SCIENCE CHINA
Life Sciences

The chimeric genes in the hybrid lineage of Carassius auratus cuvieri (♀)×Carassius auratus red var. (♂)

Liu Qingfeng, Qi Yanhua, Liang Qiuli, Xu Xiujuan, Hu Fangzhou, Wang Jing, Xiao Jun, Wang Shi, Li Wuhui, Tao Min, Qin Qinbo, Zhao Rurong, Yao Zhanzhou and Liu Shaojun

Citation: SCIENCE CHINA Life Sciences ; doi: 10.1007/s11427-017-9306-7

View online: http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

Published by the Science China Press

Articles you may be interested in

Genetic basis and breeding application of clonal diversity and dual reproduction modes in polyploid *Carassius auratus gibelio*
SCIENCE CHINA Life Sciences **53**, 409 (2010);

STUDIES ON ULTRASTRUCTURE OF MATURE ERYTHROCYTE IN GOLDFISH(Carassius auratus)
Chinese Science Bulletin **34**, 330 (1989);

Two unisexual artificial polyploid clones constructed by genome addition of common carp (*Cyprinus carp*) and crucian carp (*Carassius auratus*)
Science in China Series C-Life Sciences **46**, 595 (2003);

The formation of improved tetraploid population of red crucian carp × common carp hybrids by androgenesis
Science in China Series C-Life Sciences **50**, 753 (2007);

Distant hybridization leads to different ploidy fishes
SCIENCE CHINA Life Sciences **53**, 416 (2010);

SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-017-9306-7

The chimeric genes in the hybrid lineage of *Carassius auratus cuvieri* (♀)×*Carassius auratus* red var. (♂)

Qingfeng Liu[1,2†], Yanhua Qi[1,2†], Qiuli Liang[1,2†], Xiujuan Xu[1,2], Fangzhou Hu[1,2], Jing Wang[1,2], Jun Xiao[1,2], Shi Wang[1,2], Wuhui Li[1,2], Min Tao[1,2], Qinbo Qin[1,2], Rurong Zhao[1,2], Zhanzhou Yao[1,2] & Shaojun Liu[1,2*]

[1]*State Key Laboratory of Developmental Biology of Freshwater Fish, Hunan Normal University, Changsha 410081, China;*
[2]*College of Life Sciences, Hunan Normal University, Changsha 410081, China*

Received December 19, 2017; accepted February 6, 2018; published online June 13, 2018

Hybridization can combine the genomes of different strains or species, which leads to changes of genotype and phenotype in the hybrids. In this study, we aimed to investigate the genetic variations of hybrids (WR-F_1 and WR-F_2) derived from the intraspecific hybridization of white crucian carp (*Carassius auratus cuvieri*, WCC, ♀) and red crucian carp (*Carassius auratus* red var., RCC, ♂). Here, we compared the orthologous genes in the liver transcriptomes of hybrids with those of WCC and RCC, and classified the orthologous genes into eight gene patterns within three categories (chimera, mutant, and biparental origin genes). The results revealed 19.04%, 4.17% chimeric genes and 6.90%, 5.05% mutations of orthologous genes in WR-F_1 and WR-F_2 respectively. Seventeen of twenty-three characterized genes (77%) were confirmed to be the chimeras at the genomic DNA level. The GO classification discovered that some chimeric and mutant genes were related to metabolic process, immune system and developmental process in WR-F_1. Our results provide the new evidence that hybridization can combine the parental genomes, leading to changes in the genotype of the resultant hybrids. This is the first report on the formation of chimeric genes from fish intraspecific hybridization, which is potentially interesting from the context of both evolution and the genetic breeding of fish.

intraspecific hybridization, chimeric genes, transcriptomes, sequence validation, genetic variation

Citation: Liu, Q., Qi, Y., Liang, Q., Xu, X., Hu, F., Wang, J., Xiao, J., Wang, S., Li, W., Tao, M., et al. (2018). The chimeric genes in the hybrid lineage of *Carassius auratus cuvieri* (♀)×*Carassius auratus* red var. (♂). Sci China Life Sci 61, https://doi.org/10.1007/s11427-017-9306-7

INTRODUCTION

Hybridization is defined as the mating of genetically differentiated individuals or groups to generate a new population with heterosis (Bartley et al., 2000; Chen, 2007; Zhang et al., 2014). Hybridization can combine the whole genomes from two different strains or species, which may lead to changes of both phenotype and genotype in the offspring (Liu, 2010). This breeding technique is widely used for the genetic breeding of fish. Commonly, the desired goal is to produce offspring that exhibit hybrid vigor or positive heterosis (Bartley et al., 2000). Many successful fish hybrids and polyploidy fish have been reported previously, such as common carp (*Cyprinus carpio*)×blunt snout bream (*Megalobrama amblycephala*) (Wang et al., 2017), red crucian carp (*Carassius auratus* red *var.*) (♀)×common carp (♂) (Liu et al., 2001; Liu et al., 2007), red crucian carp (♀)×blunt snout bream (♂) (Qin et al., 2010), blunt snout bream (♀)×topmouth culter (*Culter alburnus*) (♂) (Xiao et al., 2014), polyploidy crucian carp (Qin et al., 2016).

White crucian carp (*Carassius auratus cuvieri*, WCC) and red crucian carp (*Carassius auratus* red var., RCC) are the

†Contributed equally to this work
*Corresponding author (email: lsj@hunnu.edu.cn)

© Science China Press and Springer-Verlag GmbH Germany 2018

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

members of Cyprinidae (genus *Carassius*) and have the same number of chromosomes (2*n*=100). In our previous study (Wang et al., 2015), the crossing of WCC (female) and RCC (male) produced diploid hybrids (WR). These hybrids had a gray color, exhibited hybrid morphological traits which were different from the parents, and showed obvious heterosis; sequence analysis of ribosomal DNA revealed that recombination and variation existed in WR (Wang et al., 2015). The formation of WR is important, not only for fish breeding, but also to research the genomic variation in hybridization.

