.8%)
	G	1,731 (15.2%)	1,741 (15.3%)	1,782 (15.6%)	1,747 (15.3%)	1,765 (15.5%)
	T	3,041 (26.6%)	3,202 (28.1%)	3,046 (26.7%)	3,221 (28.2%)	3,522 (30.9%)
Position 1 Codons	A	1,021 (26.8%)	1,016 (26.7%)	1,027 (27.0%)	1,014 (26.7%)	1,053 (27.7%)
	C	1,007 (26.5%)	985 (25.9%)	1,006 (26.4%)	979 (25.7%)	871 (22.9%)
	G	987 (25.9%)	987 (25.9%)	987 (25.9%)	988 (26.0%)	971 (25.6%)
	T	789 (20.7%)	816 (21.5%)	788 (20.7%)	823 (21.6%)	906 (23.8%)
Position 2 Codons	A	704 (18.5%)	708 (18.6%)	703 (18.5%)	705 (18.5%)	713 (18.8%)
	C	1,035 (27.2%)	1,034 (27.2%)	1,043 (27.4%)	1,032 (27.1%)	1,020 (26.8%)
	G	522 (13.7%)	522 (13.7%)	521 (13.7%)	521 (13.7%)	510 (13.4%)
	T	1,543 (40.6%)	1,540 (40.5%)	1,541 (40.5%)	1,546 (40.6%)	1,558 (41.0%)
Position 3 Codons	A	1,687 (44.3%)	1,683 (44.2%)	1,595 (41.9%)	1,670 (43.9%)	1,633 (43.0%)
	C	1,186 (31.2%)	1,043 (27.4%)	1,222 (32.1%)	1,044 (27.4%)	826 (21.7%)
	G	222 (5.8%)	232 (6.1%)	274 (7.2%)	238 (6.3%)	284 (7.5%)
	T	709 (18.6%)	846 (22.2%)	717 (18.8%)	852 (22.4%)	1,058 (27.8%)

Table 3 Analysis of the mutations of the third codon positions in the mitochondrial protein-coding gene sequences of 2*n*NCRC (with 2*n*COC as a reference sequence)

Mutation site types	Number (percentage) of mutations	Number (percentage) of synonymous mutations	Number (percentage)of non-synonymous mutations
T→C	239 (20.59%)	231 (19.90%)	8 (0.69%)
C→T	382 (32.90%)	363 (31.27%)	19 (1.64%)
A→T	42 (3.62%)	38 (3.27%)	4 (0.34%)
A→G	141 (12.14%)	123 (10.59%)	18 (1.55%)
A→C	73 (6.29%)	72 (6.20%)	1 (0.09%)
T→A	48 (4.13%)	44 (3.79%)	4 (0.34%)
T→G	7 (0.60%)	7 (0.60%)	0 (0.00%)
G→A	123 (10.59%)	103 (8.87%)	20 (1.72%)
G→T	9 (0.78%)	8 (0.69%)	1 (0.09%)
G→C	11 (0.95%)	9 (0.78%)	2 (0.17%)
C→A	81 (6.98%)	79 (6.80%)	2 (0.17%)
C→G	5 (0.43%)	4 (0.34%)	1 (0.09%)
	1,161 (100%)	1,081 (93.11%)	80 (6.89%)

phenomenon; epigenetic changes contribute to the survival and reproduction of nascent species and may play an important role in the evolution of species. In contrast, significant epigenetic changes have also been observed in the mtDNA of zebrafish (Tables 2 and 4). Furthermore, hybridization can induce changes in DNA methylation, and DNA methylation plays a pivotal role in genome regulation and gene expression, thereby affecting growth, development and phenotypes. Subsequent to hybridization, 2*n*NCRC exhibited novel characteristics that were different from those of its parents (Wang et al., 2019; Wang et al., 2017).

Analysis of genetic variations of mitochondrial sequences in different generations of 2*n*NCRC

Via comparative analyses of the whole-mt-genome sequences of different generations of 2*n*NCRC, these changing methylation sites of 2*n*NCRC were seen to be stably inherited by subsequent generations after continuous self-crossing. In contrast to stable heredity, the mt sequences between different generations of 2*n*NCRC also underwent mutation, which will be discussed in detail in this section. We directly sequenced eight mt structural regions (genes) using three 2*n*COC, three 2*n*BSB, three 2*n*NCRC, three 2*n*NCRC-

Table 4 Comparison of codon usage (number of codons) among 2*n*NCRC, 2*n*COC, 2*n*BSB, 2*n*CC, and zebrafish[a)]

