Reproduction and Breeding 3 (2023) 82–87

Contents lists available at ScienceDirect

Reproduction and Breeding

journal homepage: www.keaipublishing.com/en/journals/reproduction-and-breeding

Full-length transcriptome reveals rapid genetic changes in triploid hybrid grass carp derived from female grass carp × male topmouth culter

Xu Huang [1], Jiawang Huang [1], Lang Qin [1], Yali Long [1], Chang Wu, Kaijun Gong, Qingwen Xiao, Qinbo Qin [*]

State Key Laboratory of Developmental Biology of Freshwater Fish, Engineering Research Center of Polyploid Fish Reproduction and Breeding of the State Education Ministry, College of Life Sciences, Hunan Normal University, Changsha, 410081, Hunan, China

ARTICLE INFO

Keywords:
Distant hybridization
Triploid hybrids
Full-length transcriptome
Fusion gene
Alternative splicing

ABSTRACT

Distant hybridization refers to the combination of different genomes from different species and further leads to genetic structure changes in hybrids. For next-generation sequencing limitations, the genetic changes with respect to hybrids or polyploids cannot be fully explained. Here, we obtained 117,526 full-length isoforms from a triploid hybrid grass carp (3nGT, 3n = 72) by the crossing of female grass carp (GC, 2n = 48) and male topmouth culter (TC, 2n = 48) using single-molecule long reads sequencing technology. Further, the gene fusion, alternative splicing (AS), transcription of homologous genes, and novel genes were analyzed. The results showed that 1563 fusion genes were detected in 3nGT including 872 heterogenous fusion genes derived from GC and TC genomes. Meanwhile, 175,644 AS events were identified and divided into five models, of which IntronR was the dominant model. In addition, 8964 novel genes were discovered, of which 2802 novel genes were annotated. Finally, scans results of *Hox* gene family revealed that both GC and TC *Hox* genes were transcribed in 3nGT, but presented asymmetric transcription patterns with a bias towards the maternal parent GC genome. This study not only provides a rich resource of transcript isoforms for the superior triploid grass carp, but also provides an important insight into rapid genetic changes in hybrid fish.

1. Introduction

Hybridization means the combination of different genome from different species, and the interaction of heterogenous genome often cause genome or gene structure changes [1–5]. Those genetic alterations, determining phenotype changes directly, were the foundation of hybrid breeding. In hybrids, obvious gene structure alterations had been reported in heterogenous diploids and polyploids [6,7]. In the allotetraploids of female *Carassius auratus* red var. × male *Cyprinus carpio* L, chimeric genes were distributed in all kinds of signal paths, which regulated the normal life activities of polyploid [8]. Similar chimeric genes were also found in allodiploids of female *Carassius auratus cuvieri* × male C. *auratus* red var [9].

The next-generation sequencing technology has developed fast in the past few years with superior technical performance and cost advantages to the Sanger sequencing [10]. However, for the limitation of short reads, it is not easy to assemble complete mRNA or cDNA in large quantities [11]. The PacBio Isoform sequencing overcame this obstacle well. At present, with this single-molecule long-read sequencing platform, outstanding performance has been achieved in transcriptome sequencing of sorghum [12], maize [13], and red clover [14]. Moreover, the long-read sequencing platform played an important role in exploring the genetic or other biological characteristics in hybrids [15] and polyploids [16–18].

In the present study, a triploid hybrid grass carp (3nGT, 3n = 72) was produced by mating female grass carp (GC, 2n = 48) and male topmouth culter (TC, 2n = 48), which contained one set of chromosomes from TC and two sets of chromosomes from GC, and their basic biological characteristics were surveyed, which presented obvious heterosis [19]. In this study, the genetic characteristic regarding the formation of a novel gene, gene fusion, and alternative splicing in 3nGT was further analyzed by the full-length transcriptome sequencing. This study provides a rich resource of transcript isoforms for the superior triploid grass carp, which will provide important reference for the analyses of dominant characters

* Corresponding author. State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences, Hunan Normal University, Changsha, 410081, Hunan, China.
E-mail address: qqb@hunnu.edu.cn (Q. Qin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.repbre.2023.06.001
Received 26 May 2023; Received in revised form 7 June 2023; Accepted 8 June 2023
2667-0712/© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

and important guidance for the hybrid breeding of fish.

2. Methods

2.1. Ethics statement

The Administration of Affairs Concerning Animal Experimentation Guidelines states 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 not rare or near extinction (first-class or second-class state protection level). Therefore, approval was not required for the experiments conducted in this study.

2.2. Animal materials

The 3nGT was produced by the crossing of female GC and male TC, and all the fish were cultured in the Engineering Center of Polyploidy Fish Breeding of the National Education Ministry, Hunan Normal University, Hunan, China. The sample was deeply anaesthetized with 100 mg/L MS-222 (Sigma-Aldrich, St Louis, MO, USA) for 10 min (25 °C) in a separate tank prior to dissection. Five tissue materials including the heart, liver, kidney, brain, and muscle were excised carefully, flash frozen by liquid nitrogen and subsequently stored at -80 °C.

