fishes

MDPI

Article

Full-Length Transcriptome Reveals Heterologous Sperm Fragments in Natural Gynogenetic Grass Carp

Lang Qin [1,†], Yuxiang Wang [1,†], Ming Wen [1,†], Jinhui Huang [1], Xu Huang [1], Qian Chen [1], Dan Peng [1], Yang Wu [1], Qianye Wei [1], Fangzhou Hu [1], Kaijun Gong [1], Chun Zhang [1,2,3], Qinbo Qin [1,2,3,4], Chang Wu [1,2,3,*] and Shaojun Liu [1,2,3,*]

[1] Engineering Research Center of Polyploid Fish Reproduction and Breeding of the State Education Ministry, College of Life Sciences, Hunan Normal University, Changsha 410081, China; qinlang9977@163.com (L.Q.); 18774083908@163.com (Y.W.); 18711180914@163.com (M.W.); 19936855793@163.com (J.H.); xiuhuang1993@163.com (X.H.); 17136371103@163.com (Q.C.); 19973686400@163.com (D.P.); 15574358130@163.com (Y.W.); sxhnwei@sina.com (Q.W.); hufangzhou90@163.com (F.H.); m18673853053@163.com (K.G.); zhangchun421@163.com (C.Z.); qqb@hunnu.edu.cn (Q.Q.)
[2] Yuelushan Laboratory, Changsha 410128, China
[3] Institute of Interdisciplinary Studies, Hunan Normal University, Changsha 410081, China
[4] Nansha-South China Agricultural University Fishery Research Institute, Guangzhou 511457, China
* Correspondence: wuchang@hunnu.edu.cn (C.W.); lsj@hunnu.edu.cn (S.L.)
† These authors contributed equally to this work.

Abstract

Grass carp (*Ctenopharyngodon idella*) is one of the most economically important cyprinid species cultured in China. The diploid gynogenetic grass carp (2nGGC, 2n = 48) was generated from the hybrid of female grass carp (GC, 2n = 48) and male topmouth culter (TC, 2n = 48, *Culter alburnus*). This study obtained the full-length transcriptome of 2nGGC from five tissues using Pacific Biosciences (Pacbio) single-molecule real-time long-read isoform sequencing. Following the mapping of long reads to GC and TC reference genomes, a total of 1848 fusion isoforms were identified. Among them, 775 were distributed across different genomes, indicating that chimeric DNA fragments of TC were embedded in the 2nGGC genome. After removing the fusion genes and redundant isoforms, 107,721 full-length transcripts were obtained from 2nGGC, providing important full-length reference sequences for further research. Finally, comparative analysis of homologous gene variation identified 34 fragments in 2nGGC containing recombinant SNPs derived from both GC and TC. These results provide evidence that natural gynogenesis represents a form of "micro-hybridization" characterized by heterogeneous DNA fragments, distinct from traditional hybridization involving chromosome-level recombination. These findings offer valuable reference for fish genetic breeding.

Keywords: grass carp; pacbio sequencing; full-length transcript; heterologous sperm fragments; gynogenesis

Key Contribution: This study provides the first full-length transcriptome of natural gynogenetic grass carp (2nGGC), revealing the presence of paternal DNA fragments and recombinant SNPs. The findings challenge the traditional view of gynogenesis as purely matrilineal and offer new insights into genetic mechanisms in hybrid fish breeding.

Academic Editor: Rex Dunham

Received: 14 September 2025
Revised: 31 October 2025
Accepted: 5 November 2025
Published: 7 November 2025

Citation: Qin, L.; Wang, Y.; Wen, M.; Huang, J.; Huang, X.; Chen, Q.; Peng, D.; Wu, Y.; Wei, Q.; Hu, F.; et al. Full-Length Transcriptome Reveals Heterologous Sperm Fragments in Natural Gynogenetic Grass Carp. *Fishes* **2025**, *10*, 570. https://doi.org/10.3390/fishes10110570

1. Introduction

Grass carp (*Ctenopharyngodon idella*, 2n = 48) is one of the four major fish species cultured in China and an important cyprinid fish in aquaculture, valued for its herbivorous feeding habits and rapid growth rate [1–4]. The publication of the grass carp draft genome and the construction of a genome database have facilitated the molecular breeding of grass carp [5–8]. However, the genome and database do not provide full-length referenced transcript data reflecting structural changes such as alternative splicing (AS) events and fusion genes, which may hinder further research into gene function and related mechanisms.

Compared with second-generation sequencing technology, PacBio single-molecule real-time long-read isoform sequencing exhibits significant advantages in obtaining full-length transcripts and analyzing gene structural changes due to the longer reads [9–11]. Hence, large-scale Pacbio sequencing has been employed for genome assembly [12–15] and epigenome characterization [16–18] since 2015, offering an alternative to next-generation sequencing. Additionally, Pacbio sequencing has been widely used in transcriptome surveys regarding sorghum [19], corn [20], cotton [21], poplar [22], and so on.

Artificial gynogenesis is a conventional method used in fish genetic breeding, whereby eggs are activated by an inactivated heterogenous sperm [23]. In wild crucian carp populations, some reproduce by natural gynogenesis [24,25]. In addition, distant hybridization is another approach to produce natural gynogenetic offspring, such as red crucian carp [23] and blunt snout bream [26,27]. Meanwhile, DNA fragments of heterologous sperm have also been found in gynogenetic fish [2,28].

Both distant hybridization and gynogenesis can alter genotypes and phenotypes. Liu et al. have proposed the concepts of macro-hybrid and micro-hybrid based on extensive experimental findings from fish distant hybridizations and gynogenesis [29]. The term "micro-hybrid" refers to offspring whose genome originates almost solely from the maternal parent, but in which certain DNA fragments derived from the paternal parent are inserted [30].

The natural gynogenetic grass carp (2nGGC, 2n = 48) was derived from the hybridization of female grass carp and male topmouth culter. Using the Pacbio sequencing technology, long-read isoforms of grass carp were obtained from five tissues of2nGGC, including the heart, kidney, muscle, liver, and brain. This study further analyzes full-length transcripts and the paternal DNA fragments in 2nGGC, providing new insights for the fundamental research of grass carp and a new reference for genetic breeding of grass carp.

