Hox genes reveal genomic DNA variation in tetraploid hybrids derived from Carassius auratus red var. (female) × Megalobrama amblycephala (male)

Y. D. Wang¹,²†, Q. B. Qin¹,²†, R. Yang¹,²†, W. Z. Sun¹,²†, Q. W. Liu¹,², Y. Y. Huo¹,², X. Huang¹,², M. Tao¹,², C. Zhang¹,², T. Li¹,² and S. J. Liu¹,²*

Abstract

Background: Allotetraploid F₁ hybrids (4nF₁) (AABB, 4n = 148) were generated from the distant hybridization of Carassius auratus red var. (RCC) (AA, 2n = 100) (♀) × Megalobrama amblycephala (BSB) (BB, 2n = 48) (♂). It has been reported that Hox gene clusters are highly conserved among plants and vertebrates. In this study, we investigated the genomic organization of Hox gene clusters in the allotetraploid F₁ hybrids and their parents to investigate the polyploidization process.

Results: There were three copies of Hox genes in the 4nF₁ hybrids, two copies in RCC and one copy in BSB. In addition, obvious variation and pseudogenization were observed in some Hox genes from 4nF₁.

Conclusion: Our results reveal the influence of polyploidization on the organization and evolution of Hox gene clusters in fish and also clarify some aspects of vertebrate genome evolution.

Keywords: Allotetraploid, Hox gene, polyploidization, pseudogenization

Background

Polyploidization is a widespread mechanism for speciation in eukaryotes, especially plants and vertebrates [1–5]. Polyploids with duplicated genomes may originate from a single species (autopolyploidy) or from different species through interspecific hybridization (allopolyploidy) [6]. Allopolyploids are prevalent in nature, suggesting there is an evolutionary advantage to obtaining multiple sets of genetic material for adaptation and development [7]. However, the molecular mechanisms underlying the processes and consequences of allopolyploidy remain unclear [8]. Polyploidy is relatively rare in animals compared with plants, and the influence of polyploidization on intragenomic variation in polyploid animals is poorly understood. In our earlier study, we successfully obtained fertile tetraploid hybrids from Carassius auratus red var. (RCC)(♀) × Megalobrama amblycephala (BSB)(♂) [9, 10]. RCC has 100 chromosomes and belongs to the Cyprinidae subfamily, while BSB has 48 chromosomes and belongs to the Cultrinae subfamily [11]. These new polyploid hybrids represent unique specimens for studying genomic changes in F₁ hybrids and could significantly contribute to our understanding of evolution.

Hox genes, a set of important developmental regulatory genes, are highly conserved and typically organized cluster [12]. In vertebrates, Hox genes encode two exons, and the highly conserved homeodomain (60 aa) is encoded by the second exon [13]. Recent research has shown that gene duplication, sequence variation, and selective pressure played crucial roles in the origin and evolution of Hox genes [14]. The earliest indications of genome duplication came from the comparative analysis of Hox genes and clusters from different chordate lineages [15–18].

* Correspondence: lsj@hunnu.edu.cn
† Equal contributors
1 State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha 410081, Hunan, People’s Republic of China
2 College of Life Sciences, Hunan Normal University, Changsha 410081, Hunan, People’s Republic of China