2.3. RNA extraction, library construction and sequencing

The Qiagen kit was used to extract high quality RNA from 3nGT tissue samples according to the manufacturer's protocol. All RNA samples were treated with RNase-free DNase I. The RNA purity and integrity were determined by a NanoDrop® 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Then, high-quality RNA was reverse-transcribed into cDNA using SMARTer® PCR cDNA Synthesis Kit. Mix cDNA samples from the heart, liver, kidney, brain, and muscle were used to build three 0.5–6 kb libraries through PCR amplification using KAPA HiFi PCR Kits. Finally, those libraries were sequenced by single-molecule sequencing technology.

2.4. RNA-seq data analysis and calibration

Firstly, the adaptors and low-quality sequences were removed before producing raw data. For the error rate of the Pacbio sequence, the production of high-quality cDNA sequence data was further performed with the quality control and correction. After the quality control, the clean polymerase reads were processed to separate insert reads with full passes >1 and accuracy >0.80. According to the position of 5'primer, 3'primer and polyA, the filtered reads were classified were divided into full-length non-chimeric or non-full-length reads. In order to improve the accuracy of isoforms, redundancy full-length non-chimeric reads were cultured to unpolished consensus by ICE package (https://github.com/PacificBiosciences/IsoSeq_SA3nUP/wiki/RunningIso-Seq-Cluster). Finally, an arrow was used to correct unpolished consensus and obtain polished high-quality consensus sequences, whose accuracy was more than 99%.

2.5. Fusion gene analysis and validation

Fusion gene refers to part or all sequences from two different genes integrated into one new chimeric gene. To detect the fusion genes, the high-quality isoforms were mapped to a reference genome of GC and TC with gmap software and extracted the mapping results using fusion_finder.py (https://github.com/Magdoll/cDNA_Cupcake). The fusion gene reads must map to two or more positions, whose distance was more than 10 kb. Besides, the coverage of the transcript corresponding to each site must reach at least 10%, and the coverage of the mapping area must be more than 99%. Further, we verified the fusion gene events by Sanger sequencing. Reference sequences of fusion gene were blasted and extracted from the reference genome of GC and TC using the software of Tbtools [20]. Primers were designed by Primer Premier 5 and sequencing results were analyzed by BioEdit.

2.6. Gene annotation, ncRNA prediction and identification of novel gene

Firstly, according to the mapping information of isoforms to the genome of GC and TC, the annotated isoforms were extracted by matchAnnot software. For the next, another new isoform, which was mapped to the genome but without annotated results, was processed to identify noncoding RNAs (ncRNAs) using ncrna_pipeline (https://bitbucket.org/arrigonialberto/lncrnas-pipeline). Lastly, those novel isoforms, which were mapped to the genome without annotated results and were not ncRNAs, were identified as novel genes. Those novel genes were reannotated to gene ontology (GO), Kyoto Encyclopedia genome (KEGG), NCBI non-redundant protein sequence (NR), Swiss- Prot, and sequence of orthologous protein clusters (COG), respectively.

2.7. Analyses and identification of alternative splicing

Alternative splicing (AS) is a process in that pre-mRNA produced different mRNAs by different splicing ways or different splicing sites. It is an important mechanism that regulates gene expression and increases proteomic diversity [21]. To detect alternative splicing (AS) events in 3nGT, all transcripts were mapped to GC and TC genomes respectively. We extracted transcripts mapped to the same gene as an AS event, and divided the AS events into the following models: intron retention (IntronR), alternative donor site (AltD), alternative acceptor site (AltA), exon skipping (ExonS), alternative position (AltP) and other. Lastly, the top 5% of AS events were enriched to the GO and KEGG database, respectively.

2.8. Analyses of Hox gene family

To survey the homologous gene transcriptional pattern in triploid hybrids, the *Hox* gene family was determined in GC and TC genomes by the tblastn software ($E = 2e^{-5}$), using the *Hox* gene sequences from *Danio rerio* as the query sequences. The cds sequences of *Hox* gene family were extracted from GC and TC genome respectively, which were used as the query sequences for the *Hox* genes selected in the full-length transcriptome of 3nGT. Next, the *Hox* genes were mapped to full-length transcripts of 3nGT using the blast software and extracted the mapped isoforms. Further, those gene from GC and TC genomes were aligned with corresponding isoforms respectively by BioEdit software. The discrepant single nucleotide polymorphism (SNP) sites in the homologous fragment of each gene were selected, counted, and analyzed, to distinguish the different genomes and survey the SNP sites recombination in 3nGT.

3. Results

3.1. The full-length sequences of 3nGT using PacBio sequencing

Using the PacBio-Sequel platform, full-length transcriptome sequencing was performed on mixed high-quality RNA samples from five tissues of 3nGT including heart, liver, kidney, brain, and muscle. As a result, the subreads with a total of 18, 990, 518,290 bases were obtained from the three SMRT cells after removing the adaptor and low-quality reads. By applying the standard Iso-Seq classification and clustering protocol, 675,390 insert reads with an average length of 2241.36 bases were further divided into 86,736 non-full-length reads and 571,493 full-length non-chimeric reads. After ICE clustering and correction, we finally obtained 356,574 consensus sequences, of which 183,481 consensus sequences with an accuracy of more than 99%. The datasets supporting the conclusions of this article were available in the

GenBank repository with access No. PRJNA980910.