2. Materials and Methods

2.1. Materials

The diploid gynogenetic grass carp (2n = 48, 2nGGC) were derived from distant hybridization of female grass carp (Leuciscinae) and male topmouth culter (Cultrinae), which belonged to different subfamilies. All samples in this paper were cultured in the Engineering Center for Polyploidy Fish Breeding of the National Education Ministry at Hunan Normal University. Specifically, 2nGGC were subjected to deep anesthesia using 100 mg/L MS-222 (Sigma-Aldrich, St. Louis, MO, USA), and were dissected to collect heart, kidney, muscle, liver, and brain samples. Total RNA was extracted from the pooled materials above using RNeasy Plus Universal kits (Qiagen, Valencia, CA, USA).

2.2. Library Preparation and PacBio Sequencing

In order to construct a complete library, the high-quality total RNA from the five tissues was mixed. First strand cDNA was synthesized by SMARTer® PCR cDNA Synthesis Kit (Takara Bio Inc., Seta 3-4-1, Kusatsu, Shiga 525-0058, Japan), and polymerase chain reaction (PCR) was performed to amplify the cDNA and prepare the library with fragments

between 0.5–6 kb. Thereafter, those fragments were sequenced using the PacBio RS II platform for further real-time single-molecule sequencing.

2.3. Quality Control and Error Correction

The subreads were obtained from polymerase reads by removing special adapters. The high-quality region finder program was used to identify the longest region with retained polymerase activity, and the Signal Noise Ratio was used to filter out low-quality areas and short reads from the subreads. Based on the presence and relative positions of the 5′ primer, 3′ primer, and polyA, the subreads were divided into full-length non-chimeric reads and non-full-length reads. Subsequently, the ICE algorithm module was used to cluster subreads, and DSGCon was used to obtain alignment sequences [31]. The arrow algorithm was used to correct the nucleotide indels and inserts.

2.4. Mapping to the Reference Genome and Fusion Isoforms Detection

The fusion gene can be defined as the concatenation of two genes, caused by genome translocation, interstitial deletion, or chromosomal inversion [32]. The clustered isoforms were mapped into the grass carp genome [5] by GMAP [33]. Candidate fusion genes were identified as those with reads mapped to two positions distributed in different chromosomes or separated by more than 10 kb [34]. Meanwhile, each locus was required to have corresponding transcripts covering at least 10%, with the coverage of the mapped region exceeding 99%.

2.5. Annotation, LncRNA Prediction and Novel Genes Analysis

The matchAnnot was used to extract annotation information for each isoform from its best-mapped reference genome. However, isoforms mapping to regions with no genome annotations were subjected to lncRNA prediction [35]. The rest of the isoforms that were neither annotated to the genome nor predicted as lncRNAs were considered novel genes. Functional annotation of the novel genes was performed against the Gene Ontology (GO) database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database, Cluster of Orthologous Groups of proteins (COG) database, Swiss-Prot database, and NR database. The CDS and corresponding proteins of the annotated novel genes were predicted using ANGEL 2 software.

2.6. Alternative Splicing Events Analysis and Validation

In eukaryotes, pre-mRNA, as a direct transcriptional product, can undergo alternative splicing to generate multiple transcript isoforms. However, due to the limitations of short-read next-generation sequencing technology, the detection of AS events remains a challenge. Full-length transcripts were mapped to the reference genome, and related messages were stored in gff3 file. The AS events were separated from gff3 file and classified into five patterns using Rmats 2. GO and KEGG enrichment analyses were performed for the top 5% genes exhibiting AS events.

2.7. Analyses of Paternal DNA Fragments

Mao et al. [2,28] reported that some DNA fragments derived from the paternal koi carp (*Cyprinus carpio haematopterus*) were transmitted into gynogenetic grass carp. This study provided the first molecular-level evidence of paternal DNA fragment incorporation in gynogenetic grass carp. To analyze the paternal DNA fragments in 2nGGC, the full-length isoforms were mapped to the GC and TC genomes, respectively. To accurately identify heterologous fragments, the transcripts were cut into 100 bp-sized fragments. Next, these fragments were mapped to the GC genome, and sequences exhibiting less than 90% alignment similarity were screened and retained for further analysis. Moreover, the

screened fragments were mapped to the TC genome, and the fragments showing more than 99% alignment similarity were selected for further analysis. Finally, these selected fragments and corresponding genome fragments were extracted by TBtools v1.09876, and homologous fragments were blasted and analyzed by BioEdit v7.0.9.0.

3. Results

3.1. Pacbio Sequencing and Data Output

Based on three libraries derived from five tissue samples, and after filtering low-quality subreads and removing adapters, a total of 20,410,682,675 (20.4 Gb) bases of raw data were obtained (BioProject ID: PRJNA1248656), with a mean subread length of 439 bases (Figure 1A). In addition, 606,116 full-length non-chimeric reads (Figure 1B) and 81,589 non-full-length reads were identified from the analysis of inserted reads, which accounted for 86.03% and 11.58% of inserted reads, respectively. After ICE cluster and correction, a total of 337,928 consistent sequences were obtained, of which 198,465 (58.73%) were high-quality (HQ) sequences with an accuracy exceeding 99% (Figure 1C). Redundant isoforms were removed after mapping to the reference genome, and 107,721 full-length isoforms were successfully obtained.