Amino acid	Codon	2*n*COC	2*n*NCRC	2*n*BSB	2*n*CC	Zebrafish
Lys	AAA	71	70	70	70	81
	AAG	6	9	9	7	8
Asn	AAC*	77	75	74	76	65
	AAT*	43	48	47	45	56
Thr	ACA	146	150	145	149	140
	ACC*	105	100	102	94	77
	ACG	9	12	12	15	13
	ACT*	41	42	49	45	62
Ser	AGC*	44	34	41	37	38
	AGT*	8	14	15	13	14
Met	ATG	48	45	51	47	52
Ile	ATA	127	130	126	129	139
	ATC*	140	92	147	92	107
	ATT*	156	195	139	195	201
Gln	CAA	99	95	86	94	84
	CAG	2	6	11	7	10
His	CAC*	81	78	75	78	68
	CAT*	23	26	29	26	34
Pro	CCA	117	124	116	116	114
	CCC	50	56	69	58	40
	CCG	11	4	9	9	13
	CCT	34	28	23	27	40
Arg	CGA	45	52	50	52	53
	CGC*	14	10	12	9	4
	CGG	8	4	3	4	9
	CGT*	9	12	11	11	9
Leu	CTA	294	269	285	261	180
	CTC*	101	83	92	85	38
	CTG	41	39	33	43	29
	CTT*	78	99	102	99	146
Glu	GAA	95	91	82	93	80
	GAG	6	10	21	8	20
Asp	GAC*	65	57	62	56	54
	GAT*	11	19	14	20	30
Ala	GCA	130	127	122	128	132
	GCC*	146	128	147	132	96
	GCG	9	8	15	5	13
	GCT*	52	70	47	69	87
Gly	GGA	140	133	125	137	116
	GGC	43	48	42	47	39
	GGG	33	41	42	36	45
	GGT	33	29	34	30	36
Val	GTA	118	112	102	110	111
	GTC*	31	26	58	25	21
	GTG	18	26	29	27	22
	GTT*	57	62	45	65	69
Stop	TAA	7	6	8	6	7
	TAG	3	4	2	4	4
Tyr	TAC*	65	56	68	56	57
	TAT*	50	58	45	59	55
Ser	TCA	87	88	80	87	108
	TCC	59	60	64	60	19
	TCG	9	4	7	4	9
	TCT	30	33	36	34	57
Trp	TGA	112	109	107	110	107
	TGG	8	11	14	10	11
Cys	TGC	20	20	16	20	15
	TGT	5	5	9	5	14
Leu	TTA	99	127	91	128	181
	TTG	11	9	16	12	26
Phe	TTC*	145	120	153	119	88
	TTT*	79	106	72	109	148

a) The asterisks (*) indicate codons in which mutations had occurred at the third position (C was replaced by T) for the same amino acid in the mitochondria.

F_2, and three 2*n*NCRC-F_3 as DNA templates, respectively. We focused on selecting eight structural regions (genes) for sequencing, such as the non-coding CR, which is the region with the largest variation in mt sequence; the *12S rRNA* and *16S rRNA* genes, which are the slowest evolving genes in the mt sequence; and the *COI*, *Cytb*, NADH 2 (*ND2*), NADH 4 (*ND4*), and NADH 5 (*ND5*) genes, which provide information regarding phylogenetic evolution.

Figure 4, Tables 5 and 6, and Figures S1–S8 in Supporting Information directly show the genetic variations in these eight mt structural regions (genes) in different generations of 2*n*NCRC. In these structural regions (genes) (only the consistent base sites of different generations of 2*n*NCRC are counted here), most base sites were conserved: 67.61% of the sites were conserved in the CR; 88.76% to 90.36% were conserved in the two rRNAs (*12S rRNA* and *16S rRNA*); and 71.33%–79.75% were conserved in the five protein-coding genes (*COI*, *Cytb*, *ND2*, *ND4*, and *ND5*). In these structural regions (genes) (only the consistent base sites of different generations of 2*n*NCRC are counted here), 15.93% of the base sites in the CR, 5.83% to 6.71% of the base sites in the two rRNAs, and 6.90%–12.50% of the base sites in the five protein-coding genes were consistent with the female parent 2*n*COC. It is widely believed that animal mtDNA follows the mechanism of maternal inheritance. Recombination of mtDNA occurs commonly in most plants, fungal and protist species but was traditionally thought to be rare or absent in animals (Gillham, 1994; Rokas et al., 2003). In this study, the mt-genome sequence of 2*n*NCRC did not follow the mechanism of maternal inheritance, instead exhibiting partial embedding of paternal base sites due to the influence of distant hybridization. For example, 4.33% of the base sites in the CR, 0.52%–1.90% of the base sites in the two rRNAs, and 3.68%–5.87% of the base sites in the five protein-coding genes were consistent with the male parent 2*n*BSB. The paternal mtDNA fragments were stably embedded in the eight mt structural regions (genes) of different generations (F_1–F_3) of 2*n*NCRC to form chimeric DNA fragments. In addition, some base sites had undergone mutation (only the consistent base sites of different generations of 2*n*NCRC are counted here): 7.26% of the sites were mutated in the CR; 1.57%–2.56% were mutated in the two rRNAs; and 6.37%–7.26% were mutated in the five protein-coding genes. Most of these mutation sites were consistent with 2*n*CC, suggesting that along with the influence of distant hybridization, 2*n*NCRC has a tendency to mutate to 2*n*CC at the mtDNA level.

Notably, the eight mt structural regions (genes) of 2*n*NCRC exhibited inconsistent base sites in different generations (Figure 4, Tables 5 and 6, Figures S1–S8 in Supporting Information). That is, some of the base sites of the first generation (F_1) of 2*n*NCRC were not stably inherited by the second generation (F_2), resulting in inconsistent sites, but almost all the base sites were stably inherited from the second generation (F_2) by the third generation (F_3). For example, 4.88% of the base sites in the CR, 0.84%–0.95% of the base sites in the two rRNAs, and 2.58%–4.82% of the base sites in the five protein-coding genes were inconsistent between 2*n*NCRC F_1 and 2*n*NCRC F_2-F_3. Figure 4, Tables 5 and 6, and Figures S1–S8 also show that the eight mt structural regions (genes) of 2*n*NCRC F_1, which was formed by a cross between 2*n*COC (♀) and 2*n*BSB (♂), were unstable, and different individuals in the same population (2*n*NCRC F_1) exhibited multiple patterns of genetic variation. Among the same base sites, some were inherited from the female parent 2*n*COC, while others were inherited from the male parent 2*n*BSB or were mutated. There were 2.17% polymorphic base sites in the CR of 2*n*NCRC F_1, 0.48%–0.63% polymorphic base sites in two rRNAs, and 1.35%–2.20% polymorphic base sites in the five protein-coding genes. Our results revealed the occurrence of rapid homoploid speciation in the first generation (2*n*NCRC F_1) of a distant hybridization and indicated the instability of the newly established homodiploid mt-genome.