© The Author(s). 2017 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
| Species | Locus | Size (bp) | Exon1 (bp) | Intron (bp) | Exon 2 (bp) | GenBank accession no. |
|---------|-------|-----------|------------|-------------|-------------|----------------------|
| RCC     | HoxA4ai | 1181      | 89-500     | 501-974     | 975-1181    | JX282274             |
| RCC     | HoxA4a  | 1184      | 89-500     | 501-977     | 978-1184    | JX282275             |
| RCC     | HoxA9ai | 867       | 1-381      | 382-670     | 671-867     | JX282276             |
| RCC     | HoxA9a  | 819       | 1-381      | 382-622     | 623-819     | JX282277             |
| RCC     | HoxA2bi | 1486      | 1-314      | 315-901     | 902-1486    | JX282278             |
| RCC     | HoxA2b  | 1448      | 1-314      | 315-863     | 864-1448    | JX282279             |
| RCC     | HoxD4ai | 960       | 1-315      | 316-735     | 736-960     | JX282280             |
| RCC     | HoxD4a  | 952       | 1-315      | 316-719     | 728-952     | JX282281             |
| BSB     | HoxA4a  | 1188      | 89-500     | 501-981     | 982-1188    | JX282282             |
| BSB     | HoxA9a  | 879       | 1-381      | 382-682     | 683-879     | JX282283             |
| BSB     | HoxA2b  | 1479      | 1-311      | 312-894     | 895-1479    | JX282284             |
| BSB     | HoxD4a  | 911       | 1-306      | 307-686     | 687-911     | JX282285             |
| 4nF1    | HoxA4ai | 1183      | 89-500     | 501-976     | 977-1183    | JQ901468             |
| 4nF1    | HoxA4a  | 1169      | 89-500     | 501-962     | 963-1169    | JX282286             |
| 4nF1    | HoxA4a  | 1177      | 89-500     | 501-970     | 971-1177    | JX282287             |
| 4nF1    | HoxA9ai | 867       | 1-381      | 382-670     | 671-867     | JX282288             |
| 4nF1    | HoxA9a  | 817       | 1-381      | 382-620     | 621-817     | JX282289             |
| 4nF1    | HoxA9a  | 863       | 1-381      | 382-666     | 667-863     | JX282290             |
| 4nF1    | HoxA2bi | 1486      | 1-314      | 315-901     | 902-1486    | JX282291             |
| 4nF1    | HoxA2b  | 1448      | 1-314      | 315-863     | 864-1448    | JX282292             |
| 4nF1    | HoxA2b  | 1475      | 1-314      | 315-890     | 891-1475    | JX282293             |
| 4nF1    | HoxA11bi| 1251      | 3-590      | 591-1153    | 1154-1251   | JX282294             |
| 4nF1    | HoxA11b  | 1411     | 3-590      | 591-1313    | 1314-1411   | JX282295            |
| 4nF1    | HoxA11b  | 1437     | 3-590      | 591-1339    | 1340-1437   | JX282296            |
| 4nF1    | HoxB1bi | 733       | 1-477      | 478-567     | 568-733     | JX282297            |
| 4nF1    | HoxB1b  | 734       | 1-477      | 478-568     | 569-734     | JX282298            |
| 4nF1    | HoxB1b  | 731       | 1-477      | 478-565     | 566-731     | JX282299            |
| 4nF1    | HoxB5b  | 1196      | 1-561      | 562-990     | 991-1196    | JX282300            |
| 4nF1    | HoxB5b  | 1195      | 1-561      | 562-989     | 990-1196    | JX282301            |
| 4nF1    | HoxB5b  | 1190      | 1-561      | 562-984     | 985-1190    | JX282302            |
| 4nF1    | HoxB6bi | 807       | 1-169      | 170-667     | 668-807     | JX282303            |
| 4nF1    | HoxB6b  | 819       | 1-169      | 170-679     | 680-819     | JX282304            |
| 4nF1    | HoxB6b  | 812       | 1-169      | 170-672     | 673-819     | JX282304            |
| 4nF1    | HoxC4ai | 1176      | 1-410      | 411-935     | 936-1176    | JX282306            |
| 4nF1    | HoxC4ai | 1173      | 1-410      | 411-932     | 933-1173    | JX282307            |
| 4nF1    | HoxC4ai | 1169      | 1-410      | 411-928     | 929-1169    | JX282308            |
| 4nF1    | HoxC4a-1| 1179     | 1-410      | 411-938     | 939-1179    | JX282309            |
| 4nF1    | HoxD4ai | 960       | 1-315      | 316-735     | 736-960     | JX282310            |
| 4nF1    | HoxD4ai | 952       | 1-315      | 316-719     | 728-952     | JX282311            |
| 4nF1    | HoxD4a  | 942       | -          | -           | -           | JX282312            |
| 4nF1    | HoxD4a  | 897       | -          | -           | -           | JX282313            |
| 4nF1    | HoxD10a  | 1481     | -          | -           | -           | JX282314            |
| 4nF1    | HoxD10a  | 1554     | 1-589      | 590-1324    | 1325-1554   | JX282315            |
| 4nF1    | HoxD10a  | 1495     | 1-592      | 593-1265    | 1266-1495   | JX282316            |

Ψ denotes a pseudogene
In general, polyploidization plays an important role in fish evolution [19]. The purpose of this research was to study the effects of allopolyploidization on Hox gene organization and evolution. In this article, three distinct Hox duplicates were observed in the 4nF1 genome, compared with two copies in RCC and one copy in BSB. Our data reveal the genetic variation and evolutionary characteristics of the Hox gene family in 4nF1 and provide new insights into their evolutionary patterns.