Previous studies indicated that hybridization can combine valuable traits from two species into a single group, including disease resistance, good growth, and superior flesh quality (Bartley et al., 2000). These important phenotypic variations could result from the interaction between the two parental genomes, including allelic heterozygosity, and/or epigenetic changes, which result in changes at epigenetic, genetic, and gene expression levels (Mallet, 2007). In addition, hybridization can also cause drastic genetic and genomic imbalances, including chromosomal rearrangements (Kenton et al., 1993); transpositions (Wendel, 2000); deletions and insertions (Ozkan et al., 2001); dosage imbalances (Galili et al., 1986); a high rate of DNA mutations and recombination (Arnheim et al., 1980; Hu et al., 1998); and other non-Mendelian phenomena (Soltis and Soltis, 1995). However, the molecular mechanisms underlying these changes are rarely reported. RNA-seq is a recently developed method which can analyze transcriptomes in a very high-throughput manner (Wang et al., 2009). In this study, we present an analysis of the transcriptomes from the livers of WCC, RCC, WR-F_1, and WR-F_2 to reveal the genomic variation of the hybrid fish.

RESULTS

Formation of experimental fish

Hybrid fish were generated by crossing female *Carassius auratus cuvieri* with male *Carassius auratus* red var. (Figure 1A and B). One type of F_1 hybrid was obtained in the present study, which was diploid (2*n*=100), as previously reported (Wang et al., 2015). The diploid hybrid was bisexually fertile showing a normal pattern of gamete development (Figure 1C). Meanwhile, as previously reported (Wang et al., 2015), an F_2 population was created by F_1 self-crossing (Figure 1D).

Body weights of WCC, RCC, and WR for two years

Mean BWs of WCC, RCC and WR-F_1 are shown in Table 1. The mean body weight of 3 months old WR was 140 g, which was similar to the WCC (147 g), but significantly higher than the RCC (80 g) (P<0.05). The mean body weight of 8 months old WR was 328 g, which was higher than the WCC (310 g), and significantly higher than the RCC (217 g) (P<0.05). The mean body weight of 1 year old WR was 350 g, which was similar to the WCC (395 g), but significantly higher than the RCC (292 g) (P<0.05). The mean body weight of 2 years old WR (750 g) was similar to the WCC (762 g), but significantly higher than the RCC (410 g) (P<0.05).

Liver transcriptome analysis

Some of the detailed data about RNA-seq are showed in Table 2. In total, we identified 6,712 orthologous genes among the transcriptomes of WCC, RCC, and WR-F_1; 4,475 orthologous genes among the transcriptomes of WCC, RCC, and WR-F_2. Chimeric patterns were identified based on these orthologous genes and the distribution of variations. We classified eight gene patterns within three categories (chimera, mutant, and biparental origin genes) in hybrids (Figures 2 and 3, Table 3). The first category included patterns 1–3, which were with single chimeric fragment consisting of continuous, alternating variations from parent-specific variants, either with or without offspring-specific mutations. Pattern 1 and pattern 2 included genes that had a single chimeric fragment consisting of continuous, alternating variations from parent-specific variants but without offspring-specific mutations. Pattern 3 included genes that had chimeric fragments consisting of continuous, alternating variations from parent-specific variants but with offspring-specific mutations. Chimeric genes from the first category comprised 19.04% of the genes in overlapping mapped regions in WR-F_1 and 4.18% of the genes in overlapping mapped regions in WR-F_2.

The second category included chimeric genes with offspring-specific variations. These genes are grouped into patterns 4–6, which consisted of genes derived from both progenitors but with mutations unique to the offspring. Chimeric genes from the second category comprised 6.90% of genes in the overlapping mapped regions of WR-F_1 and 5.05% of genes in the overlapping mapped regions of WR-F_2.

Finally, the third category included pattern 7 and pattern 8, in which genes were derived exclusively from one parent. Pattern 7 included genes from the WCC parent while pattern 8 included genes from the RCC parent. The third category comprised 31.08% of genes in the overlapping mapped regions of WR-F_1 and 36.20% of genes in the overlapping mapped regions of WR-F_2.

Chimeric genes validated in parents and hybrids

We confirmed the presence of chimeric genes at the genomic level by both PCR amplification and Sanger sequencing. In total, Sanger sequencing validated 22 of the 30 genes tested,

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

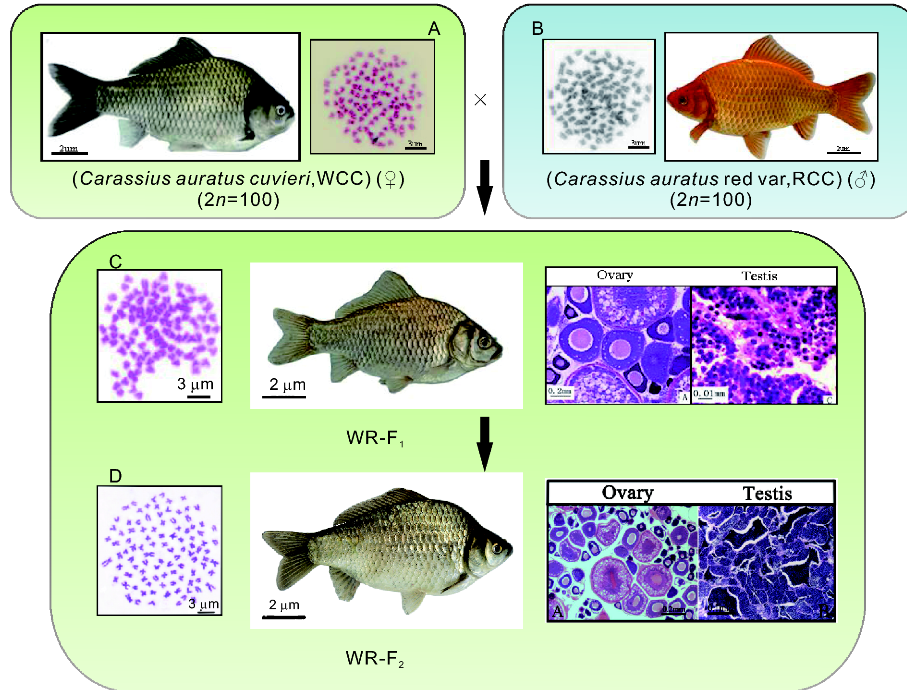

Figure 1 Formation of *Carassius auratus cuvieri*×*Carassius auratus* red var. hybrids. A, 100 chromosomes were observed in *Carassius auratus cuvieri*. B, 100 chromosomes were observed in *Carassius auratus* red var. C, Following hybridization, an F_1 generation was obtained, which also exhibited 100 chromosomes. Gonadal histology of the hybrid showed that the fish were fertile. D, An F_2 generation was created by F_1 self-crossing, which also exhibited 100 chromosomes. Gonadal histology of the hybrid showed that the fish were fertile.