DISCUSSION

Scientists generally believe that mtDNA has some unique characteristics, e.g., maternal inheritance, lack of recombination, and elevated mutation rates; therefore, mtDNA has often been considered a useful tool in phylogeography, evolutionary biology, population genetics and phylogenetic studies (Gissi et al., 2008; You et al., 2014; Li and Gui, 2018; Zhang and Gui, 2018). Following fertilization, the nuclear genome regulates the destruction of paternal mtDNA, maintaining the maternal inheritance mode of the mt-genome (Morgan et al., 2013). Unfortunately, the molecular mechanisms underlying paternal mtDNA elimination are only partially elucidated. In fact, it appears likely that a different combination of mechanisms operates depending on the species in question (Luo et al., 2018). Genetic studies have reported paternal leakage of mtDNA and recombination in animals, including fish (Ciborowski et al., 2007; Morgan et al., 2013; Wolff et al., 2008), e.g., the nematode *M. javanica* (Lunt and Hyman, 1997), the mussel sister-species *M. galloprovincialis* (Ladoukakis and Zouros, 2001) and *M. trossulus* (Burzyński et al., 2003), the flatfish *P. flesus* (Hoarau et al., 2002), *Drosophila* (Dokianakis and Ladoukakis, 2014; Nunes et al., 2013), mouse (Gyllensten et al., 1991), sheep (Zhao et al., 2004), humans (Kraytsberg et al., 2004; Luo et al., 2018), and the triploid crucian carp (Guo et al., 2006; Liu, 2014). Documented paternal leakage of mtDNA is more commonly observed in hybrid zones because these environments are thought to exhibit less stringent mechanisms for prevention of paternal leakage (Morgan et al., 2013; Wolff et al., 2008). In hybrid zones, intraspecific paternal