Results

Sequence information for RCC, BSB and 4nF1 clones

Using 11 pairs of degenerate primers (Additional file 1: Table S1), we obtained partial sequence information for eight putative Hox genes from RCC, four putative Hox genes from BSB, and 32 putative Hox genes from the 4nF1. All these fragments were between 700 and 1500 bp long and included the exon1-intron-exon2 region (Table 1). To avoid biased amplification of only one Hox gene copy, we selected 20 clones of each gene from 4nF1, 20 clones of each gene from RCC and 80 clones from BSB (20 clones for each Hox gene PCR fragment). All fragments from RCC, BSB and the 4nF1 were confirmed to be Hox gene sequences, and each included the homeobox. All Hox sequences have been submitted to GenBank; their accession numbers are listed in Table 1.

Molecular organization of the Hox gene sequence

We comparatively analysed the inferred amino acid sequences of the Hox genes in 4nF1 with those in zebrafish, fugu, medaka, and BSB (Additional file 1: Table S2), which indicated that the 4nF1 sequences were similar to those of the other species. The organization of the Hox clusters in 4nF1 is shown in Fig. 1. The clusters can be summarized as HoxAai, HoxAaaii, HoxAaaiii, HoxAbi, HoxAbi, HoxAbii, HoxBai, HoxBaii, HoxBaiii, HoxBbi, HoxBbii, HoxBbiii, HoxCai, HoxCaii, HoxCaiii, HoxCbi, HoxCbi, HoxChii, HoxDaii, HoxDaiii, and HoxDaiii (Table 1). Among these copies, we found that HoxD4aiiiΨ, HoxD9aΨ, and HoxD10aΨ in 4nF1 were pseudogenes (Fig. 2). Two deletions at codons 316 and 317 in the coding region of HoxD4aiiiΨ suggested that it was a pseudogene. The alignment of the putative HoxD4a sequences is shown in Fig. 2a. HoxD9aΨ has become a pseudogene because a stop codon prematurely terminates expression of the full-length functional product (Fig. 2b). An insertion was observed at codon 593 in the HoxD10aΨ coding region; alignment of the putative HoxD10a duplicated sequences is shown in Fig. 2c. HoxD10aΨ had an inserted G nucleotide compared with HoxD10aaiii, whereas a T in HoxD4aΨ was replaced by a G compared with HoxD4ai. Thus, non-functionalization is a possible fate for some duplicated Hox genes. The GC levels of the pseudogenes tended to be lower than that of their counterpart genes (Additional file 1: Table S3). For instance, in 4nF1, the exons of the pseudogene HoxD4aiiiΨ exhibited a GC content of 50.1%, which was lower than that of its functional counterparts HoxD4ai and HoxD4aii (51.3%, 52.1%). As shown in Additional file 1: Table S3, the exon GC content of the pseudogene HoxD10aaiiΨ was 49.4%, which was lower than those of its putative functional counterparts HoxD10aii and HoxD10aiii (49.6% and 49.9%, respectively) in 4nF1. Similarly, the exon GC content of the pseudogene HoxD9aΨ (43.3%) was slightly lower than that of its putative functional HoxB1b paralogues (50.1%, 50.2%, and 50.2%). During duplication, one copy typically remains functional, whereas the other copy may lose its function, which generally leads to a decreased GC level for the non-functional gene.