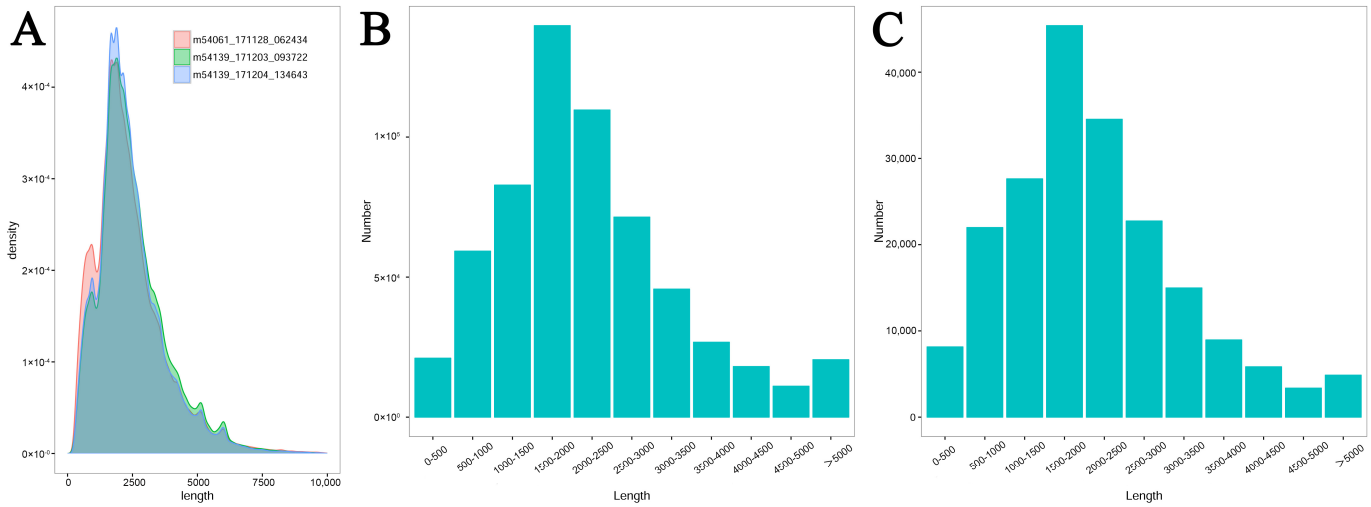

Figure 1. The number and length distribution of PacBio SMRT sequencing. (**A**) The number and length distribution of 234,838 CCS sequences. (**B**) The number and length distribution of 606,116 full-length non-chimeric reads. (**C**) The number and length distribution of 198,465 high-quality consensus sequences.

3.2. Fusion Isoform Detection

The fusion isoform was formed by two fragments located either on the same chromosome or on different chromosomes. A total of 1848 fusion isoforms were identified in 2nGGC from the mapped isoforms, including 1081 fusion isoforms distributed within the GC genome, 12 within the TC genome, and 775 distributed within both the GC and TC genomes (Figure 2). These findings indicated that, in addition to the grass carp genome, the natural gynogenetic grass carp also possessed the DNA fragments of the male parent.

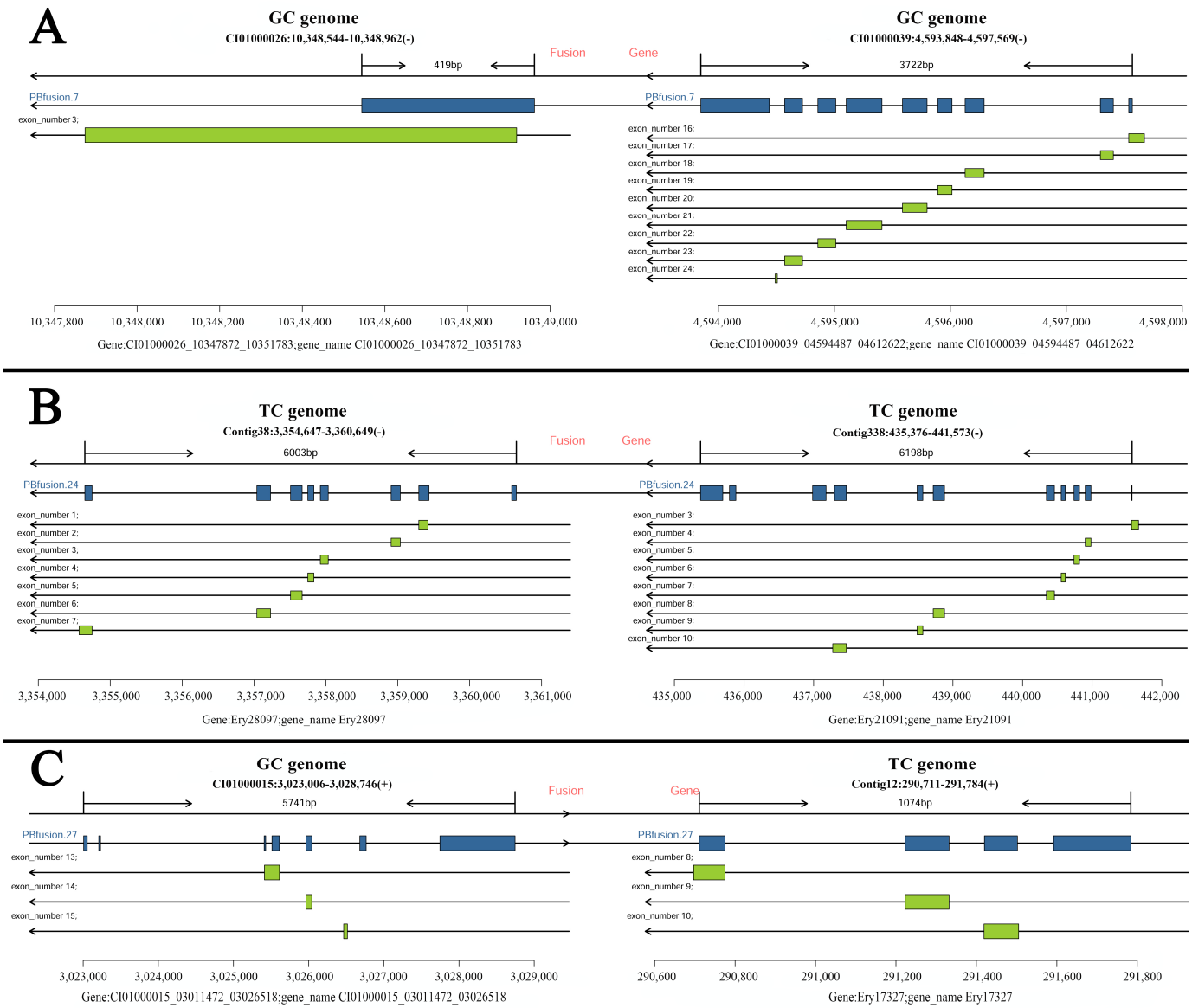

Figure 2. Three fusion gene models in 2nGGC. (**A**) The gene fusion model between different chromosomes of the GC genome. (**B**) The gene fusion model between different chromosomes of the TC genome. (**C**) The gene fusion model in GC and TC genomes.