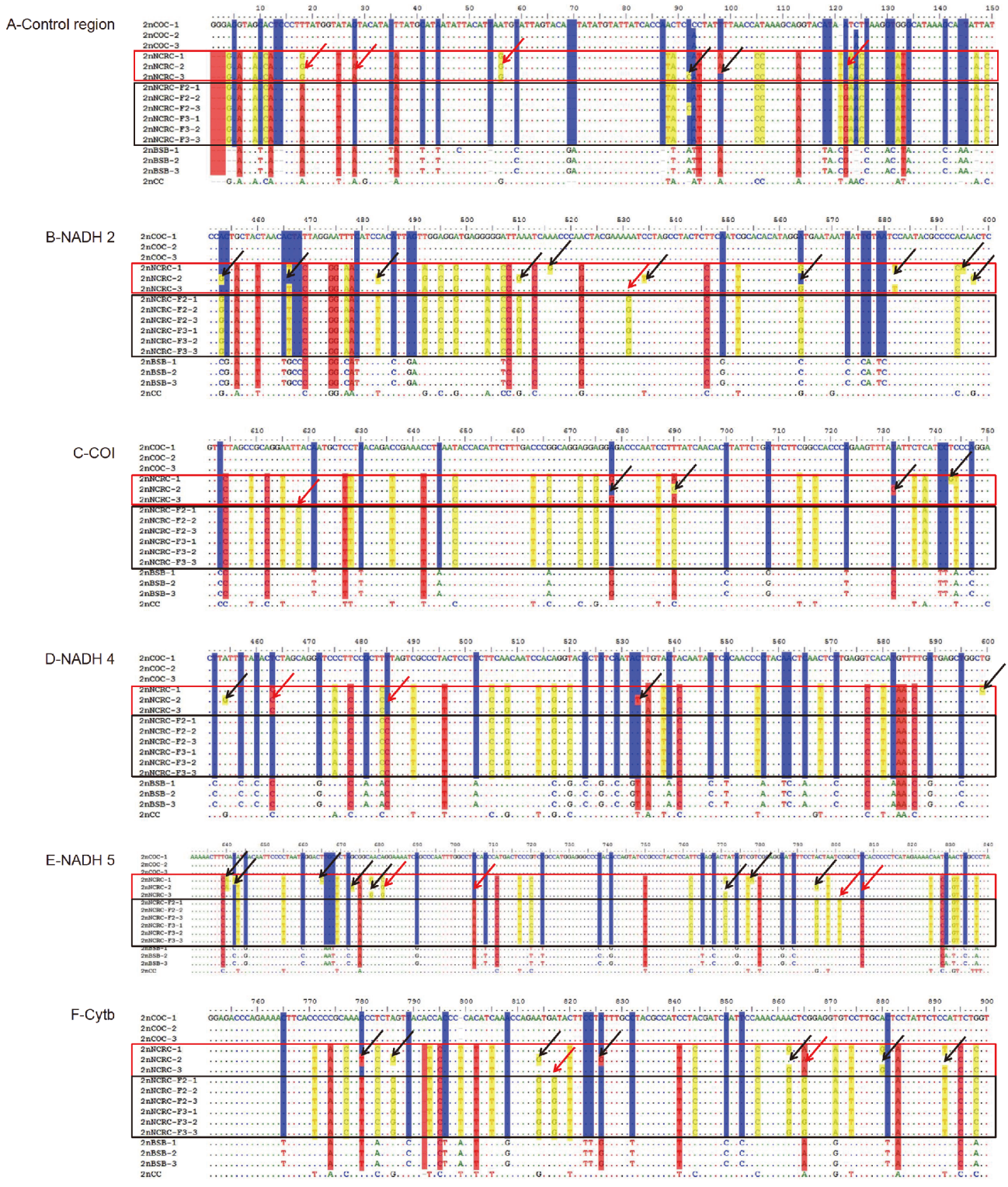

Figure 4 Nucleotide sequence alignment of mitochondrial structural regions (genes) of three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC. A, Partial nucleotide sequence alignment for the control region (CR). B, Partial nucleotide sequence alignment for NADH 2 (*ND2*). C, Partial nucleotide sequence alignment for *COI*. D, Partial nucleotide sequence alignment for NADH 4 (*ND4*). E, Partial nucleotide sequence alignment for NADH 5 (*ND5*). F, Partial nucleotide sequence alignment for *Cytb*. The dots indicate sequence identity, and the hyphens represent insertions/deletions. Blue denotes loci from 2*n*COC that were different from 2*n*BSB; red variants represent loci from 2*n*BSB that were different from 2*n*COC; yellow represents mutations that differed from both 2*n*COC and 2*n*BSB. The red box shows the nucleotide sequence alignment of 2*n*NCRC (2*n*NCRC-F_1); the black box shows the nucleotide sequence alignment of 2*n*NCRC-F_2 and 2*n*NCRC-F_3. The red arrow indicates that the base sites in the mitochondrial structural regions (genes) were inconsistent between 2*n*NCRC F_1 and 2*n*NCRC F_2-F_3; the black arrow indicates the polymorphic base sites in mitochondrial structural regions (genes) of 2*n*NCRC F_1.

Table 5 Base site composition of eight mitochondrial structural regions (genes) of 2*n*NCRC in different generations

Name of genes	Size (bp)	Number (percentage) of conserved base sites	Number (percentage) of base sites of maternal consistency			Number (percentage) of base sites of paternal consistency			Number (percentage) of mutation sites		
			F_1–F_3	F_1	F_2–F_3	F_1–F_3	F_1	F_2–F_3	F_1–F_3	F_1	F_2–F_3
Control region	923	624 (67.61%)	147 (15.93%)	2 (0.22%)	4 (0.43%)	40 (4.33%)	1 (0.11%)	5 (0.54%)	67 (7.26%)	4 (0.43%)	9 (0.98%)
12S rRNA	954	862 (90.36%)	64 (6.71%)	0 (0.00%)	0 (0.00%)	5 (0.52%)	0 (0.00%)	0 (0.00%)	15 (1.57%)	1 (0.10%)	1 (0.10%)
16S rRNA	1,681	1,492 (88.76%)	98 (5.83%)	0 (0.00%)	1 (0.06%)	32 (1.90%)	1 (0.06%)	1 (0.06%)	43 (2.56%)	1 (0.06%)	4 (0.24%)
NADH 2	1,047	759 (72.49%)	111 (10.60%)	0 (0.00%)	5 (0.48%)	57 (5.44%)	2 (0.19%)	0 (0.00%)	76 (7.26%)	2 (0.19%)	12 (1.15%)
COI	1,551	1,237 (79.75%)	107 (6.90%)	0 (0.00%)	8 (0.52%)	66 (4.26%)	0 (0.00%)	3 (0.19%)	101 (6.51%)	2 (0.13%)	6 (0.39%)
NADH 4	1,381	1,035 (74.95%)	125 (9.05%)	2 (0.14%)	9 (0.65%)	81 (5.87%)	2 (0.14%)	3 (0.22%)	88 (6.37%)	4 (0.29%)	8 (0.58%)
NADH 5	1,824	1,301 (71.33%)	228 (12.50%)	5 (0.27)	6 (0.33%)	94 (5.15%)	2 (0.11%)	5 (0.27%)	125 (6.85%)	4 (0.22%)	19 (1.04%)
Cytb	1,141	875 (76.69%)	95 (8.33%)	0 (0.00%)	8 (0.70%)	42 (3.68%)	2 (0.18%)	3 (0.26%)	74 (6.49%)	3 (0.26%)	15 (1.31%)

Table 6 Polymorphic base sites of eight mitochondrial structural regions (genes) of 2*n*NCRC F_1

Name of genes	Size (bp)	Number of unstable base sites of F_1			
		Maternal or paternal	Maternal or mutant	Paternal or mutant	Conserved or mutant
Control region	923	1 (0.11%)	4 (0.43%)	1 (0.11%)	14 (1.52%)
12S rRNA	954	0 (0.00%)	0 (0.00%)	0 (0.00%)	6 (0.63%)
16S rRNA	1,681	1 (0.06%)	1 (0.06%)	1 (0.06%)	5 (0.30%)
NADH 2	1,047	1 (0.10%)	5 (0.48%)	0 (0.00%)	17 (1.62%)
COI	1,551	8 (0.52%)	4 (0.26%)	2 (0.13%)	7 (0.45%)
NADH 4	1,381	5 (0.36%)	4 (0.29%)	2 (0.14%)	13 (0.94%)
NADH 5	1,824	4 (0.22%)	4 (0.22%)	1 (0.05%)	26 (1.43%)
Cytb	1,141	8 (0.70%)	3 (0.26%)	0 (0.00%)	13 (1.14%)