Phylogenetic relationships

For most genes, such as HoxA4a, HoxB1b, and HoxD10a, three distinct orthologues of the zebrafish genes were identified in 4nF1. These duplicated genes shared a high identity percentage for the deduced amino acid sequences (Additional file 1: Tables S2 and S3). An identity analysis of the putative amino acid sequences suggested that the duplicated sequences were more closely related to each other than to the reported zebrafish orthologues except for the HoxC4aiii sequences. For instance, the percentage nucleotide identity between the HoxA11bi, HoxA11bi, and HoxA11bi orthologues from 4nF1 and HoxA11b from zebrafish was only 89.9%, 89.9%, and 92.4%, respectively. Conversely, the identity between the paralogues HoxA11bi and HoxA11bi,
HoxA11bi and HoxA11bii, and HoxA11bii and HoxA11biii in 4nF1, was 98.6%, 96.4%, and 96.0%, respectively (Additional file 1: Table S2 and Fig. 2a). The identity between HoxB1bi and HoxB1bii, HoxB1bii and HoxB1biii, and HoxB1bii and HoxB1biii was 99.5%, 95.7% and 96.2%, whereas the similarity to their zebrafish orthologs was 91.0%, 90.6% and 91.5% (Additional file 1: Table S2 and Fig. 3b). These results showed that HoxA11bi, HoxA11bii, and HoxA11biii as well as HoxB1bi, HoxB1bii and HoxB1biii all share a mostly closed ancestral cluster and are true orthologues of the zebrafish genes HoxA11b and HoxB1b. Analysis of the sequences obtained for HoxC4a suggested that four distinct copies of this gene exist in 4nF1, which were named HoxC4ai, HoxC4aii, HoxC4aiii and HoxC4a-1. The putative amino acid sequence of HoxC4a-1 shares approximately 100%, 100% and 99% similarity to those of HoxC4ai, HoxC4aii, and HoxC4aiii, respectively. However, the nucleotide
similarity to all three sequences is 100%, which suggests the mutation was synonymous.

To evaluate the speciation of 4nF₁, the nucleotide identity percentages among all known representatives of the HoxA4a, HoxA9a, HoxA2b, and HoxD4a gene groups in RCC, BSB, and the 4nF₁ were examined (Table 2, Fig. 4). The identities of orthologous ‘i’ or ‘ii’ genes between 4nF₁ and RCC were much higher than those between 4nF₁ and BSB. For example, the nucleotide identity percentages of the orthologous HoxA4ai, HoxA9ai, HoxA2bi, and HoxD4ai genes between 4nF₁ and RCC were 99.5%, 99.4%, 99.6% and 99.6%, respectively. Conversely, the similarity of these genes between 4nF₁ and BSB was 97.0%, 92.3.0%, 97.2%, and 93.7%, respectively. Although similarly high identity was observed, the ‘iii’ gene in 4nF₁ did not exhibit higher similarity to the gene in RCC or BSB for all four Hox sequence groups, suggesting no
obvious orthologous relationship between the two species. Thus, we speculated that the ‘ii’ genes were variants of RCC or BSB genes. For example, the *HoxA4*ii, *HoxA9*ii, *HoxA2b*ii, and *HoxD4a*ii genes from 4nF1 and the *HoxA4a*, *HoxA9a*, *HoxA2b*, and *HoxD4a* genes from BSB shared 98.0%, 92.2%, 97.7%, and 94.0% identity (Table 2).

**Discussion**

**The structure of cloned Hox gene sequences**

Prior PCR surveys and genomic library screening have identified interesting variability in *Hox* gene content among teleosts [12, 15, 16, 20, 21]. Luo et al. [22] estimated there were 14–16 *Hox* gene clusters in goldfish. Our data suggested 18–21 *Hox* gene clusters were present in 4nF1, with each was located on a different acrocentric chromosome. The *Hox* gene clusters in 4nF1 were approximately the sum of the clusters in RCC and BSB, except that some clusters were lost. The topology of the *Hox* gene maximum likelihood tree (Fig. 3) further suggested that some of the *Hox* genes orthologous to zebrafish genes were present in two copies in RCC, one copy in BSB, and three copies in 4nF1. However, the third copy did not exhibit notably higher similarity to the gene in RCC or BSB. We speculated that variation and reorganization of the genome likely occurred during polypliodization, resulting in new copies in 4nF1. This might be evidence that allopolyploidization induces a variety of rapid genomic changes in a 4nF1 population [23, 24]. Using sequence alignment in 4nF1, we isolated 32 fragments that can be characterized as *HoxA*, *HoxB*, *HoxC*, and *HoxD* family genes. However, amplified RCC and BSB DNA were only characterized as the *HoxA* and *HoxD* genes. We speculated that the increase in the number of 4nF1 genes might be related to polypliodization. This situation was also observed in our previous study [25, 26]; the number of 4nF1 fragments increased, and some genes from RCC and BSB were lost. At present, although we have no precise data explaining this outcome, we speculate that allotetraploidization might lead to rapid changes in 4nF1 genome diversity. Our study is the first to evaluate the organization of *Hox* clusters in a 4nF1 population. This theory is also strongly supported by other studies examining *Hox* genes in other gene families [27], and DNA content [28].