Table 1 Comparison of the mean body weight of WCC, RCC, and WR-F_1 (g)[a)]

	3 months	4 months	6 months	8 months	10 months	1 year	2 years
WCC	147±10	180±14	245±18	310±20	372±36	395±42	762±48
RCC	80±6*	113±8*	156±10*	217±17*	246±23*	292±33*	410±45*
WR-F_1	140±8	158±13	242±20	328±26	334±30	350±34	750±40

a) Values are mean±SE (g). *, Statistically significant differences between RCC and WR-F_1 (P<0.05).

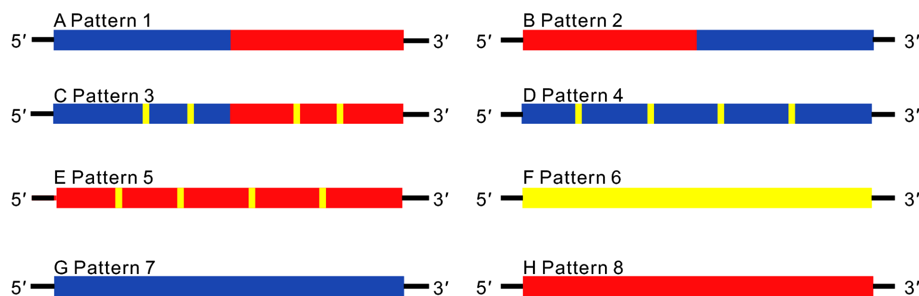

Figure 2 (Color online) Schematic diagrams of gene patterns for the offspring arising from the hybridization of WCC (W) and RCC (R) fish. Blue bars marked W variation denote offspring fragments with the WCC-specific variants; red bars marked R variation show RCC-specific variants; and yellow bars marked F-variation show offspring-specific variants. Genes were classified into three categories. The first category includes patterns 1–3 (A–C, respectively) in which chimeric genes had single or multiple chimeric fragments consisting of continuous, alternating variations from parent-specific variants, either with or without offspring-specific mutations. The second category included patterns 4–6 (D–F, respectively), which consisted of genes derived from either or both progenitors but with mutations unique to offspring. The third category included patterns 7–8 (G–H, respectively) where genes were derived exclusively from one parent.

indicating a 73% success rate for the bioinformatic identification of chimeric and other patterns. There were 23 characterized chimeric genes in the 30 tested genes, and 17 of 23 characterized chimeric genes (74%) were confirmed to the chimeras by Sanger sequencing. For example, we were able to confirm that ZFP36 ring finger protein like 2

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

A

zfp3612

Gynogenetic bred goldfish genome scaffold57

753098

752438

Sequence by Sanger method

1

661

Contig 20156 by *de novo* assembly

2

662

B

Figure 3 (Color online) Diagram and alignment of *zfp36l2* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2). A, Schematic map of relationships between the reference genome, Sanger sequences and contigs of the *de novo* assembly. B, Sequence alignment for *zfp36l2*. Blue denotes loci from WCC and different from RCC; yellow variants represent loci from RCC and different from WCC; red represents mutations that differ from both WCC and RCC.

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

Table 2 *De novo* assembled transcripts for four individuals

	WCC	RCC	WR-F_1	WR-F_2
Number of unigene	64,972	73,421	60,586	66,619
Number of reads	41,116,994	52,040,914	41,344,836	51,490,270
N50	841	1,154	819	850

Table 3 Gene numbers of each variation pattern in hybrids using white crucian carp as a reference

Categories	Chimeric genes	Genes with specific mutations	Genes of biparental origin
Patterns	1–3	4–6	7–8
Gene number (WR-F_1)	1,278	463	2,086
Percentage (WR-F_1)	19.04%	6.90%	31.08%
Gene number (WR-F_2)	187	226	1,620
Percentage (WR-F_2)	4.18%	5.05%	36.20%

(*zfp36l2*), SET domain containing 5 (*setd5*), and some other genes showed chimeric fragments consisting of continuous, alternating variations from parent-specific variants with or without offspring-specific mutations (Figure 3, Figures S1–S16 in Supporting Information). In addition, *zfp36l2*, *setd5*, somatostatin receptor 2 (*sst2*), elastin microfibril interface (*emilin1*), ZFP36 ring finger protein 1 (*zfp36*), apolipoprotein B (*apob*), and small nuclear ribonucleoprotein U1 subunit 70 (*snrnp70*) showed chimeric fragments consisting of continuous, alternating variations from parent-specific variants with or without offspring-specific mutations both in the WR-F_1 and WR-F_2 (Figure 3, Figures S1–S6 in Supporting Information). Analysis is showed as follows: ring finger protein 139 (*rnf139*) had many locus mutations with the deletion of some bases (Figure S17 in Supporting Information); desmoplakin a (*dspa*) has one fragment from WCC and exhibited some mutations (Figure S18 in Supporting Information); RNA binding motif protein 25 (*rbm25*) was exclusively inherited from RCC (Figure S19 in Supporting Information); mannosidase endo-alpha (*manea*), and serine-aspartate repeat-containing protein I-like (*loc107984728*), were exclusively inherited from RCC (Figures S20 and S21 in Supporting Information). Moreover, we mapped these validated genes from the hybrids transcriptome to homologous genes in the goldfish reference genome to determine whether any structural changes existed (Figure 3, Figures S1–S21 in Supporting Information). The results showed the structure of these validated genes is consistent with the goldfish reference genome; there is no structural changes in these validated genes.

Enrichment analysis of chimeric and mutant genes in WR-F_1

In WR-F_1, GO analysis classified 1,731 annotated chimeric and mutant genes into 15, 16, and 23 main subgroups (at level 2) about the cellular component, molecular function, and biological process categories, respectively (Figure 4). Among the biological process category, enrichment was observed in metabolic processes (GO:0008152), immune system processes (GO:0002376), and developmental processes (GO:0032502), which may be related to metabolism, immunology, and development in WR-F_1.

In contrast to GO enrichment, pathway enrichment summarizes the complex networks of genes in maps for improved visualization and comprehension. Multiple pathways, including the Metabolic pathways, Biosynthesis of secondary metabolites, PI3K-Akt signaling pathway, T cell receptor signaling pathway and NF-kappa B signaling pathway, were clearly enriched in chimeric and mutant genes (Table 4).