leakage is difficult to detect, but paternal mtDNA inherited phenomena have been reported in human (Schwartz and Vissing, 2002); paternal leakage in interspecies hybrids may be more common than previously believed, as many cases of paternal leakage involve interspecies hybrids, which might interfere with mitonuclear interactions and alter the proper function of the mechanisms that eliminate paternal mitochondria (Breton and Stewart, 2015). For example, Kondo et al. reported that among 331 lines from a *Drosophila simulans* (♀)×*Drosophila mauritiana* (♂) interspecies cross, four lines exhibited clear evidence of paternal leakage of mtDNA (Kondo et al., 1990). Gyllensten et al. detected paternal mtDNA in interspecies mt congenic mice derived from a backcrosses between *Mus musculus* and *Mus spretus* (Gyllensten et al., 1991). Among reports on hybrid fish, Guo et al. also reported a case of mtDNA recombination in one triploid fish that derived from distant hybridization of Japanese crucian carp (*Carassius auratus cuvieri*, 2*n*=100, ♀)×allotetraploid hybrids (4*n*=200, ♂) (Guo et al., 2006). For interspecies hybrids (e.g., experimental crosses with both fruit fly and cattle), it is thought that differences in sequences of genes encoding the destruction-related proteins might result in failure of the mechanism via which the nuclear genome regulates the destruction of paternal mtDNA to maintain the maternal inheritance mode (Sherengul et al., 2006; Sutovsky et al., 2000).

In natural environments, hybridization has been extensively observed in many groups of fish, from sharks to some teleosts (Dehal and Boore, 2005). In our previous research, we reported the spontaneous occurrence of a newborn homodiploid crucian carp-like fish (2*n*NCRC) that originated from the interspecific hybridization of female common carp (2*n*COC)×male blunt snout bream (2*n*BSB). The phenotypes and genotypes of 2*n*NCRC differed from those of its parents but were closely related to those of 2*n*CC. Our previous results indicated that a source of crucian carp might be present in the hybridization pathway of female common carp×male blunt snout bream at the phenotype and genomic DNA levels (Wang et al., 2019; Wang et al., 2017; Wang et al., 2018). In this study, we provide direct evidence of mtDNA recombination in the 2*n*NCRC lineage from F_1 to F_3 that originated from 2*n*COC (♀)×2*n*BSB (♂). The paternal mtDNA fragments (0.52%–5.87% of the base sites) were stably embedded in the eight mt structural regions (genes) of different generations (F_1–F_3) of 2*n*NCRC to form chimeric DNA fragments. Along with this chimeric process, some base sites (1.57%–7.26%) in the eight mt structural regions (genes) of different generations (F_1–F_3) of 2*n*NCRC underwent mutations. Notably, most of these mutation sites were consistent with 2*n*CC, suggesting that along with the influence of distant hybridization, 2*n*NCRC has a tendency to mutate to 2*n*CC at the mtDNA level. Furthermore, some of

the base sites (0.84%–4.88%) in the eight mt structural regions (genes) of the first generation (F_1) of 2*n*NCRC were not stably inherited by the second generation (F_2), resulting in inconsistent sites, but almost all the base sites were stably inherited from the second generation (F_2) by the third generation (F_3), indicating that 2*n*NCRC had eventually established a stable genetic lineage (F_1–F_3). In addition, different individuals in the same population (2*n*NCRC F_1) exhibited multiple patterns of genetic variation in the eight mt structural regions (genes). Among the same base sites (0.48%–2.20%), some were inherited from the female parent 2*n*COC, while others were inherited from the male parent 2*n*BSB or were mutated. Our results revealed the occurrence of rapid homoploid speciation in the first generation (2*n*NCRC F_1) of a distant hybridization and indicated the instability of the newly established homodiploid mt-genome. Our analysis results also showed that the mtDNA organization and nucleotide composition of 2*n*NCRC were more similar to those of 2*n*CC than to those of the parents. Compared to the parents, the T in the overall nucleotide composition of the 13 protein-coding genes of 2*n*NCRC was mainly derived from C (cytosine methylation) in 2*n*COC, especially at the third codon positions, and these substitutions were almost synonymous mutations. To a certain extent, epigenetic changes contribute to the survival and reproduction of nascent species and may play an important role in the evolution of species. Hybridization can induce changes in DNA methylation, and DNA methylation plays a pivotal role in genome regulation and gene expression, thereby affecting growth, development and phenotype (Bird, 1992; Chen et al., 2008; Li et al., 2011).

This comparative study of the sequence identities of mtDNA between 2*n*NCRC and its parents provided direct evidence for that the paternal mtDNAs were transmitted into the mt-genomes of homodiploid lineage. Our observations regarding the complete mt-genome indicated that homologous recombination (via mechanisms involving double-strand break repair and/or break-induced replication pathways) could occur in any part of the 2*n*NCRC mt-genome. For the mt homologous recombination, some previous studies can be used as supporting evidence. Thyagarajan et al. observed that mitochondria in cultured human cells contain enzymes that catalyse nuclear recombination, indicating that these enzymes were imported into the mitochondrion as well as the nucleus and that the molecular mechanism underlying mtDNA recombination was similar to that underlying nuclear recombination (Thyagarajan et al., 1996). Kraytsberg et al. observed mtDNA recombinants in the muscle tissue of a man with paternal inheritance of the mt-genome, where heterologous (paternal and maternal) mtDNAs were mixed and thus may have opportunity to recombine (Kraytsberg et al., 2004). The results of this research also suggested that the enzymes that were responsible for replicating mtDNA stopped replicating maternal mtDNA and jumped to the corresponding paternal mtDNA position to replicate paternal mtDNA (Guo et al., 2006; Kraytsberg et al., 2004). Recently, it has also been reported that mitochondrial endonuclease G relocates from the intermembrane space of paternal mitochondria to the matrix after fertilization where it proceeds to degrade or eliminate paternal mtDNA (Luo et al., 2018; Zhou et al., 2016). It is not difficult to imagine what this experimental result means.