**The significance of polypliodization**

Polypliodization likely increases genomic variation rates, which can result in the formation of new polyploid species [29]. First, the process of polypliodization can itself generate species that are reproductively isolated from their diploid progenitors, increasing the number of species as a by product. For example, several studies have indicated that a polypliodization event occurred in an ancestor of teleost fish shortly after this lineage diverged from the lineage leading to tetrapods [30–32]. Second, an entirely different trait can result in increased rates of polypliodization [6]. Synonymous mutations increase genomic variation. For example, the putative amino acid sequence of *HoxC4*-1 shares approximately 100%, 100%, and 99% similarity with those of *HoxC4a*, *HoxC4b*, and *HoxC4c*, respectively. The identity of their nucleotide sequences is 100%. In the polypliodization process, genome duplication produces abundant genomic DNA, so the organism maintains the dosage balance or rapidly stabilizes the duplicated genomes via retention/exclusion of redundancy. Lynch et al. [33] suggested there are three outcomes in the evolution of duplicate genes: non-functionalization, neo-functionalization and sub-functionalization. Interestingly, we found some pseudo-genes in 4nF1, such as *HoxD4a*ii, *HoxD9a*ii, and *HoxD10a*ii. Pseudo-genes are formed either by random mutations that create stop codons and prematurely terminate full-length functional product expression or by insertions/deletions that shift the reading frame, rendering the translated protein non-functional. We speculate that dosage effects generated selection pressure from the loss of *Hox* genes or the formation of pseudo-genes after whole genome duplication. This pressure is consistent with the expectation that there are *Hox* clusters in the 4nF1 genome that have lost functional *Hox* genes due to the reduction of redundancy following the polypliodization event. However, 4nF1 required genetic recombination, mutation, and pseudogenization to reduce the amount of incompatible genetic material and improve fertility [34]. Thus, we unexpectedly

### Table 2

|        | HoxA4a (%) | HoxA9a (%) | HoxA2b (%) | HoxD4a (%) |
|--------|------------|------------|------------|------------|
| 4nF1   | 97.4/97.5  | 91.8/89.5  | 96.9/96.9  | 98.3/97.7  |
| 4nF1   | 98.5/98.0  | 98.9/98.9  | 97.5/97.6  | 96.2/61.6  |
| RCC i  | 99.5/99.5  | 94.9/99.4  | 96.9/100.0 | 99.6/100.0 |
| RCC ii | 97.2/96.6  | 92.3/90.1  | 96.5/96.6  | 98.3/97.7  |
| BSB    | 90.7/96.6  | 92.3/91.6  | 97.2/97.9  | 93.7/94.9  |
| 4nF1   | 98.3/98.5  | 91.6/88.5  | 96.9/95.9  | 95.3/60.5  |
| RCC i  | 97.5/98.0  | 92.0/89.0  | 96.8/96.9  | 98.3/97.7  |
| RCC ii | 98.8/98.5  | 94.9/92.1  | 98.3/97.6  | 100.0/100.0|
| BSB    | 96.7/97.0  | 91.0/86.4  | 96.4/96.3  | 93.3/92.7  |
| 4nF1   | 98.7/98.5  | 99.1/98.4  | 97.4/97.6  | 95.9/61.6  |
| RCC i  | 98.2/98.5  | 92.2/89.0  | 96.5/95.9  | 95.3/60.5  |
| RCC ii | 98.0/98.5  | 92.2/90.6  | 97.7/97.3  | 94.0/59.4  |
| BSB    | 97.2/97.0  | 92.5/91.1  | 97.1/97.9  | 94.0/94.9  |
| RCC i:BSB| 96.9/97.0  | 91.5/88.0  | 96.2/96.6  | 93.3/92.7  |