DISCUSSION

The significance of hybridization

Hybridization is a significant and useful strategy for creating new types of fish stocks in genetic breeding. This process combines the genetic material from parental species into hybrids offspring, resulting in changes in the gene regulation and expression in the hybrid progeny (Mallet, 2007; Liu, 2010; Stelkens and Seehausen, 2009). Thus, the hybrids can exhibit many advantageous traits, such as rapid growth, high harvest rates, and strong environmental tolerances (Bartley et al., 2000; Li et al., 2013). Hybridization can be divided into distant hybridization (above-specific or interspecific crossing) and intraspecific hybridization. At the molecular level, distant hybridization leads to the emergence of chimeric genes and mutations of orthologous genes (Liu et al., 2016). As reported previously, the allopolyploid fish offspring has many chimeric genes (>9%) and mutations of orthologous gene. The number of chimeric and mutant genes is steady in different offspring (Liu et al., 2016). According to our results in the present study, intraspecific hybridization can result in chimeric genes, and both the first generation and the second

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

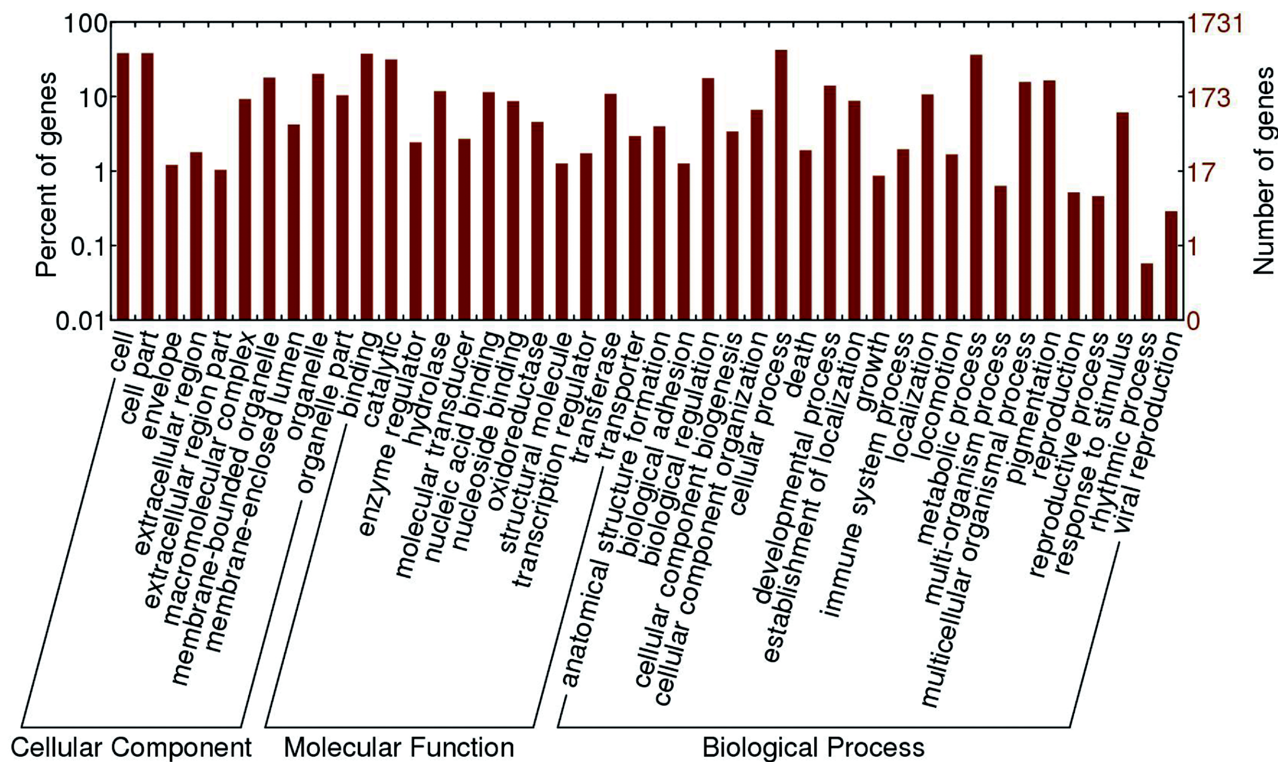

Figure 4 Gene ontology classification of chimeric and mutant genes in WR-F_1. Chimeric and mutant genes are annotated in three categories: biological processes, molecular function, and cellular component.

Table 4 Pathways enriched with chimeric and mutant genes in WR-F_1

Pathways	Pathway ID	Number of annotated chimeric and mutant genes in WR-F_1 (1,731)
Metabolic pathways	KO01100	203 (11.73%)
Biosynthesis of secondary metabolites	KO01110	70 (4.04%)
Pathways in cancer	KO05200	51 (2.95%)
Biosynthesis of antibiotics	KO01130	47 (2.72%)
Protein processing in endoplasmic reticulum	KO04141	45 (2.60%)
Microbial metabolism in diverse environments	KO01120	43 (2.48%)
PI3K-Akt signaling pathway	KO04151	41 (2.37%)
NF-kappa B signaling pathway	KO04064	20 (1.16%)
T cell receptor signaling pathway	KO04660	18 (1.04%)
Toll-like receptor signaling pathway	KO04620	17 (0.98%)

generation have chimeric genes. But different from distant hybridization, the chimeric genes of WR-F_2 (4.17%) were lower than WR-F_1 (19.04%). According to this phenomenon, we presume that the steadiness of chimeric genes in distant hybridization is necessary for progeny surviving, because the offspring is hard to survive in distant hybridization; however, the offspring is easy to survive in intraspecific hybridization, the less chimeric genes in WR-F_2 may suggest the genome of the hybrids was more stable with the increase of generation. In a word, whether in distant hybridization or in intraspecific hybridization, the existence of chimeric genes is crucial. In addition, in our research, the patterns of chimeric genes (8 patterns) are less than that of previous study in allopolyploid fish (18 patterns) (Liu et al., 2016). It may be the otherness of parental genome in distant hybridization is higher than that in intraspecific hybridization, so it caused more violent genomic oscillations.

Chimeric gene in hybrid lineage of *Carassius auratus cuvieri*×*Carassius auratus* red var

The genetic mechanisms underlying the creation of hybrid traits in hybrids are poorly understood. The basic features of the complex processes include the distribution of transpo-

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

sable elements, gene insertion or loss, and chromosomal rearrangement (Nei and Nozawa, 2011; Rieseberg, 2001). Changes of sequence in promoters, or other regulatory sequences, may also cause changes in gene expression and contributes to the evolution of new traits (Chen, 2007). Both chimeric and mutant genes are able to produce structural changes in genes which could reduce the activity or fidelity of an encoded enzyme, by influencing normal transcriptional activity; thus, chimeric genes may be responsible for some of the novel traits observed in hybrids (Koyama et al., 2007).