In this study, we observed that the paternal mtDNA fragments in the newborn homodiploid crucian carp-like fish that originated from the interspecific hybridization of female common carp×male blunt snout bream. Our findings demonstrate for the first time that the paternal mtDNA can be transmitted to the offspring of homodiploid cyprinid. We first revealed the instability of the mtDNA of F_1 of homoploid hybrid species formed by distant hybridization but eventually established a relatively genetically stable hybrid fish lineage (F_1–F_3). The inheritable chimeric DNA fragments and mutant loci in the mt-genomes of different generations of 2*n*NCRC provided important evidence of the mtDNA change process in the newborn lineage derived from hybridization of different species. Stable homologous recombination in the mt-genomes of different generations of 2*n*NCRC will have an important effect on our understanding of mtDNA mutation and repair mechanisms and rates of mutation accumulation. There is solid evidence that homologous recombination is essential for DNA repair in yeast (Ling et al., 1995) and is expected to play a similar role in animal mtDNA (Howell, 1997; Thyagarajan et al., 1996). Furthermore, with the development of next-generation sequencing technology, an increasing number of mt-genomes and transcriptomes are being examined to infer phylogenetic relationships (Bazinet et al., 2013; Kawahara and Breinholt, 2014; Qin et al., 2015). We plan to study nuclear DNA markers, including via transcriptome analysis, in a future study of the genetic variation characteristics of this newborn homodiploid crucian carp-like fish.

MATERIALS AND METHODS

Ethics statement

The guidelines established by the Administration of Affairs Concerning Animal Experimentation state that approval from the Science and Technology Bureau of China and the Department of Wildlife Administration is not necessary when the fish in question are neither rare nor near extinction (first- or second-class state protection level). Therefore, approval was not required for the experiments conducted in this study.

Animals and crossing procedure

All of the nature material, such as common carp (abbreviated

as 2*n*COC), blunt-snout bream (abbreviated as 2*n*BSB) were obtained from the State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha, China. The protocols for crosses and culturing were described previously (Wang et al., 2017). The newborn homodiploid crucian carp-like fish (abbreviated as 2*n*NCRC) obtained by artificial hybridization was cultured in ponds at the State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha, China, and fed with artificial feed. The 2*n*NCRC self-cross resulted in a single offspring. Hereafter, the newborn homodiploid crucian carp-like fish produced by 2*n*COC (♀)×2*n*BSB (♂) is referred to as 2*n*NCRC and the self-cross offspring as 2*n*NCRC-F_2 and then the self-cross offspring as 2*n*NCRC-F_3. All fishes were deeply anaesthetized with 100 mg L^{-1} MS-222 (Sigma-Aldrich, St. Louis, MO, USA) prior to dissection.

DNA extraction, PCR amplification, cloning and sequencing

Total genomic DNA from the peripheral blood cells of 2*n*COC, 2*n*BSB, 2*n*NCRC, 2*n*NCRC-F_2, and 2*n*NCRC-F_3 by routine approaches (Sambrook and Russell David, 1989) were used as templated, respectively. The highly conserved and newly designed PCR primers (You et al., 2014) were used to amplify up to 22 contiguous and overlapping fragments of the complete mt sequences in 2*n*NCRC, 2*n*NCRC-F_2, and 2*n*NCRC-F_3. Furthermore, the non-coding CR and the *12S rRNA*, *16S rRNA*, *COI*, *Cytb*, NADH 2 (*ND2*), NADH 4 (*ND4*), and NADH 5 (*ND5*) genes of three 2*n*COC, three 2*n*BSB, three 2*n*NCRC, three 2*n*NCRC-F_2, and three 2*n*NCRC-F_3 were amplified by PCR using the primers listed in Table S1 in Supporting Information. The PCRs were performed in a volume of 50 μL with approximately 10–30 ng of genomic DNA, 1.5 mmol L^{-1} $MgCl_2$, 250 μmol L^{-1} dNTP, 0.4 μmol L^{-1} each primer, and 1.25 U of Taq polymerase (TaKaRa, Dalian, China). The thermal program consisted of an initial denaturation step at 94°C for 5 min, followed by 30 cycles of 94°C for 35 s, 50–60°C for 60 s, and 72°C for 60–150 s and a final extension step at 72°C for 10 min (You et al., 2014). A majority of the PCR products were directly sequenced, and some fragments that were difficult to sequence using PCR products were cloned into the pMD18-T vector (TaKaRa, Dalian, China), the plasmids were transformed into *E. coli* DH5a and purified (Guo et al., 2006). At least three clones from each PCR product were sequenced with vector-specific primers using the primer walking method on an ABI 3730XL automatic sequencer (ABI PRISM 3730, Applied Biosystems, CA, USA) (You et al., 2014).

Sequence analysis

The sequence homology and variation among the fragments amplified from 2*n*COC, 2*n*BSB, 2*n*NCRC, 2*n*NCRC-F_2, 2*n*NCRC-F_3, and 2*n*CC were analysed using BioEdit (Hall, 1999) and the DNAstar 5.0 software package (DNAstar Inc.). The protein-coding genes, non-coding control region, ribosomal RNA genes, and transfer RNA genes were identified using BLAST (http://www.ncbi.nlm.nih.gov) searches and by comparison with other cyprinid mt-genomic sequences. All sequences were analysed using the BLAST, ClustalW (http://www.ebi.ac.uk/) (Thompson et al., 1994) and MEGA 4.0 (Tamura et al., 2007) programs to determine identity (You et al., 2014). The whole-mt-genomes of common carp (KF856965.1), blunt snout bream (NC_010341.1), and local crucian carp (GU086395.1) were retrieved from the GenBank database.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Natural Science Foundation of China (31802287, 31730098, 31430088), the China Postdoctoral Science Foundation (2018M632969), the earmarked fund for China Agriculture Research System (CARS-45), the Key Research and Development Program of Hunan Province (2018NK2072), Hunan Provincial Natural Science and Technology Major Project (2017NK1031), and the Cooperative Innovation Center of Engineering and New Products for Developmental Biology of Hunan Province (20134486).*

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., et al. (2013). Hybridization and speciation. J Evol Biol 26, 229–246.