Values before slashes (/) denote nucleotide identity, values after slashes denote amino acid identity.
obtained autotetraploids with greater fertility among the 4nF1 progeny, and we successfully established an autotetraploid fish line [35]. Our characterization of the Hox gene clusters in tetraploid hybrids improves our understanding of the evolutionary processes occurring after Hox gene duplication in vertebrates.

**Conclusions**

We identified three copies of Hox genes in 4nF1, two copies in RCC and one copy in BSB. In addition, obvious variation and pseudogene generation were observed in some 4nF1 Hox genes. These results reveal the effects of polyploidization on the organization and evolution of Hox gene clusters in fish and also help to clarify aspects of vertebrate genome evolution.

**Methods**

**DNA extraction**

Specimens of 4nF1 (4n = 148), RCC (2n = 100), and BSB (2n = 48) were obtained from the Engineering Research Center of Polyploid Fish Breeding and Reproduction of the State Education Ministry at Hunan Normal University.
Fish treatments were carried out according to the regulations for protected wildlife and the Administration of Affairs Concerning Animal Experimentation, and approved by the Science and Technology Bureau of China. Approval from the Department of Wildlife Administration was not required for the experiments conducted in this paper. The fish were deeply anesthetized with 100 mg/L MS-222 (Sigma-Aldrich, St Louis, MO, USA) before dissection. Narcotic drugs were fed before blood sampling. Total genomic DNA was isolated from peripheral blood cells using the standard phenol chloroform extraction procedures described by Sambrook et al. [36].

Cloning and sequencing of Hox genes
We amplified fragments of Hox genes from genomic DNA by PCR amplification using several combinations of degenerate primers (Table 1). PCR was performed in 50-μL reaction volumes using Taq DNA polymerase (TaKaRa, Dalian, China). A typical PCR programme consisted of a denaturation step at 94°C for 5 min; 35 cycles of 98°C for 15 s, 55°C for 45 s and 72°C for 1 min; and a final elongation step at 72°C for 10 min. PCR products were cloned into a T vector and sequenced with an automated ABI 3700 DNA sequencer. The sequences were BLAST searched against the non-redundant protein database maintained at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) to determine their identity.

Sequence comparison and analysis
Sequence homology and variation among the fragments amplified from RCC, BSB and the 4nF1 were analysed in BioEdit [37, 38]. Partial DNA sequences for each gene were verified using a BLASTx search. To increase the probability of detecting duplicated paralogues and circumventing PCR errors, we sequenced 20 clones for each gene from 4nF1, RCC and BSB. The obtained sequences were screened for Hox gene fragments using different BLAST searches (BLASTn, BLASTp, and BLASTx) against GenBank (http://www.ncbi.nlm.nih.gov/Blast.cgi). Then, we evaluated the organization of the 4nF1 Hox clusters compared to RCC and BSB to characterize the Hox genes.

Phylogenetic analysis
Using Clustal X 1.81, the derived amino acid sequences of these fragments were aligned with the Hox genes from BSB, zebrafish, fugu, medaka and other teleosts retrieved from GenBank [38]. Regions of sequences that were difficult to align were removed from the alignment. Gaps were also removed from the alignment. The maximum likelihood method implemented in the online software RAxML was used to construct a phylogenetic tree [39].

Additional file

Additional file 1: Table S1. The PCR Gene-specific degenerate primers. Gene-specific degenerate primers designed based on the alignment and identification of consensus orthologous Hox gene sequences from zebrafish (Danio rerio), medaka (Oryzias latipes), pufferfish (Fugu rubripes), mouse (Mus musculus), cichlids, and humans (Homo sapiens). Table S2. The Percentage of the amino acid. Percentage amino acid identity between paralogous Hox sequences obtained from 4nF1 and reported orthologues from zebrafish, fugu, and medaka. Table S3. Comparison of GC levels among duplicated genes. (DOCX 20 kb)

Abbreviations
4nF1: Allotetraploid F1 hybrids; BSB: Blunt snout bream; RCC: Red crucian carp

Acknowledgements
We would like to sincerely thank many researchers who help to complete this manuscript, including Drs. Yao Zhanzhou and Zhao Rurong.