3.2. Fusion gene analysis

After high-quality isoforms mapping to the GC-TC genome, 1563 fusion genes were identified and divided into three types. The first type including 427 sequences (27.32%) showed a gene fusion of different chromosomes or scaffolds in the GC genome; the second type including 264 sequences (16.89%) showed a gene fusion of different chromosomes or scaffolds in the TC genome; the third type including 872 sequences (55.79%) showed a gene fusion between GC genome and TC genome. The gene fusion in GC or TC genome may occur before the hybridization, however, the fusion genes between different subgenomes were formed following the hybridization. In this case, we verified the 3 fusion genes randomly from the third type by Sanger sequencing. As a result, the PBfusion.557 transcript combined the creatine kinase, M-type (CKM) gene of TC, and skeletal muscle alpha-actin (acta1) gene of GC (Fig. 1).

3.3. Gene annotation, LncRNA prediction and novel gene discorvery

As shown in Table 1, 355,269 transcripts were obtained after removing fusion genes, and 353,538 transcripts were mapped to the reference genome by GMAP software, accounting for 99.51% of the total transcripts. According to the mapped transcripts, collapse_isoforms_by_sam.py was used to remove the redundancy and got 117,526 isoforms. Then, matchAnnot software was used to compare and extract the mapping results with the annotation information of reference genomes. There were 102,908 isoforms mapped to 31,247 known genes, and 14,618 isoforms mapped to the area without annotation information in reference genomes. A total of 5654 LncRNAs were predicted from 14,618 unannotated isoforms by ncrna_pipeline prediction, in which the average sequence length was 1649 bp with a minimum length of 201 bp