Baack, E.J., and Rieseberg, L.H. (2007). A genomic view of introgression and hybrid speciation. Curr Opin Genet Dev 17, 513–518.

Bazinet, A.L., Cummings, M.P., Mitter, K.T., and Mitter, C.W. (2013). Can RNA-Seq resolve the rapid radiation of advanced moths and butterflies (Hexapoda: Lepidoptera: Apoditrysia)? An exploratory study. PLoS ONE 8, e82615.

Bird, A. (1992). The essentials of DNA methylation. Cell 70, 5–8.

Boore, J.L. (1999). Animal mitochondrial genomes. Nucleic Acids Res 27, 1767–1780.

Breton, S., and Stewart, D.T. (2015). Atypical mitochondrial inheritance patterns in eukaryotes. Genome 58, 423–431.

Broughton, R.E., Milam, J.E., and Roe, B.A. (2001). The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. Genome Res 11, 1958–1967.

Burzyński, A., Zbawicka, M., Skibinski, D.O.F., and Wenne, R. (2003). Evidence for Recombination of mtDNA in the Marine Mussel *Mytilus trossulus* from the Baltic. Mol Biol Evol 20, 388–392.

Chen, M., Ha, M., Lackey, E., Wang, J., and Chen, Z.J. (2008). RNAi of *met1* reduces DNA methylation and induces genome-specific changes in gene expression and centromeric small RNA accumulation in *Arabidopsis* allopolyploids. Genetics 178, 1845–1858.

Ciborowski, K.L., Consuegra, S., García de Leániz, C., Beaumont, M.A., Wang, J., and Jordan, W.C. (2007). Rare and fleeting: an example of interspecific recombination in animal mitochondrial DNA. Biol Lett 3, 554–557.