Funding
This research was supported by National Natural Science Foundation of China Grants 31210103918, 31430088 and 31730098, the earmarked fund for China Agriculture Research System (CARS-45), the Natural Science Foundation of Hunan Province for Distinguished Young Scholars (Grants No. 2017JJ1022), the Major Program of Educational Commission of Hunan Province (Grant No. 17A133), the Cooperative Innovation Center of Engineering and New Products for Developmental Biology of Hunan Province (20134486), the Construction Project of Key Discipline of Hunan Province and China, Hunan Normal University Graduate Research Innovation Project (CG2017B184).

Availability of data and materials
The dataset(s) supporting the conclusions of this article is (are) available in the GenBank repository with access No. JX282274- JX282316 and hyperlink to dataset(s) in https://www.ncbi.nlm.nih.gov/nuccore/JX282274-JX282316. The dataset(s) supporting the conclusions of this article is (are) included within the article (and its additional file(s)).

Authors’ contributions
YW, QQ and SL designed the study, performed the analyses, performed the technical discussions, and prepared and drafted the manuscript. WS, QL, XH and YH participated in data simulation and discussions. RY, TL, CZ and MT were involved in the statistical analysis. All authors have read and approved the final manuscript.

Ethics approval
All experiments were approved by the Animal Care Committee of Hunan Normal University and followed the guidelines statement of the Administration of Affairs Concerning Animal Experimentation of China. All samples were raised in natural ponds, all dissections were performed under sodium pentobarbital anaesthesia, and all efforts were made to minimize suffering.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

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

Received: 20 May 2017 Accepted: 13 September 2017
Published online: 11 October 2017

References
1. Ohno S, Wolf U, Atkin NB. Evolution from fish to mammals by gene duplication. Heredity. 1968;59(1):169–87.
2. Masterson J. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science. 1994;264(5157):421–4.

3. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature. 1997;387(6634):708–13.

4. Conrail L. Genetic and epigenetic interactions in allopolyploid plants. Plant molecular biology. 2000;43(2-3):387–99.

5. Liu S, Liu Y, Zhou G, Zhang X, Luo C, Feng H, He X, Zhu G, Yang H. The formation of tetraploid stocks of red crucian carp common carp hybrids as an effect of interspecific hybridization. Aquaculture. 2001;192(2):171–86.

6. Otto SP. The evolutionary consequences of polyploidy. Cell. 2007;131(3):452–62.

7. Mallar J. Hybrid speciation. Nature. 2007;446(7133):279–83.

8. Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJ, Bierne N, Boughman J, Brooksford A, Buerkle CA, Buggs R, et al. Hybridization and speciation. Journal of evolutionary biology. 2011;25(2):229–46.

9. Liu S, Qin Q, Xiao J, Lu W, Shen J, Li W, Liu J, Duan W, Zhang C, Tao M, et al. The formation of the polyploid hybrids from different subfamily fish crossing and its evolutionary significance. Genetics. 2007;176

10. Qin Q, Wang Y, Wang J, Dai J, Liu Y, Liu S. Abnormal chromosome behavior during meiosis in the allotetraploid of Carassius auratus red var. (female symbol) × Megalobrama amblycephala (male symbol). BMC genetics. 2014;15:59.

11. Yu X-J, Zhou T, Li Y-C, Li K, Zhou M. Chromosomes of Chinese fresh-water fishes. In: Beijing (in Chinese); Science Publishing House, 1989.

12. Santini S, Bernardi G. Organization and base composition of telipha Hox genes: implications for the evolution of Hox clusters in fish. Gene. 2005;346:51–61.

13. Koh EG, Lam K, Christoffels A, Erdmann MV, Brenner S, Venkatesh B. Hox gene clusters in the Indonesian calecanth, Latimeria menadoensis. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(3):1084–8.