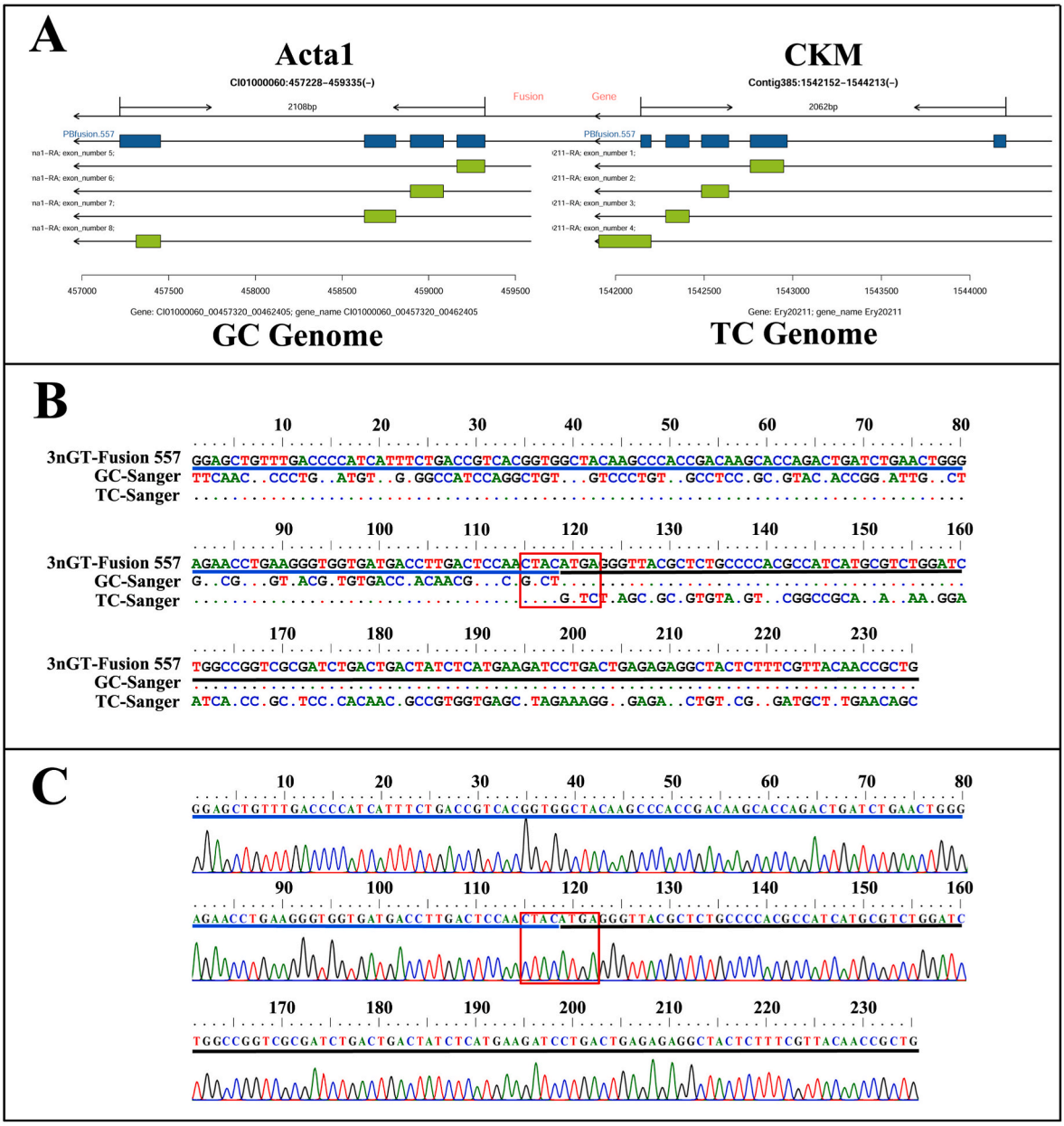

Fig. 1. Structural model and verification of PBfusion.557 (CKM-acta1). (A) The mapping result and diagram of PBfusion.557. (B) The nucleotide sequence alignment of PBfusion.557. (C) Sanger sequencing peak map of PBfusion.557. Red boxes mean fusion site of PBfusion.557. (For interpretation of the references to color in this figure legend the reader is referred to the web version of this article).

Table 1
Summary of mapping and annotation results of isoforms.

Summary of mapping result	Total mapped number	Total mapped rate	Number of collapse_isoforms	Isoforms mapped to known gene	Total mapped gene	Isoforms mapped to unknown gene	Number of ncRNAs	Number of novel isoforms
355,269	353,538	99.51%	117,526	102,908	31,247	14,618	5654	8964

and a maximum length of 6728 bp. Lastly, 8964 novel isoforms, which were mapped to reference genomes but neither annotated information nor ncRNAs, were identified as novel genes. For 8964 novel isoforms, multiple databases, including COG, KEGG, NR, SwissProt, and GO, were used to annotate new genes. A total of 2802 genes (31.26%) were annotated by those databases (Fig. 2). In addition, the function of the annotated novel genes was analyzed by the GO and KEGG database, in which the novel genes in GO database were clustered in the process of protein binding, DNA binding, DNA integration, and they were clustered in the process of signal transduction, endocrine system, cellular community, and immune system in KEGG database.

3.4. Alternative splicing analysis

There were 102,908 isoforms mapped to 31,247 known genes, indicating that some different isoforms mapped to the same gene sequence in the genome. Those different isoforms were extracted to be an AS event. In this study, AS events were identified and divided into five models (Table 2). The models of ExonS contained 12,045 AS events and involved 6613 genes; the models of AltD contained 7478 AS events and involved 4276 genes; the models of AltA contained 7726 AS events and involved 4459 genes; the models of IntronR contained 29,590 AS events and involved 9603 genes; the models of AltP contained 22,948 AS events and involved 6311 genes; and 95,857 AS events were divided into Other, which contained 8002 genes. The IntronR and AltP were the main AS models. In addition, we also performed GO enrichment and KEGG enrichment on the top 5% genes, in which tight junction, insulin signaling pathway, and AMPK signaling pathway were main pathways in KEGG, while molecular function and binding were the main GO term in GO database.

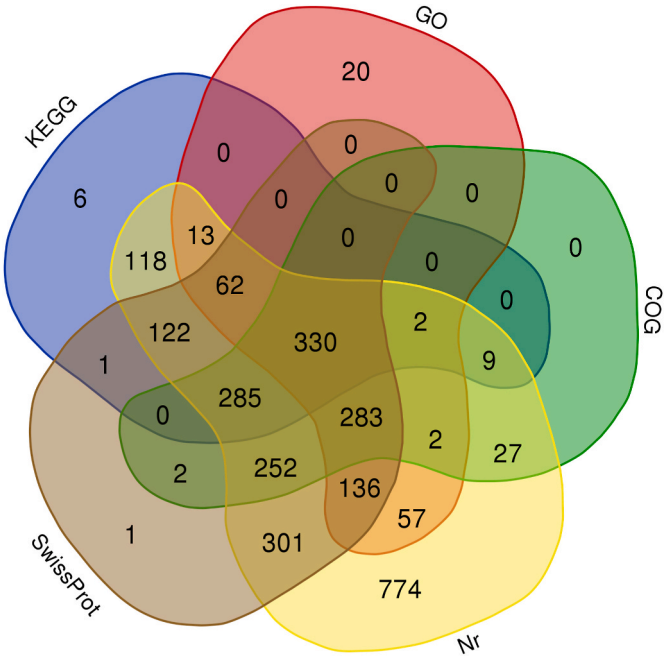

Fig. 2. Novel genes annotated in COG, KEGG, NR, SwissProt and GO database.

Table 2
Types and numbers of alternative splicing (AS).