In this study, we identified 1,731 chimeric and mutant genes in WR-F_1 using GO and pathway enrichment. GO enrichment analysis revealed that some of the chimeric and mutant genes were related to metabolism, immune and developmental processes, such as metabolic process (GO:0008152), primary metabolic process (GO:0044238), pigment metabolic process (GO:0042440), immune system process (GO:0002376), immune response (GO:0006955), developmental process (GO:0032502), embryonic development (GO:0009790), growth (GO:0040007), pigmentation (GO:0043473). Interestingly, ''metabolic pathways'', ''biosynthesis of secondary metabolites'', ''T cell receptor signaling pathway'', ''NF-kappa B signaling pathway'', "Toll-like receptor signaling pathway" and some metabolism-specific and immune-specific pathways are annotated in the KEGG (Table 4). We measured the body weight of WR-F_1, WCC and RCC for two years (n=50 for each fish). The mean body weight of two-year-old WR-F_1 was 750 g, which was closed to WCC (762 g) but significantly higher than the RCC (292 g) ($P<0.05$) (Table 1). We hypothesize that some chimeric and mutant genes, such as *sst2* and *igfals* (Figures S2 and S10 in Supporting Information), which were related to growth and development (Lopez et al., 1997; Voss, 2010) may have contributed to the rapid growth of WR-F_1. In conclusion, these enrichment analyzes provide a rich and detailed resource for science researchers to probe the relationship between hybrid traits and chimeric genes.

Potential mechanisms related to chimeric genes

Structural changes in the genome might severely constrain the survival of hybrid offspring. Rapid gene loss, genome restructuring, and altered patterns of gene expression are common in the genomes of newly-formed natural or artificial hybrids (Adams and Wendel, 2005; Hegarty et al., 2012; Madlung et al., 2005; Schnable et al., 2011; Xiong et al., 2011). As reported previously, a high level of genomic restructuring occurred in allopolyploid fish offspring, including genetic recombination, offspring-specific mutation and significant alteration of gene expression (Liu et al., 2016). In addition, some hybrid plants, such as *Brassica* (Gaeta et al., 2007) and wheat (Brenchley et al., 2012; Feldman et al., 1997), also exhibit relatively high levels of genomic rearrangements.

In the present study, we found chimeric genes (WR-F_1:19.04%, WR-F_2:4.18%) in the hybrids which were produced from close hybridization of the WCC (♀) and RCC (♂) genomes, to a certain extent, representing a "genetic melting" of the two organisms. WCC and RCC belong to different breed. Using them as parents for the hybridization, it increases nuclear-nuclear and nuclear-cytoplasmic incompatibility between WCC and RCC, which easily gives rise to drastic genomic imbalances in the hybrids.

The presence of chimeric and mutated genes may be caused by the struggling for survival in the initial stage of hybridization. The high frequency of chimeric and mutated genes may be driven by erroneous DNA excision between homologous parental genes, large-scale DNA repair via recombination or non-homologous end-joining, or even transposon activity (Bao and Yan, 2012; Fedoroff, 2012; Levin and Moran, 2011; Liu et al., 2016; Wang et al., 2006). Abnormalities of DNA or RNA repair, such as dysfunction of *RAD* (Delmas et al., 2009; Huang et al., 2009) or other genes and pathways (Cui et al., 2012; Kobayashi et al., 2008; van Gent et al., 2001; Zhao et al., 2007), may also have contributed to the high rate of homologous recombination in the hybrids. Some transposable elements are capable of mobilizing an adjacent sequence, and new chimeric genes were generated (Collins and Bell, 2004; Shapiro, 2005; Yang et al., 2008). The hybrids combined whole genomes from WCC and RCC, which may have allowed the reciprocal exchange of homologous chromosomes, and gene conversion, to occur more easily, leading to a large number of genetic changes in the hybrids. Furthermore, replication slippage, or the imprecise cutting of an unpaired duplication during large-loop mismatch repair, may represent additional mechanisms for the formation of chimeric genes (Rogers et al., 2009). In addition, double-strand break (DSB) is more likely to occur at tandem repeats locus during the process of DNA replication, thus, a chimera is formed by recombination between different sequences during the process of DSB repair (Ye et al., 2017) However, the causative mechanism(s) involved could not be identified by our present analyses.

In conclusion, hybridization offers a means, unlike genome-editing technique which acts on one or a few genes, provides genetic variation at hundreds or thousands of genes. The chimeric and mutated genes are a common phenomenon whether in distant hybridization or in close hybridization. And the chimeric and mutated genes are likely to be related to hybrid traits.

MATERIALS AND METHODS

Fish cross and sampling

WCC, RCC, WR-F_1, and WR-F_2 (1 year old) were collected from the State Key Laboratory of Developmental Biology of

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

Freshwater Fish, Hunan Normal University, Changsha, China. The protocols for crosses and culturing were described previously (Wang et al., 2015). All fish were euthanized using 2-phenoxyethanol (Sigma, USA) before being dissected. Livers were removed surgically, and stored in RNAlater (Invitrogen, USA) following the manufacturer's instructions for total RNA extraction. Animal experimenters were certified under a professional training course for laboratory animal practitioners held by the Institute of Experimental Animals, Hunan Province, China.

Measurement of body weight for WCC, RCC, and WR-F_1

In May of 2014, 200 WCC, 200 RCC, and 200 WR-F_1 were selected and reared in a 500-m^2 pond. After feeding for 3 months, body weights (BWs) of WCC, RCC, and WR-F_1 were measured for two years (n=50 for each fish).

RNA isolation, cDNA library construction, sequencing and quality control

Total RNA was isolated from liver tissues of each type of fish using the TRIzol Reagent (Invitrogen) in accordance with the manufacturer's protocol. Total RNA samples exhibited 28S:18S ratios ranging from 1.8–2.0 while the integrity of the total RNA ranged from 8.0 to 10.0. Next, we constructed four cDNA libraries which represented genotype of each type of fish (three WCC, three RCC, three WR-F_1, and three WR-F_2) using 2 μg mRNA. Then, each library was sequenced using an Illumina HiSeq™ 2500. Raw reads were filtered using the Trimmomatic software (Babraham Bioinformatics, USA) (Xu et al., 2015) to obtain the paired-end clean reads. *De novo* assembly of transcriptomes was then carried out with a short-reads assembly program (Trinity, USA) (Dion-Côté et al., 2014), using three independent software modules known as Inchworm, Chrysalis, and Butterfly. Principal component analysis (PCA) of three liver transcriptomes was then applied to examine the contribution of each transcript to the separation of the classes (Anders and Huber, 2010; Reeb and Steibel, 2013).