3.3. Isoform Annotation, LncRNA Prediction and Novel Genes Analyses

Following the removal of the fusion isoforms, full-length transcriptome sequencing yielded 376,948 isoforms, of which 374,939 (99.47%) were mapped to the reference GC genome using GMAP software (Table 1). Next, collapse_isoforms_by_sam.py was used to filter out the redundant sequences, yielding 107,721 full-length isoforms. Subsequently, the mapping results were annotated by the reference genomes and extracted using matchAnnot software. A total of 92,628 isoforms were mapped to 21,702 known genes based on their genomic locations relative to the reference genome, while 15,093 isoforms aligned to regions of the reference genomes lacking annotation information. A total of 5994 lncRNAs were predicted from 15,093 unannotated isoforms using ncrna_pipeline, with an average sequence length of 1649 bp, a minimum length of 201 bp, and a maximum length of 6728 bp. Lastly, the rest of the 9099 isoforms were not predicted to lncRNAs and were identified as novel isoforms. The 9099 novel isoforms were compared against the NR, GO, KEGG, COG, and Swiss-Prot databases, with 2776 genes (30.51%) being annotated (Figure 3). In addition, the functions of the annotated novel genes were analyzed by the GO and KEGG databases.

The novel genes in the GO database were clustered in protein-binding, DNA-binding, and DNA-integration processes, and the novel genes in the KEGG database were clustered in signal transduction, endocrine system, cellular community, and immune system processes.

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 LncRNAs	Number of Novel Isoforms
376,948	374,939	99.47%	107,721	92,628	21,702	15,093	5994	9099

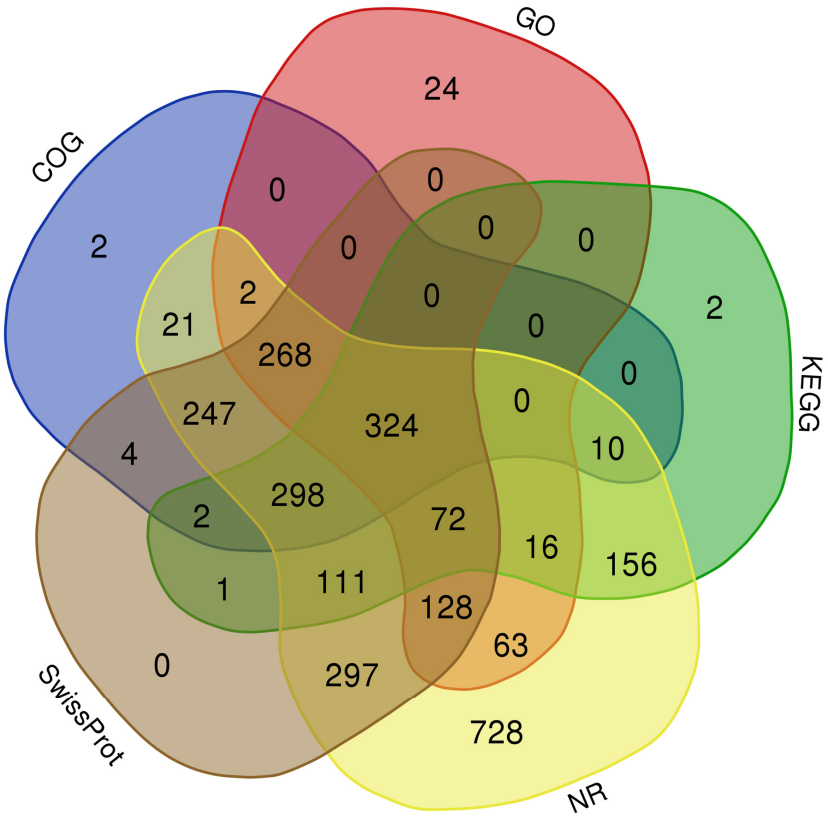

Figure 3. Functional annotations of the novel genes in NR, COG, KEGG, Swiss Prot, and GO databases.

3.4. Alternative Splicing Analysis

After mapping the transcripts to the reference grass carp genome, a total of 140,408 AS events were identified, which occurred in 34,561 genes. In this study, 140,408 AS events were identified and divided into five models: ExonS, AltD, AltA, IntronR, AltP (Table 2). The models of ExonS contained 10,934 AS events and involved 5773 genes; the models of AltD contained 6651 AS events and involved 3654 genes; the models of AltA contained 7582 AS events and involved 4242 genes; the models of IntronR contained 32,952 AS events and involved 8678 genes; the models of AltP contained 22,291 AS events and involved 5382 genes. The IntronR and AltP were the main AS models, accounting for 23.5% and 15.9% of all AS events, respectively. In addition, 59,998 AS events were categorized as Other, which involved 6832 genes. In order to analyze the function of genes involved in AS events, the top 5% genes were subjected to GO enrichment and KEGG enrichment analyses. In the KEGG analysis, tight junction, carbon metabolism, and the AMPK signaling pathway were the most enriched pathways, while the GO enrichment analysis revealed

that pyrophosphatase activity, nucleoside-triphosphatase activity, and cytoskeleton were the main GO terms.

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

Type	AS_Number	Gene_Number
Exons	10,934	5773
AltD	6651	3654
AltA	7582	4242
IntronR	32,952	8678
Other	59,998	6832
AltP	22,291	5382

3.5. Analyses of Paternal DNA Fragments

After mapping the transcripts to the GC and TC genomes, a total of 34 fragments were identified in 2nGGC. To verify whether the fragment originated from heterologous sperm, one of the fragments was randomly selected for Sanger sequencing. The fragment of PB27109.2 was amplified, sequenced, and compared at the DNA level among GC, TC, and 2nGGC (Figure 4). The results of the comparative analysis in 2nGGC and their parents showed that SNP loci in 2nGGC can not only originate from the GC genome but also from TC, indicating that the homologous recombination events can also occur in natural gynogenic grass carp. In addition, SNP mutations were also found in 2nGGC. Our results provide direct evidence that natural gynogenesis represents a "micro-hybrid", characterized by the incorporation and normal expression of paternal genetic material.