Type	AS_Number	Gene_Number
ExonS	12,045	6613
AltD	7478	4276
AltA	7726	4459
IntronR	29,590	9603
Other	95,857	8002
AltP	22,948	6311

3.5. Analyses of Hox gene family

After homologous gene alignment, 48 *Hox* genes were screened and extracted in both GC and TC genomes. A total of 18 *Hox* genes, including 28 isoforms, were founded and extracted from the 3nGT full-length sequences (Fig. 3). After aligning each gene with homologous isoforms, 4 gene transcription patterns were determined in 3nGT according to the species-specific SNPs of GC and TC. The first pattern is that both GC and TC *Hox* genes (*HoxA3a*, *HoxB3a*, and *HoxB8a*) were transcribed with the form of alternative splicing. The second pattern is that only GC *Hox* genes were found in 3nGT, including *HoxA4a*, *HoxB5a*, *HoxB9a*, *HoxC5a*, *HoxC8a*, *HoxC9a*, *HoxC10a*, *HoxD3a*, and *HoxD9a*. The third pattern is that only TC *Hox* genes were found in 3nGT, including *HoxB2a*, *HoxB7a*, *HoxB10a*, and *HoxC3a*. The fourth pattern is chimeric genes, such as *HoxD10a* involving 1 isoform, which had GC and TC species-specific SNPs in the same isoform of 3nGT. In addition, except for *HoxB4a* and *HoxD10a*, the remaining 25 isoforms were divided into 16 isoforms (64%) from GC and 9 isoforms (36%) from TC, which was consistent with the ratio of GC and TC chromosome numbers in 3nGT. Interestingly, only the subtypes of "a" were transcribed in 3nGT. For example, we found that only *HoxB5a* and *HoxB8a* were transcribed, but did not find *HoxB8b* or *HoxB5b* in 3nGT.

4. Discussion

In nature, distant hybridization and polyploidization were the important ways to produce new hybrids and polyploids, which accelerated the speciation and evolution [22,23]. Compared with autopolyploids, the allopolyploids refer to the combination of different subgenome, and following experience the interaction of subgenome [24,25]. However, it is rare in animals compared with that in plants [26,27]. In our previous study, a triploid hybrid from female GC × male TC was produced, and their biological characteristics were also been documented [19]. Polyploidization increases the complexity of the transcriptome, which refers to the interaction of subgenome [28], and the next-generation with the short reads has an obstacle in surveying the complete transcripts. In this study, 117,526 full-length transcripts of 3nGT were obtained from the Pacbio platform to analyze the genetic characteristics such as gene fusion, AS events, LncRNAs, novel genes, and *Hox* gene family in triploid hybrids.

Hybridization may result in the combination of subgenomes and rearrangement of chromosomes [29], accelerating the formation of the fusion gene. However, the fusion gene was rarely reported in hybrid species, but usually reported in humans, especially in certain diseases, such as the *FUS-CREB3L1* fusion gene in low-grade fibromyxoid sarcoma [30], and *EML4-ALK* fusion gene in non-small-cell lung cancer [31]. Following the full-length transcripts sequencing and mapping to

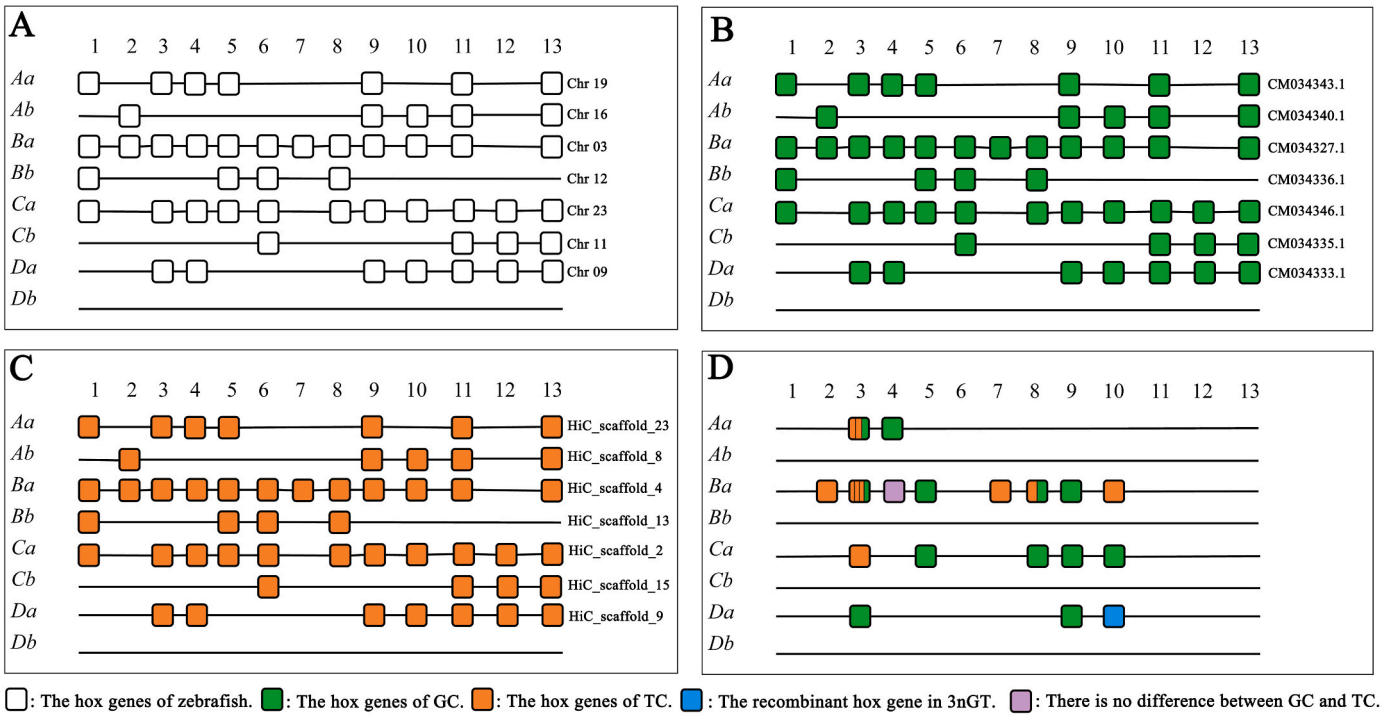

Fig. 3. Location of *Hox* gene family in Zebrafish (A), GC (B), TC (C) genome and the transcription in 3nGT (D).