Contig annotation was performed using five public databases (nonredundant (Nr); Swiss-Prot; Kyoto Encyclopedia of Genes and Genomes (KEGG); Clusters of Orthologous Groups (COG) and Gene Ontology (GO)). First, BLASTX alignment (e-value$\leq 1\times10^{-6}$) was used between contigs and protein databases, and the best-aligned results were used to decide the sequence direction of contigs. After screening the sequences (alignment length≤100 bp), the accession numbers of the genes were obtained from the BLASTX results. Then, GO terms of annotation sequences were obtained through Ensembl BioMart (Flicek et al., 2013). WEGO software was used to analyze the GO annotation (Ye et al., 2006). The KEGG pathway was then used to identifiy the relevant biological pathways (Kanehisa et al., 2008). Then, to identify putative orthologs between hybrids and the original parents, the sequences from WCC, RCC, and WR-F_1 assembled sequences were aligned with the reciprocal BLAST (BLASTN) hit with an e-value of 1×10^{-20} (Blanc and Wolfe, 2004); the sequences from WCC, RCC, and WR-F_2 assembled sequences were aligned with the reciprocal BLAST (BLASTN) hit with an e-value of 1×10^{-20} (Blanc and Wolfe, 2004). Three sequences were defined as orthologs if each of them was the best hit of the other and if the sequences were aligned over 300 bp. Meanwhile, the nucleotide sequence was aligned using the BioEdit program (version 7.0.9) (Hall, 1999).

Variation and detection of chimeric patterns

High-quality reads were remapped to goldfish reference genome (Liu et al., 2016) with Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009) to detect variants among WCC, RCC, WR-F_1 and WR-F_2, and their distributions. Because divergence within most shared copies of both WCC and RCC was less than 5%, the maximum mismatch of hits was set as 5 per 100 bp read. This parameter setting (maximum mismatch of 5%) ensured that reads from half of the orthologs were mapped to the genome as the mapping ratio of the red crucian carp is about 50%. After obtaining the BAM files, we recorded the mapped region of each sample on the reference genome. Variations from regions overlapped by both parents and hybrids were extracted from the alignments using both *mpileup* from the SAMtools package (Li et al., 2009), and the GATK (DePristo et al., 2011; McKenna et al., 2010; Van et al., 2013) pipeline for RNA-seq. Candidate variations were filtered based on a variation-quality score ≥20, and depth >3 reads. VCFtools (Danecek et al., 2011) was used to compare variations among both parents and hybrids identified by both methods.

The distribution patterns of variations in WR-F_1 and WR-F_2 compared to the parental were analyzed and the distributions of chimeric loci were retained for downstream analysis. Mutation patterns were defined in the following manner: first, as chimeric-variants, single or multiple fragments consisting of continuous, alternating variations from parents-specific variants; second, offspring-specific-variants, including locus mutations, DNA insertions or DNA deletions; third, as parent-variants, where the sequence was the same as one parent but different from the other parent. Within a gene-region, several segmental fragments potentially consisted of different exons. Thus, offspring mutations were associated with segments. If parent-specific variations aligned alternately within a continuous fragment, the gene was classified as a chimera containing a parental crossing hotspot (patterns 1–3). Genes from the offspring with spe-

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

cific variations but no chimeric patterns were identified as having mutations (patterns 4–6). Genes with single or multiple exons derived only from one parent were classified as being of maternal-origin or paternal-origin (patterns 7 and 8). Only genes with fragments per kilobase of transcript per million mapped reads (FPKM) values >0.3 were classified by pattern. Redundant genes (the same gene name and pattern of variation) in the red crucian carp-based analyses were removed.

PCR validation for 30 chimeric genes

Thirty chimeric genes were chosen randomly to validate the chimeric pattern. Genomic DNAs were extracted from the liver tissue of three WCC, three RCC, three WR-F_1, and three WR-F_2 according to Yue and Orban (Yue and Orban, 2005). Polymerase chain reaction (PCR) primers were designed based on the conserved regions of homologous gene from WCC, RCC, and WR (Table S1 in Supporting Information). The amplification reaction mixture (50 μL) consisted of 40 ng genomic DNA, 3.0 mmol L^{-1} $MgCl_2$, 0.4 mmol L^{-1} of each dNTP, 0.8 μmol L^{-1} of each primer, 1× PCR buffer, and 2.50 U Taq polymerase (TaKaRa, Dalian). The thermal program consisted of an initial denaturation step of 94°C for 5 min, followed by 35 cycles (94°C for 30 s, 51°C–59°C for 30 s, and 72°C for 1–5 min) and a final extension step of 72°C for 10 min. The PCR products were separated on a 1.2% agarose gel, purified using a Gel Extraction Kit (Sangon Biotech Co., Ltd., Shanghai), ligated into a pMD18-T vector, and transferred into *Escherichia coli* DH5α. The positive clones were then sequenced (10 clones for each PCR fragment). To determine sequence homology and variation, the sequences were aligned using BioEdit (Hall, 1999). In addition, those validated genes identified from goldfish reference-genome were combined to detect the structural changes.

Analysis of chimeric and mutant genes in WR-F_1

To analyze the function of chimeric and mutant genes in WR-F_1, we investigated the biological processes and pathways in which they were involved. WEGO software was used to classify the GO terms of the chimeric and mutant genes, while the KEGG pathway was used to identify the most relevant biological pathways.

Data availability statement

The complete clean reads for these libraries have been uploaded to the NCBI Sequence Read Archive site (http://www.ncbi.nlm.nih.gov/sra/; accession Nos. SRX1999729, SRX3041132, SRX3041134, and SRR5936567).