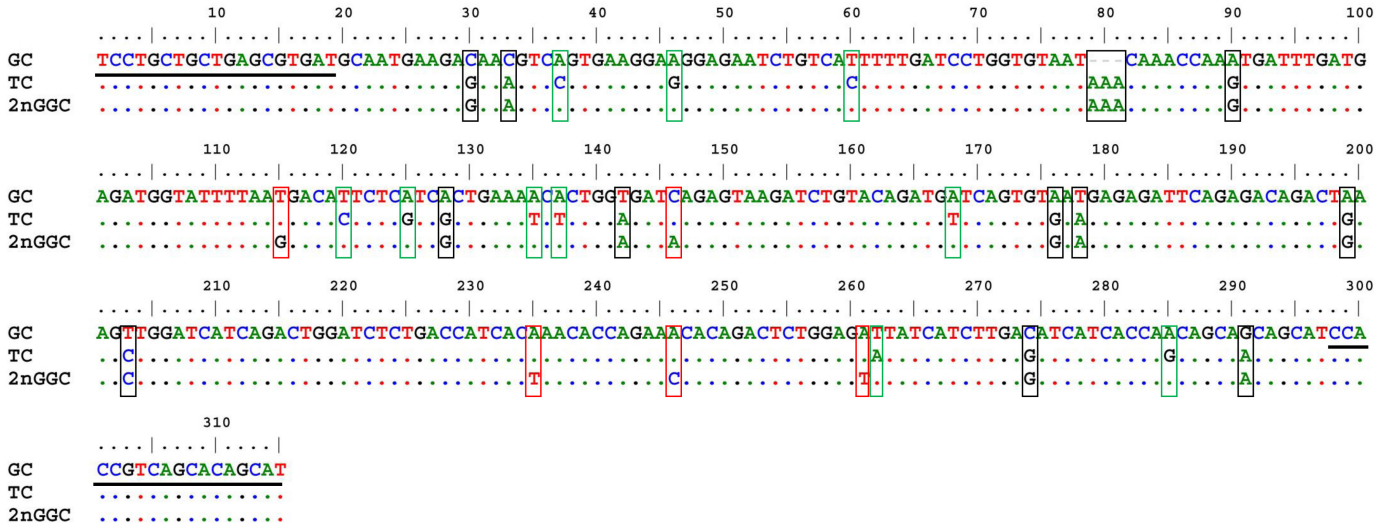

Figure 4. The nucleotide sequence alignment of PB27109.2 in GC, TC, and 2nGGC. The black boxes represent the SNPs of 2nGGC derived from TC; green boxes represent the SNPs of 2nGGC derived from GC; the red boxes represent mutational SNPs. The primers are underlined.

4. Discussion

Distant hybridization allows the transfer of genomes from one species to another, and gynogenesis is a specific type of reproduction [36]; both can increase genetic variation and accelerate speciation and evolution [37–39]. Natural gynogenesis is a unique reproductive strategy that is coupled with distant hybridization [30]. The induction of gynogenesis through distant hybridization has been reported in individual cases. Conventionally, this process was believed to occur without fusion between the egg and sperm, with chromosome doubling in the egg forming a natural gynogenesis diploid [40]. For example, in a cross

between a female common carp and a male blunt snout bream, a natural gynogenetic diploid common carp was observed in the offspring [41]. Distant hybridization between female blunt snout bream and male Chinese perch produced natural gynogenetic blunt snout bream [42]. In our previous study, a diploid gynogenetic grass carp (2nGGC, 2n = 48) from female GC × male TC was produced [43]. In this study, 107,721 full-length isoforms of 2nGGC were obtained by full-length transcriptome sequencing, and gene fusion, AS events, lncRNAs, novel genes, and paternal DNA fragments were analyzed. Our results provide a reference for full-length transcripts for future studies on genetic regulation and other basic research.

Gynogenesis is an important breeding technique that has been widely used for sex control breeding, improvement of species, germplasm purification, and in a wide variety of research areas [5,44]. Gynogenetic offspring were once thought to be completely matrilineal. However, the residue of paternal genetic material in the offspring has led to new insights into gynogenesis. In this study, following full-length transcriptome sequencing, 775 fusion isoforms derived from both GC and TC genomes were identified. Meanwhile, 34 fragments in 2nGGC exhibiting homologous recombinant sites from GC and TC were detected. These results provide direct evidence for the presence of paternal genetic material in the 2nGGC. In the present study, micro-chromosomes in the offspring of gynogenetic gibel carp (*Carassius auratus gibelio*) were found and sequenced; the presence of micro-chromosomes was also confirmed by fluorescent labeling [45]. Liu et al. introduced the concept of "micro-hybrid", whose genome almost originates solely from the maternal parent, but also contains paternal DNA fragments [30]. The residual paternal DNA in the offspring of gynogenesis allowed the co-existence of both parents' DNA in the offspring, providing conditions for the formation of recombinant genes [46,47]. The *HoxC6b* gene in gynogenetic grass carp combined the sequence structure of both parents' *HoxC6b* gene, which implied that the *HoxC6b* gene of koi carp (*Cyprinus carpio haematopterus*) and common grass carp (*Ctenopharyngodon idella*) have been recombined, forming a new recombinant gene in the gynogenetic grass carp [28]. Recombinant genes are the product of interactions between the genetic material of parents and are widely found in hybrid fish. Notably, recombination in natural gynogenetic offspring provides important evidence supporting gynogenesis as a new type of hybridization with a small amount of paternal genetic information.

Our study identified paternal (topmouth culter) DNA fragments and recombinant SNPs in natural gynogenetic grass carp (2nGGC), challenging the conventional view that gynogenesis is a strictly matrilineal mode of reproduction. This finding suggested that gynogenesis represents a cryptic form of hybridization—termed "micro-hybrid"—that could serve as an evolutionary mechanism for maintaining genetic diversity. Furthermore, it provided a novel perspective for understanding how certain fish, such as silver crucian carp and grass carp [5,48], underwent multiple independent origins in nature and successfully adapted, shedding light on their remarkable evolutionary plasticity.

5. Conclusions

In conclusion, a high-quality full-length transcriptome was obtained for 2nGGC, and gene fusions, AS events, lncRNAs, and novel genes were characterized. Meanwhile, 787 fusion genes involving the TC genome and 34 paternal DNA fragments in 2nGGC were further analyzed. This study not only provides a rich resource of transcript isoforms for the natural gynogenetic grass carp but also provides an important insight into potential "hybridization effects" in gynogenesis (Figure 5).