reference genomes, 872 species-species heterologous fusion isoforms were identified (0.74%) in 3nGT. The reported study held the opinion that the fusion genes were formed by the insertion mechanism [32]. The fragments may be randomly inserted into 3nGT genome. However, it may be a common phenomenon in hybrids following hybridization. The formation of heterologous fusion genes may format new functions and corresponding regulatory mechanisms, which may another way to explore the production of heterosis. Although the function of heterologous fusion genes was not clear, it will potentially accelerate the fusion of heterologous genome in hybrids.

The single nucleotide polymorphism (SNP) is an effective molecular marker to distinguish different species and even detect homologous gene recombination or gene chimerism in hybrids [33–35]. *Hox* genes are a group of important developmental regulatory genes with highly conserved and typical Cluster structure, involving a variety of life activities [36,37]. According to the chromosome distribution of *Hox* gene, the subtypes of "a" and "b" were in homologous chromosomes, but only the subtypes of "a" were transcribed and detected in 3nGT, which means that only one set of chromosomes of GC was transcribed or expressed for *Hox* family. The potential reason is that the existence of two identical sets of chromosomes in the nucleus would induce dosage compensation [38]. However, about one-third of the transcripts originated from TC and two-thirds from GC according to the species-SNPs in 3nGT, which was consistent with the ratio of GC and TC chromosome numbers (48:24). Similarly, genome-wide expression level dominance was biased toward maternal genome in the triploids from female GC × male blunt snout bream [38]. Although only one set of chromosomes of GC was transcribed or expressed for *Hox* family, the transcripts of triploid were biased toward maternal genome with the form of AS events. The AS events may play an important role in hybrids to maintain the stability of the genome and normal life activities. In addition, SNPs recombination is another important phenomenon in hybrids. The recombined SNP sites in allotetraploid of C. *auratus* red var. (♀) × *Cyprinus carpio* L. (♂) related to mutagenesis, repair, and cancer-related pathways, which played important roles in the regulation of different processes [8]. In this way, we suspected that the SNP sites recombination between homologous gene was another way to accelerate the fusion of heterologous genome in hybrids.

In summary, we obtained 117,526 full-length transcripts from a triploid hybrid using the PacBio-Iso-seq platform, and further analyzed the gene fusion, AS events, LncRNAs, novel genes, and the transcription of *Hox* gene family in 3nGT. In addition, the heterologous gene fusion and homologous gene transcription provided important insights into the interaction of heterologous subgenomes, especially in the hybridization and formation of heterogeneous species. Those data provided an important reference for the regulation of heterogenous genes and further understood the interaction of heterogenous genomes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was financially supported by grants from the National Natural Science Foundation of China (Grant Nos. 32102781, U19A2040 and 31730098), Key Research and Development Program of Hunan Province (Grants No. 2020NK2016), The Science and Technology Innovation Program of Hunan Province (Grants No. 2022RC1162), Natural Science Foundation of Hunan Province, China (Grant No. 2021JJ40345), the Project of Hunan Provincial Education Department (Grant No. 22C0003), earmarked fund for China Agriculture Research System (Grant No. CARS-45), Hunan Provincial Natural Science and Technology Major Project (Grant No.2017NK1031), Cooperative Innovation Center of Engineering and New Products for Developmental Biology of Hunan Province (Grant No.20134486).

References

[1] F. Hu, H. Zhong, C. Wu, S. Wang, Z. Guo, M. Tao, C. Zhang, D. Gong, X. Gao, C. Tang, Z. Wei, M. Wen, S. Liu, Development of fisheries in China, Reprod. Breed 1 (2021) 64–79.