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

Acknowledgements *This work was supported by National Natural Science Foundation of China (31210103918, 31430088, 91631305), the earmarked fund for China Agriculture Research System (CARS-45), 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 and Natural Science Foundation of Hunan Province (14JJ2062).*

Adams, K.L., and Wendel, J.F. (2005). Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8, 135–141.
Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11, R106.
Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O., and Zimmer, E. (1980). Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc Natl Acad Sci USA 77, 7323–7327.
Bao, J., and Yan, W. (2012). Male germline control of transposable elements. Biol Reproduct 86, 1–14.
Bartley, D.M., Rana, K., and Immink, A.J. (2000). The use of inter-specific hybrids in aquaculture and fisheries reviews in Fish. Rev Fish Biol Fisheries 10, 325–337.
Blanc, G., and Wolfe, K.H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell 16, 1667–1678.
Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G.L.A., D'Amore, R., Allen, A.M., McKenzie, N., Kramer, M., Kerhornou, A., Bolser, D., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature 491, 705–710.
Chen, Z.J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu Rev Plant Biol 58, 377–406.
Collins, S., and Bell, G. (2004). Phenotypic consequences of 1,000 generations of selection at elevated CO_2 in a green alga. Nature 431, 566–569.
Cui, P., Lin, Q., Ding, F., Hu, S., and Yu, J. (2012). The transcript-centric mutations in human genomes. Genomics Proteomics Bioinformatics 10, 11–22.
Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M. A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics 27, 2156–2158.
Delmas, S., Shunburne, L., Ngo, H.P., and Allers, T. (2009). Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. PLoS Genet 5, e1000552.
DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43, 491–498.
Dion-Côté, A.M., Renaut, S., Normandeau, E., and Bernatchez, L. (2014). RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. Mol Biol Evol 31, 1188–1199.
Fedoroff, N.V. (2012). Transposable elements, epigenetics, and genome evolution. Science 338, 758–767.
Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A.A., and Vega, J.M. (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: a possible mechanism for differentiation of homoeologous chromosomes. Genetics 147, 1381–1387.
Flicek, P., Amode, MR., Barrell, D., Beal, K., Billis K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2013). Ensembl 2014. Nucleic Acids Res 41, D49–D55.
Gaeta, R.T., Pires, J.C., Iniguez-Luy, F., Leon, E., and Osborn, T.C. (2007).

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

Genomic changes in resynthesized brassica napus and their effect on gene expression and phenotype. Plant Cell 19, 3403–3417.
Galili, G., Levy, A.A., and Feldman, M. (1986). Gene-dosage compensation of endosperm proteins in hexaploid wheat *Triticum aestivum*. Proc Natl Acad Sci USA 83, 6524–6528.
Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41, 95–98.
Hegarty, M.J., Abbott, R.J., and Hiscock, S.J. (2012). Allopolyploid Speciation in Action: The Origins and Origins and Evolution of Senecio cambrensis. (Berlin: Springer Berlin Heidelberg), pp. 245–270.
Hu, W., Timmermans, M.C., and Messing, J.(1998). Interchromosomal recombination in *Zea mays*. Genetics 150, 1229–1237.
Huang, J., Huen, M.S.Y., Kim, H., Leung, C.C.Y., Glover, J.N.M., Yu, X., and Chen, J. (2009). RAD18 transmits DNA damage signalling to elicit homologous recombination repair. Nat Cell Biol 11, 592–603.
Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., et al. (2008). KEGG for linking genomes to life and the environment. Nucleic Acids Res 36, D480–D484.
Kenton, A., Parokonny, A.S., Gleba, Y.Y., and Bennett, M.D. (1993). Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Genet Genomics 240, 159–169.
Kobayashi, J., Iwabuchi, K., Miyagawa, K., Sonoda, E., Suzuki, K., Takata, M., and Tauchi, H. (2008). Current topics in DNA double-strand break repair. J Radiat Res 49, 93–103.
Koyama, H., Ito, T., Nakanishi, T., and Sekimizu, K. (2007). Stimulation of RNA polymerase II transcript cleavage activity contributes to maintain transcriptional fidelity in yeast. Genes Cells 12, 547–559.
Levin, H.L., and Moran, J.V. (2011). Dynamic interactions between transposable elements and their hosts. Nat Rev Genet 12, 615–627.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.
Li, L., Zhong, Z.Z., Zeng, M., Liu, S.J., Zhou, Y., Xiao, J., Wang, J., and Liu, Y. (2013). Comparative analysis of intermuscular bones in fish of different ploidies. Sci China Life Sci 56, 341–350.
Liu, S.J. (2010). Distant hybridization leads to different ploidy fishes. Sci China Life Sci 53, 416–425.
Liu, S., Liu, Y., Zhou, G., Zhang, X., Luo, C., Feng, H., He, X., Zhu, G., and Yang, H. (2001). The formation of tetraploid stocks of red crucian carp×common carp hybrids as an effect of interspecific hybridization. Aquaculture 192, 171–186.
Liu, S., Qin, Q., Xiao, J., Lu, W., Shen, J., Li, W., Liu, J., Duan, W., Zhang, C., Tao, M., et al. (2007). The formation of the polyploid hybrids from different subfamily fish crossings and its evolutionary significance. Genetics 176, 1023–1034.
Liu, S., Luo, J., Chai, J., Ren, L., Zhou, Y., Huang, F., Liu, X., Chen, Y., Zhang, C., Tao, M., et al. (2016). Genomic incompatibilities in the diploid and tetraploid offspring of the goldfish×common carp cross. Proc Natl Acad Sci USA 113, 1327–1332.
Lopez, F., Estève, J.P., Buscail, L., Delesque, N., Saint-Laurent, N., Théveniau, M., Nahmias, C., Vaysse, N., and Susini, C. (1997). The tyrosine phosphatase SHP-1 associates with the sst2 somatostatin receptor and is an essential component of sst2-mediated inhibitory growth signaling. J Biol Chem 272, 24448–24454.
Madlung, A., Tyagi, A.P., Watson, B., Jiang, H., Kagochi, T., Doerge, R.W., Martienssen, R., and Comai, L. (2005). Genomic changes in synthetic *Arabidopsis* polyploids. Plant J 41, 221–230.
Mallet, J. (2007). Hybrid speciation. Nature 446, 279–283.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297–1303.
Nei, M., and Nozawa, M. (2011). Roles of mutation and selection in speciation: from hugo de vries to the modern genomic era. Genome Biol Evol 3, 812–829.
Ozkan, H., Levy, A.A., and Feldman, M. (2001). Allopolyploidy-induced rapid genome evolution in the wheat (Aegilops-Triticum) group. Plant Cell 13, 1735–1747.
Qin, Q., He, W., Liu, S., Wang, J., Xiao, J., and Liu, Y. (2010). Analysis of 5S rDNA organization and variation in polyploid hybrids from crosses of different fish subfamilies Journal of experimental zoology Part B. Mol Dev Evol 314, 403–411.
Qin, Q., Wang, J., Hu, M., Huang, S., and Liu, S. (2016). Autotriploid origin of *Carassius auratus* as revealed by chromosomal locus analysis. Sci China Life Sci 59, 622–626.
Reeb, P.D., and Steibel, J.P. (2013). Evaluating statistical analysis models for RNA sequencing experiments. Front Genet 4, 178.
Rieseberg, L.H. (2001). Chromosomal rearrangements and speciation. Trends Ecol Evol 16, 351–358.
Rogers, R.L., Bedford, T., and Hartl, D.L. (2009). Formation and longevity of chimeric and duplicate genes in *Drosophila melanogaster*. Genetics 181, 313–322.
Schnable, J.C., Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci USA 108, 4069–4074.
Shapiro, J.A. (2005). A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. Gene 345, 91–100.
Soltis, D.E., and Soltis, P.S. (1995). The dynamic nature of polyploid genomes. Proc Natl Acad Sci USA 92, 8089–8091.
Stelkens, R., and Seehausen, O. (2009). Genetic distance between species predicts novel trait expression in their hybrids. Evolution 63, 884–897.
Van, G.A., Carneiro, M.O., Hartl, C., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protocols Bioinform 43, 11–33.
van Gent, D.C., Hoeijmakers, J.H.J., and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. Nat Rev Genet 2, 196–206.
Voss, K.A. (2010). A new perspective on deoxynivalenol and growth suppression. Toxicol Sci 113, 281–283.
Wang, H.C., Chou, W.C., Shieh, S.Y., and Shen, C.Y. (2006). Ataxia telangiectasia mutated and checkpoint kinase 2 regulate BRCA1 to promote the fidelity of DNA end-joining. Cancer Res 66, 1391–1400.
Wang, J., Xiao, J., Zeng, M., Xu, K., Tao, M., Zhang, C., Duan, W., Liu, W B., Luo, K.K., Liu, Y., et al. (2015). Genomic variation in the hybrids of white crucian carp and red crucian carp: evidence from ribosomal DNA. Sci China Life Sci 58, 590–601.
Wang, S., Ye, X., Wang, Y., Chen, Y., Lin, B., Yi, Z., Mao, Z., Hu, F., Zhao, R., Wang, J., et al. (2017). A new type of homodiploid fish derived from the interspecific hybridization of female common carp×male blunt snout bream. Sci Rep 7, 4189.
Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genetics 10, 57–63.
Wendel, J.F. (2000). Genome evolution in polyploids. Plant Mol Biol 42, 225–249.
Xiao, J., Kang, X., Xie, L., Qin, Q., He, Z., Hu, F., Zhang, C., Zhao, R., Wang, J., Luo, K., et al. (2014). The fertility of the hybrid lineage derived from female *Megalobrama amblycephala*×male *Culter alburnus*. Animal Reproduct Sci 151, 61–70.
Xiong, Z., Gaeta, R.T., and Pires, J.C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. Proc Natl Acad Sci USA 108, 7908–7913.
Xu, K., Wen, M., Duan, W., Ren, L., Hu, F., Xiao, J., Wang, J., Tao, M., Zhang, C., Wang, J., et al. (2015). Comparative analysis of testis transcriptomes from triploid and fertile diploid cyprinid fish. Biol Reproduct 92, 95.
Yang, S., Arguello, J.R., Li, X., Ding, Y., Zhou, Q., Chen, Y., Zhang, Y.,