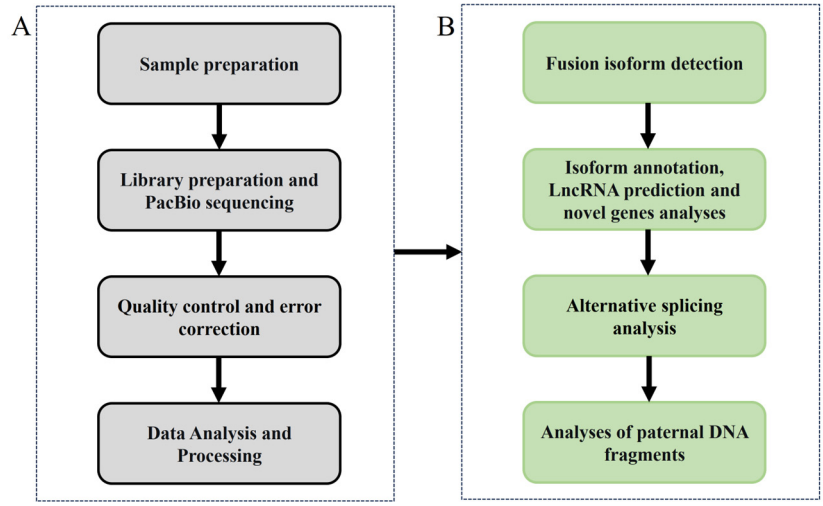

Figure 5. Summary of this paper. (**A**) Experimental methods. (**B**) Experimental results.

Author Contributions: L.Q., Y.W. (Yuxiang Wang) and M.W. conceived and designed the experiments. J.H., X.H., D.P., Y.W. (Yang Wu) and Q.W. performed the bioinformatics analysis and prepared the manuscript, the table, and the figures. Q.C. and K.G. conducted the experiment. F.H., C.Z., Q.Q., C.W. and S.L. collected the samples. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by grants from Yuelushan Laboratory Breeding Program (Grant No. YLS-2025-ZY02043), the National Natural Science Foundation of China (Grant Nos. 32102781, 32293250, 32293251), Key Research and Development Program of Hunan Province (Grants No. 2023WK2001), The Science and Technology Innovation Program of Hunan Province (Grants No. 2022RC1162), Natural Science Foundation of Hunan Province, China (Grant No. 2024JJ6314), Changsha Municipal Natural Science Foundation (Grant No. kq2402160).

Institutional Review Board Statement: The present study was approved by the Ethics Committee of Hunan Normal University (protocol code 2023-610, date: 4 November 2023).

Data Availability Statement: Data will be made available on request. Assembly accession of the GC reference genome (GCA_051903545.1). Assembly accession of the TC reference genome (GCA_009869775.1).

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

Abbreviations

The following abbreviations are used in this manuscript:

2nGGC	2n Gynogenetic Grass Carp
GC	Grass Carp
Pacbio	Pacific Biosciences
SNP	Single-Nucleotide Polymorphism
AS	Alternative Splicing
PCR	Polymerase Chain Reaction
ICE	In-Context Learning
GMAP	Genomic Mapping and Alignment Program
LncRNA	Long non-coding RNA
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
COG	Cluster of Orthologous Groups
NR	Non-Redundant

CDS	Coding Sequence
HQ	Higher Quality
ExonS	Exon Skip
AltD	Alternate Donor Site
AltA	Alternate Acceptor Site
IntronR	Intron Retention
AltP	Alternate Promoter
AMPK	AMP-activated Protein Kinase

References

1. Li, L.; Liang, X.F.; He, S.; Sun, J.; Wen, Z.Y.; He, Y.H.; Cai, W.J.; Wang, Y.P.; Tao, Y.X. Transcriptome analysis of grass carp (*Ctenopharyngodon idella*) fed with animal and plant diets. *Gene* **2015**, *574*, 371–379. [CrossRef]
2. Mao, Z.; Fu, Y.; Wang, Y.; Wang, S.; Zhang, M.; Gao, X.; Luo, K.; Qin, Q.; Zhang, C.; Tao, M.; et al. Evidence for paternal DNA transmission to gynogenetic grass carp. *BMC Genet.* **2019**, *20*, 3. [CrossRef]
3. He, W.G.; Xie, L.; Li, T.; Liu, S.; Xiao, J.; Hu, J.; Wang, J.; Qin, Q.; Liu, Y. The formation of diploid and triploid hybrids of female grass carp x male blunt snout bream and their 5S rDNA analysis. *BMC Genet.* **2013**, *14*, 110. [CrossRef]
4. Chilton, E.W.; Muoneke, M.I. Biology and management of grass carp (*Ctenopharyngodon idella*, Cyprinidae) for vegetation control: A North American perspective. *Rev. Fish Biol. Fish.* **1992**, *2*, 283–320. [CrossRef]
5. Wang, Y.; Lu, Y.; Zhang, Y.; Ning, Z.; Li, Y.; Zhao, Q.; Lu, H.; Huang, R.; Xia, X.; Feng, Q.; et al. The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation. *Nat. Genet.* **2015**, *47*, 625–631. [CrossRef]
6. Chen, Y.; Shi, M.; Zhang, W.; Cheng, Y.; Wang, Y.; Xia, X.Q. The Grass Carp Genome Database (GCGD): An online platform for genome features and annotations. *Database* **2017**, *2017*, bax051. [CrossRef] [PubMed]
7. Wang, Y.; Liu, W.; Li, Z.; Qiu, B.; Li, J.; Geng, G.; Hu, B.; Liao, A.; Cai, Y.; Wen, M. Improvement and application of genetic resources of grass carp (*Ctenopharyngodon idella*). *Reprod. Breed.* **2024**, *4*, 126–133. [CrossRef]
8. Tan, H.; Wang, Y.; Hu, B.; Zhang, Y.; Liao, A.m.; Liu, W.; Gen, C.; Luo, K.; Tao, M.; Zhang, C. Analysis of the genetic characteristics and variations in disease-resistant grass carp based on whole-genome resequencing and transcriptome sequencing. *Reprod. Breed.* **2024**, *4*, 22–31. [CrossRef]
9. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* **2015**, *13*, 278–289. [CrossRef] [PubMed]
10. Li, Y.; Dai, C.; Hu, C.; Liu, Z.; Kang, C. Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry. *Plant J.* **2017**, *90*, 164–176. [CrossRef] [PubMed]
11. Ardui, S.; Ameur, A.; Vermeesch, J.R.; Hestand, M.S. Single molecule real-time (SMRT) sequencing comes of age: Applications and utilities for medical diagnostics. *Nucleic Acids Res.* **2018**, *46*, 2159–2168. [CrossRef]
12. Zhang, J.; Chen, L.L.; Sun, S.; Kudrna, D.; Copetti, D.; Li, W.; Mu, T.; Jiao, W.B.; Xing, F.; Lee, S. Building two indica rice reference genomes with PacBio long-read and Illumina paired-end sequencing data. *Sci. Data* **2016**, *3*, 160076. [CrossRef]
13. Korlach, J.; Gedman, G.; Kingan, S.B.; Chin, C.S.; Howard, J.T.; Audet, J.N.; Cantin, L.; Jarvis, E.D. De novoPacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *Gigascience* **2017**, *6*, gix085. [CrossRef]
14. Jung-Ting, C.; Pakala, S.B.; Geraldo, J.A.; Lapp, S.A.; Humphrey, J.C.; Barnwell, J.W.; Kissinger, J.C.; Galinski, M.R. High-Quality Genome Assembly and Annotation forPlasmodium coatneyi, Generated Using Single-Molecule Real-Time PacBio Technology. *Genome Announc.* **2016**, *4*, e00883-16.
15. Giordano, F.; Aigrain, L.; Quail, M.A.; Coupland, P.; Bonfield, J.K.; Davies, R.M.; Tischler, G.; Jackson, D.K.; Keane, T.M.; Li, J. De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms. *Sci. Rep.* **2017**, *7*, 3935. [CrossRef] [PubMed]
16. Yang, Y.; Sebra, R.; Pullman, B.S.; Qiao, W.; Peter, I.; Desnick, R.J.; Geyer, C.R.; Decoteau, J.F.; Scott, S.A. Quantitative and multiplexed DNA methylation analysis using long-read single-molecule real-time bisulfite sequencing (SMRT-BS). *BMC Genom.* **2015**, *16*, 350. [CrossRef] [PubMed]
17. Yang, Y.; Scott, S.A. *DNA Methylation Profiling Using Long-Read Single Molecule Real-Time Bisulfite Sequencing (SMRT-BS)*; Humana Press: New York, NY, USA, 2017; pp. 125–134.
18. Kelleher, P.; Murphy, J.; Mahony, J.; Sinderen, D.V. *Identification of DNA Base Modifications by Means of Pacific Biosciences RS Sequencing Technology*; Humana Press: New York, NY, USA, 2018.
19. Abdel-Ghany, S.E.; Hamilton, M.; Jacobi, J.L.; Ngam, P.; Devitt, N.; Schilkey, F.; Ben-Hur, A.; Reddy, A.S. A survey of the sorghum transcriptome using single-molecule long reads. *Nat. Commun.* **2016**, *7*, 11706. [CrossRef]
20. Wang, B.; Tseng, E.; Regulski, M.; Clark, T.A.; Hon, T.; Jiao, Y.; Lu, Z.; Olson, A.; Stein, J.C.; Ware, D. Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. *Nat. Commun.* **2016**, *7*, 11708. [CrossRef]