[2] S.J. Liu, Distant hybridization leads to different ploidy fishes, Sci. China Life Sci. 53 (2010) 416–425.
[3] S.J. Liu, Fish Distant Hybridization, China Social Sciences Press, 2014.
[4] C. Wu, J. Li, C. Cai, Q. Qin, C. Huang, Z. Chen, F. Hu, J. Hu, H. Huang, J. Luo, A new type of hybrid golden pompano "Chenhai No. 1" produced by the hybridization of (Trachinotus ovatus♀× Trachinotus blochii♂)♀× T. ovatus, Reprod. Breed 2 (2022) 78–82.
[5] Z. Zhang, J. Chen, L. Li, M. Tao, C. Zhang, Q. Qin, J. Xiao, Y. Liu, S. Liu, Research advances in animal distant hybridization, Sci. China Life Sci. 57 (2014) 889–902.
[6] T.V. Danilova, A.R. Akhunova, E.D. Akhunov, B. Friebe, B.S. Gill, Major structural genomic alterations can be associated with hybrid speciation in Aegilops markgrafii (Triticeae), Plant J. 92 (2017) 317–330.
[7] M. Navashin, Chromosome alterations caused by hybridization and their bearing upon certain general genetic problems, Cytologia 5 (1934) 169–203.
[8] S. Liu, J. Luo, J. Chai, L. Ren, Y. Zhou, F. Huang, X. Liu, Y. Chen, C. Zhang, M. Tao, B. Lu, W. Zhou, G. Lin, C. Mai, S. Yuan, J. Wang, T. Li, Q. Qin, H. Feng, K. Luo, J. Xiao, H. Zhong, R. Zhao, W. Duan, Z. Song, Y. Wang, J. Wang, L. Zhong, L. Wang, Z. Ding, Z. Du, X. Lu, Y. Gao, R.W. Murphy, Y. Liu, A. Meyer, Y.P. Zhang, Genomic incompatibilities in the diploid and tetraploid offspring of the goldfish x common carp cross, Proc. Natl. Acad. Sci. U. S. A 113 (2016) 1327–1332.
[9] Q. Liu, Y. Qi, Q. Liang, X. Xu, F. Hu, J. Wang, J. Xiao, S. Wang, W. Li, M. Tao, The chimeric genes in the hybrid lineage of Carassius auratus cuvieri (♀)× Carassius auratus red var.(♂), Sci. China Life Sci. 61 (2018) 1079–1089.
[10] B. Ji, X. Zhang, S. Zhang, H. Song, Z. Kong, Insights into the bacterial species and communities of a full-scale anaerobic/anoxic/oxic wastewater treatment plant by using third-generation sequencing, J. Biosci. Bioeng. 128 (2019) 744–750.
[11] X. Zhang, G. Li, H. Jiang, L. Li, J. Ma, H. Li, J. Chen, Full-length transcriptome analysis of Litopenaeus vannamei reveals transcript variants involved in the innate immune system, Fish Shellfish Immunol. 87 (2019) 346–359.
[12] S.E. Abdel-Ghany, M. Hamilton, J.L. Jacobi, P. Ngam, A.S.N. Reddy, A survey of the sorghum transcriptome using single-molecule long reads, Nat. Commun. 7 (2016), 11706.
[13] B. Wang, E. Tseng, M. Regulski, T.A. Clark, T. Hon, Y. Jiao, Z. Lu, A. Olson, J. C. Stein, D. Ware, Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing, Nat. Commun. 7 (2016), 11708.
[14] Y. Chao, J. Yuan, S. Li, S. Jia, L. Han, L. Xu, Analysis of transcripts and splice isoforms in red clover (Trifolium pratense L.) by single-molecule long-read sequencing, BMC Plant Biol. 18 (2018) 300.
[15] L. Ren, X. Yan, X. Gao, J. Cui, P. Yan, C. Wu, W. Li, S. Liu, Maternal effects shape the alternative splicing of parental alleles in reciprocal cross hybrids of Megalobrama amblycephala x Culter alburnus, BMC Genom. 21 (2020) 457.
[16] B. Dauphin, J.R. Grant, D.R. Farrar, C.J. Rothfels, Rapid allopolyploid radiation of moonwort ferns (Botrychium; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions, Mol. Phylogenet. Evol. 120 (2018) 342–353.
[17] N.V. Hoang, A. Furtado, P.J. Mason, A. Marquardt, L. Kasirajan, P. P. Thirugnanasambandam, F.C. Botha, R.J. Henry, A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing, BMC Genom. 18 (2017) 1–22.
[18] M. Wang, P. Wang, F. Liang, Z. Ye, J. Li, C. Shen, L. Pei, F. Wang, J. Hu, L. Tu, K. Lindsey, D. He, X. Zhang, A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation, New Phytol. 217 (2018) 163–178.
[19] C. Wu, X. Huang, F. Hu, Y. Ouyang, L. Zhao, S. Wang, W. Li, J. Fan, C. Zhang, L. Ren, Q. Qin, K. Luo, M. Tao, S. Liu, Production of diploid gynogenetic grass carp and triploid hybrids derived from the distant hybridization of female grass carp and male topmouth culter, Aquaculture 504 (2019) 462–470.
[20] C. Chen, H. Chen, Y. Zhang, H.R. Thomas, M.H. Frank, Y. He, R. Xia, TBtools-an integrative toolkit developed for interactive analyses of big biological data, Mol. Plant 13 (2020) 1194–1202.
[21] M. Sammeth, S. Foissac, R. Guigó, A general definition and nomenclature for alternative splicing events, PLoS Comput. Biol. 4 (2008), e1000147.