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7

Zhao, R., Brunet, F., Peng, L., et al. (2008). Repetitive element-mediated recombination as a mechanism for new gene origination in *Drosophila*. PLoS Genet 4, 78–87.
Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., Wang, J., Li, S., Li, R., Bolund, L., et al. (2006). WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34, W293–W297.
Ye, L., Jiao, N., Tang, X., Chen, Y., Ye, X., Ren, L., Hu, F., Wang, S., Wen, M., Zhang, C., et al. (2017). Chimeras linked to tandem repeats and transposable elements in tetraploid hybrid fish. Mar Biotechnol 19, 401–409.
Yue, G.H., and Orban, L. (2005). A simple and affordable method for high-throughput DNA extraction from animal tissues for polymerase chain reaction. Electrophoresis 26, 3081–3083.
Zhang, Z.H., Chen, J., Li, L., Tao, M., Zhang, C., Qin, Q.B., Xiao, J., Liu, Y., and Liu, S.J. (2014). Research advances in animal distant hybridization. Sci China Life Sci 57, 889–902.
Zhao, G.Y., Sonoda, E., Barber, L.J., Oka, H., Murakawa, Y., Yamada, K., Ikura, T., Wang, X., Kobayashi, M., Yamamoto, K., et al. (2007). A critical role for the ubiquitin-conjugating enzyme Ubc13 in initiating homologous recombination. Mol Cell 25, 663–675.

SUPPORTING INFORMATION

Figure S1 Diagram and alignment of *setd5* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S2 Diagram and alignment of *sst2* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S3 Diagram and alignment of *emilin1* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S4 Diagram and alignment of *zfp36* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S5 Diagram and alignment of *apob* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S6 Diagram and alignment of *snrnp70* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1 and WR-F_2).

Figure S7 Diagram and alignment of *npc1* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S8 Diagram and alignment of *minpp1a* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S9 Diagram and alignment of *sall1a* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S10 Diagram and alignment of *igfals* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S11 Diagram and alignment of *c3ar1* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S12 Diagram and alignment of *znf638* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S13 Diagram and alignment of *acin1b* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S14 Diagram and alignment of *march7* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR).

Figure S15 Diagram and alignment of *zc3h12a* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S16 Diagram and alignment of *rnf225* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S17 Diagram and alignment of *rnf139* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S18 Diagram and alignment of *dspa* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S19 Diagram and alignment of *rbm25* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F1).

Figure S20 Diagram and alignment of *manea* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Figure S21 Diagram and alignment of *loc107984728* from white crucian carp (WCC), red crucian carp (RCC) and hybrids (WR-F_1).

Table S1 Primers sequences used in the PCR

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

Downloaded to IP: 58.20.50.216 On: 2018-06-13 17:23:34 http://engine.scichina.com/doi/10.1007/s11427-017-9306-7