21. Wang, M.; Wang, P.; Liang, F.; Ye, Z.; Li, J.; Shen, C.; Pei, L.; Wang, F.; Hu, J.; Tu, L.; et al. A global survey of alternative splicing in allopolyploid cotton: Landscape, complexity and regulation. *New Phytol.* **2017**, *217*, 163–178. [CrossRef]
22. Filichkin, S.A.; Hamilton, M.; Dharmawardhana, P.D.; Singh, S.K.; Sullivan, C.; Ben-Hur, A.; Reddy, A.S.N.; Jaiswal, P. Abiotic Stresses Modulate Landscape of Poplar Transcriptome via Alternative Splicing, Differential Intron Retention, and Isoform Ratio Switching. *Front. Plant Sci.* **2018**, *9*, 5. [CrossRef]
23. Liu, S.J.; Qin, Q.B.; Wang, Y.; Zhang, H.; Zhao, R.; Zhang, C.; Wang, J.; Li, W.; Chen, L.; Xiao, J.; et al. Evidence for the formation of the male gynogenetic fish. *Mar. Biotechnol.* **2010**, *12*, 160–172. [CrossRef] [PubMed]
24. Xiao, J.; Zou, T.M.; Chen, Y.B.; Chen, L.; Liu, S.J.; Tao, M.; Zhang, C.; Zhao, R.R.; Zhou, Y.; Long, Y.; et al. Coexistence of diploid, triploid and tetraploid crucian carp (*Carassius auratus*) in natural waters. *BMC Genet.* **2011**, *12*, 20. [CrossRef]
25. Zhou, L.; Wang, Y.; Gui, J.-F. Genetic evidence for gonochoristic reproduction in gynogenetic silver crucian carp (*Carassius auratus* gibelio Bloch) as revealed by RAPD assays. *J. Mol. Evol.* **2000**, *51*, 498–506. [CrossRef]
26. Wu, P.; Zeng, Y.; Qin, Q.; Wu, C.; Wang, Y.; Zhao, R.; Tao, M.; Zhang, C.; Tang, C.; Liu, S. Comparative analysis of the texture, composition, antioxidant capacity and nutrients of natural gynogenesis blunt snout bream and its parent muscle. *Reprod. Breed.* **2022**, *2*, 149–155. [CrossRef]
27. Fan, S.; Tang, Z.; Wang, Y.; Zhou, Z.; Wu, C.; Luo, K.; Hu, J.; Gong, D.; Li, S.; Tao, M.; et al. The bisexual natural gynogenetic blunt snout bream lineage derived from the distant hybridization of female blunt snout bream and male Bleeker's yellow tail. *Aquac. Rep.* **2024**, *37*, 102206. [CrossRef]
28. Mao, Z.; Fu, Y.; Wang, S.; Wang, Y.; Luo, K.; Zhang, C.; Tao, M.; Liu, S. Further evidence for paternal DNA transmission in gynogenetic grass carp. *Sci. China Life Sci.* **2020**, *63*, 1287–1296. [CrossRef]
29. Liu, Q.; Liao, A.; Tao, M.; Qin, Q.; Luo, K.; Zhang, C.; Wang, S.; Zhou, Y.; Hu, F.; Wang, Y.; et al. Macro-Hybrid and Micro-Hybrid of Fish. *Rev. Aquac.* **2025**, *18*, e70106. [CrossRef]
30. Liu, Q.; Wang, S.; Tang, C.; Tao, M.; Zhang, C.; Zhou, Y.; Qin, Q.; Luo, K.; Wu, C.; Hu, F.; et al. The research advances in distant hybridization and gynogenesis in fish. *Rev. Aquac.* **2025**, *17*, e12972. [CrossRef]
31. Sahlin, K.; Medvedev, P. De novo clustering of long-read transcriptome data using a greedy, quality-value based algorithm. In Proceedings of the International Conference on Research in Computational Molecular Biology, Washington, DC, USA, 5–8 May 2019; pp. 227–242.
32. Weirather, J.L.; Afshar, P.T.; Clark, T.A.; Tseng, E.; Powers, L.S.; Underwood, J.G.; Zabner, J.; Korlach, J.; Wong, W.H.; Au, K.F. Characterization of fusion genes and the significantly expressed fusion isoforms in breast cancer by hybrid sequencing. *Nucleic Acids Res.* **2015**, *43*, e116. [CrossRef]
33. Wu, T.D.; Watanabe, C.K. GMAP: A genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **2005**, *21*, 1859. [CrossRef]
34. Qin, Q.; Popic, V.; Wienand, K.; Yu, H.; White, E.; Khorgade, A.; Shin, A.; Georgescu, C.; Campbell, C.D.; Dondi, A. Accurate fusion transcript identification from long-and short-read isoform sequencing at bulk or single-cell resolution. *Genome Res.* **2025**, *35*, 967–986. [CrossRef]
35. Chen, X.; Sun, Y.-Z.; Guan, N.-N.; Qu, J.; Huang, Z.-A.; Zhu, Z.-X.; Li, J.-Q. Computational models for lncRNA function prediction and functional similarity calculation. *Brief. Funct. Genom.* **2019**, *18*, 58–82. [CrossRef] [PubMed]
36. Wu, C.; Chen, Q.; Huang, X.; Hu, F.; Zhu, S.; Luo, L.; Gong, D.; Gong, K.; Zhao, R.; Zhang, C.; et al. Genomic and epigenetic alterations in diploid gynogenetic hybrid fish. *Aquaculture* **2019**, *512*, 734383. [CrossRef]
37. Huang, X.; Huang, J.; Qin, L.; Long, Y.; Wu, C.; Gong, K.; Xiao, Q.; Qin, Q. Full-length transcriptome reveals rapid genetic changes in triploid hybrid grass carp derived from female grass carp × male topmouth culter. *Reprod. Breed.* **2023**, *3*, 82–87. [CrossRef]
38. Liu, S. Distant hybridization leads to different ploidy fishes. *Sci. China Life Sci.* **2010**, *53*, 416–425. [CrossRef]
39. Liu, S.; Duan, W.; Tao, M.; Zhang, C.; Sun, Y.; Shen, J.; Wang, J.; Luo, K.; Liu, Y. Establishment of the diploid gynogenetic hybrid clonal line of red crucian carp × common carp. *Sci. China Ser. C Life Sci.* **2007**, *50*, 186–193. [CrossRef]
40. Hubbs, C.L.; Hubbs, L.C. Apparent parthenogenesis in nature, in a form of fish of hybrid origin. *Science* **1932**, *76*, 628–630. [CrossRef]
41. Wang, S.; Ye, X.; Wang, Y.; Chen, Y.; Lin, B.; Yi, Z.; Mao, Z.; Hu, F.; Zhao, R.; Wang, J.; et al. A new type of homodiploid fish derived from the interspecific hybridization of female common carp × male blunt snout bream. *Sci. Rep.* **2017**, *7*, 4189. [CrossRef] [PubMed]
42. Wu, P.; Ji, W.; Zeng, Y.; Tang, J.; Wu, C.; Qin, Q.; Yi, T.; Zhou, Y.; Zhao, R.; Tao, M. The analysis of growth performance and expression of growth-related genes in natural gynogenic blunt snout bream muscle derived from the blunt snout bream (*Megalobrama amblycephala*, ♀) × Chinese perch (*Siniperca chuatsi*, ♂). *Aquaculture* **2024**, *590*, 741052. [CrossRef]
43. Wu, C.; Huang, X.; Hu, F.; Ouyang, Y.; Zhao, L.; Wang, S.; Li, W.; Fan, J.; Zhang, C.; Ren, L.; et al. Production of diploid gynogenetic grass carp and triploid hybrids derived from the distant hybridization of female grass carp and male topmouth culter. *Aquaculture* **2019**, *504*, 462–470. [CrossRef]