[22] K. Alix, P.R. Gérard, T. Schwarzacher, J. Heslop-Harrison, Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants, Ann. Bot. 120 (2017) 183–194.
[23] C. Song, S. Liu, J. Xiao, W. He, Y. Zhou, Q. Qin, C. Zhang, Y. Liu, Polyploid Organisms. Science China, vol. 55, Life sciences, 2012, pp. 301–311.
[24] P.P. Edger, T.J. Poorten, R. VanBuren, M.A. Hardigan, M. Colle, M.R. McKain, R. D. Smith, S.J. Teresi, A.D. Nelson, C.M. Wai, Origin and evolution of the octoploid strawberry genome, Nat. Genet. 51 (2019) 541–547.
[25] R.A. Gill, F. Scossa, G.J. King, A.A. Golicz, C. Tong, R.J. Snowdon, A.R. Fernie, S. Liu, On the role of transposable elements in the regulation of gene expression and subgenomic interactions in crop genomes, Crit. Rev. Plant Sci. 40 (2021) 157–189.
[26] H. Muller, Why polyploidy is rarer in animals than in plants, Am. Nat. 59 (1925) 346–353.
[27] P.Y. Novikova, I.G. Brennan, W. Booker, M. Mahony, P. Doughty, A.R. Lemmon, E. M. Lemmon, J.D. Roberts, L. Yant, Y. Van de Peer, Polyploidy breaks speciation barriers in Australian burrowing frogs Neobatrachus, PLoS Genet. 16 (2020), e1008769.
[28] T.C. Osborn, J.C. Pires, J.A. Birchler, D.L. Auger, Z.J. Chen, H.-S. Lee, L. Comai, A. Madlung, R. Doerge, V. Colot, Understanding mechanisms of novel gene expression in polyploids, Trends Genet. 19 (2003) 141–147.
[29] T. Mandáková, X. Guo, B. Özüdoğru, K. Mummenhoff, M.A. Lysak, Hybridization-facilitated genome merger and repeated chromosome fusion after 8 million years, Plant J. 96 (2018) 748–760.
[30] F. Mertens, C.D. Fletcher, C.R. Antonescu, J.M. Coindre, M. Colecchia, H. A. Domanski, E. Downs-Kelly, C. Fisher, J.R. Goldblum, L. Guillou, R. Reid, J. Rosai, R. Sciot, N. Mandahl, I. Panagopoulos, Clinicopathologic and molecular genetic characterization of low-grade fibromyxoid sarcoma, and cloning of a novel FUS/CREB3L1 fusion gene, Lab. Invest. J. Tech. Methods Pathol. 85 (2005) 408–415.
[31] M. Soda, Y.L. Choi, M. Enomoto, S. Takada, Y. Yamashita, S. Ishikawa, S.-i. Fujiwara, H. Watanabe, K. Kurashina, H. Hatanaka, M. Bando, S. Ohno, Y. Ishikawa, H. Aburatani, T. Niki, Y. Sohara, Y. Sugiyama, H. Mano, Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer, Nature 448 (2007) 561–566.
[32] G. Maire, C.W. Brown, J. Bayani, C. Pereira, D.H. Gravel, J.C. Bell, M. Zielenska, J. A. Squire, Complex rearrangement of chromosomes 19, 21, and 22 in Ewing sarcoma involving a novel reciprocal inversion–insertion mechanism of EWS–ERG fusion gene formation: a case analysis and literature review, Cancer Genet. Cytogenet. 181 (2008) 81–92.
[33] V. Tonzo, A. Papadopoulou, J. Ortego, Genomic footprints of an old affair: single nucleotide polymorphism data reveal historical hybridization and the subsequent evolution of reproductive barriers in two recently diverged grasshoppers with partly overlapping distributions, Mol. Ecol. 29 (2020) 2254–2268.
[34] J. Wang, L.H. Ye, Q.Z. Liu, L.Y. Peng, W. Liu, X.G. Yi, Y.D. Wang, J. Xiao, K. Xu, F. Z. Hu, L. Ren, M. Tao, C. Zhang, Y. Liu, Y.H. Hong, S.J. Liu, Rapid genomic DNA changes in allotetraploid fish hybrids, Heredity 114 (2015) 601–609.
[35] C. Wu, X. Huang, Q. Chen, F. Hu, L. Zhou, K. Gong, W. Fu, D. Gong, R. Zhao, C. Zhang, Q. Qin, M. Tao, S. Liu, The formation of a new type of hybrid culter derived from a hybrid lineage of Megalobrama amblycephala (♀) × Culter alburnus (♂), Aquaculture 525 (2020), 735328.
[36] S. Santini, G. Bernardi, Organization and base composition of tilapia Hox genes: implications for the evolution of Hox clusters in fish, Gene 346 (2005) 51–61.
[37] S. Wang, Q. Liu, X. Huang, C. Yang, L. Chen, M. Han, Y. Shu, M. Wang, W. Li, F. Hu, M. Wen, K. Luo, Y. Wang, R. Zhou, C. Zhang, M. Tao, R. Zhao, C. Tang, S. Liu, The rapid variation of Hox clusters reveals a clear evolutionary path in a crucian carp-like homodiploid fish lineage, Reprod. Breed 1 (2021) 149–156.
[38] L. Ren, C. Tang, W. Li, J. Cui, X. Tan, Y. Xiong, J. Chen, J. Wang, J. Xiao, Y. Zhou, J. Wang, M. Tao, C. Zhang, S. Liu, Determination of dosage compensation and comparison of gene expression in a triploid hybrid fish, BMC Genom. 18 (2017) 38.