14. Mallo M, Alonso CR. The regulation of Hox gene expression during animal development. Development. 2013;140(19):3951–63.

15. Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL, et al. Zebrafish hox clusters and vertebrate genome evolution. Science. 1998;282(5394):1711–4.

16. Amores A, Suzuki T, Yan YL, Pomeroy J, Sanger A, Amemiya C, Postlethwait JH. Developmental roles of pufferfish Hox clusters and genome evolution in ray-finned fish. Genome Res. 2004;14(1):1–10.

17. Chiu CH, Amemiya C, Dewar K, Kim CB, Ruddle FH, Wagner GP. Molecular evolution of the HoxA cluster in the three major gnathostome lineages. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(8):5492–7.

18. Chiu CH, Dewar K, Wagner GP, Takashashi K, Ruddle F, Ledje C, Bartsch P, Scemama JL, Stellwag E, Fried C, et al. Bichir HoxA cluster sequence reveals surprising trends in ray-finned fish genomic evolution. Genome Res. 2004;14(11):7–17.

19. Neuroser J, Stadler PF, He S, Meyer A. PCR survey of hox genes in the goldfish Carassius auratus auratus. Journal of experimental zoology Part B, Molecular and developmental evolution. 2007;308(3):250–8.

20. Qin Q, Liu Z, Cao L, Xiao Q, Wang Y, Liu S. Rapid genomic changes in allopolyploids of Carassius auratus red var. (female symbol) × Megalobrama amblycephala (male symbol). Scientific reports. 2016;6:34417.

21. Stephens SC. Possible Significance of Duplication in Evolution. Advances in Genetics. 1951;4:247–65.

22. Qin Q, He W, Liu S, Wang J, Xiao J, Liu Y. Analysis of SS rDNA organization and variation in polyploid hybrids from crosses of different fish subfamilies. J Exp Zool B Mol Dev Evol. 2010;314

23. Chen L, Li W, Liu S, Tao M, Long Y, Duan W, Zhang C, Xiao J, Qin Q, Luo K, et al. Novel genetic markers derived from the DNA fragments of Sox genes. Molecular and cellular probes. 2009;23(3-4):157–65.

24. Malanga-Trillo E, Laessing U, Lang DM, Meyer A, Steuerman CA. Evolution of duplicated reggie genes in zebrafish and goldfish. Journal of molecular evolution. 2002;54(2):235–45.

25. Ohno S, Muramoto J, Chistian L, Atkin NB. Diploid-tetraploid relationship among old-world members of the fish family Cypriinidae. Chromosoma. 1967;23(1):1–9.

26. Conrail L. The advantages and disadvantages of being polyploid. Nat Rev. Genet. 2005;6(11):836–46.

27. Meyer A, Scharf M. Gene and genome duplications in vertebrates: the one-to-four (to-eight in fish) rule and the evolution of novel gene functions. Current opinion in cell biology. 1999;11(6):699–704.

28. Panopoulos G, Hennig S, Grotth D, Krause P, Postkau AJ, Herwig R, Vingron M, Léhrach H. New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. Genome Res. 2003;13(6A):1056–66.

29. Seoighe C, Wolfe KH. Extent of genomic rearrangement after genome duplication in yeast. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(8):4447–52.

30. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290(5504):1151–5.

31. Qin Q, He W, Liu S, Wang J, Xiao J, Liu Y. Analysis of SS rDNA organization and variation in polyploid hybrids from crosses of different fish subfamilies. Journal of experimental zoology Part B, Molecular and developmental evolution. 2010;314(5):403–11.

32. Qin Q, Wang Y, Wang J, Dai J, Xiao J, Hu F, Luo K, Tao M, Zhang C, Liu Y, et al. The autotetraploid fish derived from hybridization of Carassius auratus red var. (female) × Megalobrama amblycephala (male). Biology of reproduction. 2014;91(4):5.

33. J S, EF F, T M. Molecular cloning: a laboratory manual. 2nd ed, 1989.

34. Wall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic Acids Symp Ser. 1999:95–8.

35. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. Nucleic acids research. 1997;25(4):4876–82.

36. Stamatakis A, Hoover P Fau - Rougemont J, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. (1076-836X Electronic).