44. Li, X.-Y.; Liu, X.-L.; Zhu, Y.-J.; Zhang, J.; Ding, M.; Wang, M.-T.; Wang, Z.-W.; Li, Z.; Zhang, X.-J.; Zhou, L.; et al. Origin and transition of sex determination mechanisms in a gynogenetic hexaploid fish. *Heredity* **2018**, *121*, 64–74. [CrossRef] [PubMed]
45. Yi, M.S.; Li, Y.Q.; Liu, J.D.; Zhou, L.; Yu, Q.X.; Gui, J.F. Molecular cytogenetic detection of paternal chromosome fragments in allogynogenetic gibel carp, *Carassius auratus* gibelio Bloch. *Chromosome Res.* **2003**, *11*, 665–671. [CrossRef] [PubMed]
46. Ren, L.; Li, W.; Qin, Q.; Dai, H.; Han, F.; Xiao, J.; Gao, X.; Cui, J.; Wu, C.; Yan, X.; et al. The subgenomes show asymmetric expression of alleles in hybrid lineages of *Megalobrama amblycephala* × *Culter alburnus*. *Genome Res.* **2019**, *29*, 1805–1815. [CrossRef] [PubMed]
47. Zhang, Z.; Chen, J.; Li, L.; Tao, M.; Zhang, C.; Qin, Q.; Xiao, J.; Liu, Y.; Liu, S. Research advances in animal distant hybridization. *Sci. China Life Sci.* **2014**, *57*, 889–902. [CrossRef] [PubMed]
48. Yang, L.; Yang, S.-T.; Wei, X.-H.; Gui, J.-F. Genetic diversity among different clones of the gynogenetic silver crucian carp, *Carassius auratus* gibelio, revealed by transferrin and isozyme markers. *Biochem. Genet.* **2001**, *39*, 213–225. [CrossRef]