Dehal, P., and Boore, J.L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3, e314.
Dokianakis, E., and Ladoukakis, E.D. (2014). Different degree of paternal mtDNA leakage between male and female progeny in interspecific *Drosophila* crosses. Ecol Evol 4, 2633–2641.
Gillham, N.W. (1994). Organelle Genes and Genomes. (Oxford University Press).
Gissi, C., Iannelli, F., and Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101, 301–320.
Gui, J.F., and Zhou, L. (2010). Genetic basis and breeding application of clonal diversity and dual reproduction modes in polyploid *Carassius auratus gibelio*. Sci China Life Sci 53, 409–415.
Guo, W., and Gui, J.F. (2008). Microsatellite marker isolation and cultured strain identification in *Carassius auratus gibelio*. Aquacult Int 16, 497–510.
Guo, X., Liu, S., and Liu, Y. (2006). Evidence for recombination of mitochondrial DNA in triploid crucian carp. Genetics 172, 1745–1749.
Gyllensten, U., Wharton, D., Josefsson, A., and Wilson, A.C. (1991). Paternal inheritance of mitochondrial DNA in mice. Nature 352, 255–257.
Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT[C]//. Nucleic acids symposium series. London: Information Retrieval Ltd., c1979-c2000., 41(41), 95–98.
Hoarau, G., Holla, S., Lescasse, R., Stam, W.T., and Olsen, J.L. (2002). Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish *Platichthys flesus*. Mol Biol Evol 19, 2261–2264.
Howell, N. (1997). mtDNA recombination: what do in vitro data mean? Am J Hum Genet 61, 19–22.
Kawahara, A.Y., and Breinholt, J.W. (2014). Phylogenomics provides strong evidence for relationships of butterflies and moths. Proc R Soc B-Biol Sci 281, 20140970.
Kondo, R., Satta, Y., Matsuura, E., Ishiwa, H., Takahata, N., and Chigusa, S. (1990). Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. Genetics 126, 657–663.
Kraytsberg, Y., Schwartz, M., Brown, T.A., Ebralidse, K., Kunz, W.S., Clayton, D.A., Vissing, J., and Khrapko, K. (2004). Recombination of human mitochondrial DNA. Science 304, 981.
Ladoukakis, E.D., and Zouros, E. (2001). Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. Mol Biol Evol 18, 1168–1175.
Li, X.Y., and Gui, J.F. (2018). Diverse and variable sex determination mechanisms in vertebrates. Sci China Life Sci 61, 1503–1514.
Li, Z.H., Lu, X., Gao, Y., Liu, S.J., Tao, M., Xiao, H., Qiao, Y.Q., Zhang, Y. P., and Luo, J. (2011). Polyploidization and epigenetics. Chin Sci Bull 56, 245–252.
Ling, F., Makishima, F., Morishima, N., and Shibata, T. (1995). A nuclear mutation defective in mitochondrial recombination in yeast.. EMBO J 14, 4090–4101.
Liu, S. (2014). Fish distant hybridization. (Beijing: China Social Sciences Press).
Lunt, D.H., and Hyman, B.C. (1997). Animal mitochondrial DNA recombination. Nature 387, 247.
Luo, S., Valencia, C.A., Zhang, J., Lee, N.C., Slone, J., Gui, B., Wang, X., Li, Z., Dell, S., Brown, J., et al. (2018). Biparental inheritance of mitochondrial DNA in humans. Proc Natl Acad Sci USA 115, 13039–13044.
Mallet, J. (2007). Hybrid speciation. Nature 446, 279–283.
Mei, J., and Gui, J.F. (2015). Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish. Sci China Life Sci 58, 124–136.
Morgan, J.A.T., Macbeth, M., Broderick, D., Whatmore, P., Street, R., Welch, D.J., and Ovenden, J.R. (2013). Hybridisation, paternal leakage and mitochondrial DNA linearization in three anomalous fish (Scombridae). Mitochondrion 13, 852–861.
Nunes, M.D.S., Dolezal, M., and Schlötterer, C. (2013). Extensive paternal mtDNA leakage in natural populations of *Drosophila melanogaster*. Mol Ecol 22, 2106–2117.
Qin, J., Zhang, Y., Zhou, X., Kong, X., Wei, S., Ward, R.D., and Zhang, A. (2015). Mitochondrial phylogenomics and genetic relationships of closely related pine moth (Lasiocampidae: *Dendrolimus*) species in China, using whole mitochondrial genomes. BMC Genomics 16, 428.
Richards, E.J., and Elgin, S.C.R. (2002). Epigenetic codes for heterochromatin formation and silencing. Cell 108, 489–500.
Rokas, A., Ladoukakis, E., and Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. Trends Ecol Evol 18, 411–417.
Sambrook, J., and Russell David, W. (1989). Molecular Cloning: A Laboratory Manual. (Vol. 3, Cold Spring Harbor Laboratory Press).
Schwartz, M., and Vissing, J. (2002). Paternal inheritance of mitochondrial DNA. N Engl J Med 347, 576–580.
Sherengul, W., Kondo, R., and Matsuura, E.T. (2006). Analysis of paternal transmission of mitochondrial DNA in *Drosophila*. Genes Genet Syst 81, 399–404.
Sutovsky, P., Moreno, R.D., Ramalho-Santos, J.., Dominko, T., Simerly, C., and Schatten, G. (2000). Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. Biol Reprod 63, 582–590.
Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24, 1596–1599.
Thompson, J.D., Higgins, D.G., and 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. Nucl Acids Res 22, 4673–4680.
Thyagarajan, B., Padua, R.A., and Campbell, C. (1996). Mammalian mitochondria possess homologous DNA recombination activity. J Biol Chem 271, 27536–27543.
Wang, S., Tang, C., Tao, M., Qin, Q., Zhang, C., Luo, K., Zhao, R., Wang, J., Ren, L., Xiao, J., et al. (2019). Establishment and application of distant hybridization technology in fish. Sci China Life Sci 62, 22–45.
Wang, S., Ye, X., Wang, Y., Chen, Y., Lin, B., Yi, Z., Mao, Z., Hu, F., Zhao, R., Wang, J., et al. (2017). A new type of homodiploid fish derived from the interspecific hybridization of female common carp × male blunt snout bream. Sci Rep 7, 4189.
Wang, Y., Yang, C., Luo, K., Zhang, M., Qin, Q., Huo, Y., Song, J., Tao, M., Zhang, C., and Liu, S. (2018). The formation of the goldfish-like fish derived from hybridization of female Koi carp×male blunt snout bream. Front Genet 9, 437.
Wolff, J.N., Gandre, S., Kalinin, A., and Gemmell, N.J. (2008). Delimiting the frequency of paternal leakage of mitochondrial DNA in chinook salmon. Genetics 179, 1029–1032.
You, C.P., Zhao, R.R., Hu, J., Liu, S.J., Tao, M., Zhang, C., Chen, Y.B., Qin, Q.B., Xiao, J., Duan, W., et al. (2014). Inheritance of the complete mitochondrial genomes in three different ploidy fishes. Curr Mol Med 14, 1322–1330.
Yue, G.H., Liew, W.C., and Orban, L. (2006). The complete mitochondrial genome of a basal teleost, the Asian arowana (*Scleropages formosus*, Osteoglossidae). BMC Genomics 7, 242.
Zhang, Q.Y., and Gui, J.F. (2018). Diversity, evolutionary contribution and ecological roles of aquatic viruses. Sci China Life Sci 61, 1486–1502.
Zhao, X., Li, N., Guo, W., Hu, X., Liu, Z., Gong, G., Wang, A., Feng, J., and Wu, C. (2004). Further evidence for paternal inheritance of mitochondrial DNA in the sheep (*Ovis aries*). Heredity 93, 399–403.
Zhou, Q., Li, H., Li, H., Nakagawa, A., Lin, J.L.J., Lee, E.S., Harry, B.L., Skeen-Gaar, R.R., Suehiro, Y., William, D., et al. (2016). Mitochondrial endonuclease G mediates breakdown of paternal mitochondria upon fertilization. Science 353, 394–399.

SUPPORTING INFORMATION

Figure S1 Nucleotide sequence alignment of control region (CR) in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S2 Nucleotide sequence alignment of 12S rRNA in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S3 Nucleotide sequence alignment of 16S rRNA in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S4 Nucleotide sequence alignment of NADH 2 (ND2) in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S5 Nucleotide sequence alignment of COI in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S6 Nucleotide sequence alignment of NADH 4 (ND4) in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S7 Nucleotide sequence alignment of NADH 5 (ND5) in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Figure S8 Nucleotide sequence alignment of Cytb in three 2*n*COC, three 2*n*BSB, three 2*n*NCRC (2*n*NCRC-F_1), three 2*n*NCRC-F_2, three 2*n*NCRC-F_3, and one 2*n*CC.

Table S1 Genes and primers used for PCR analysis

The supporting information is available online at http://life.scichina.com and https://link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.