Comparison of diploid and triploid Carassius auratus provides insights into adaptation to environmental change

Ren Li, Gao Xin, Yang Conghui, Tan Hui, Cui Jialin, Wang Shi, Li Wuhui, Zhang Chun, Tao Min, Qin Qinbo and Liu Shaojun

Citation: SCIENCE CHINA Life Sciences ; doi: 10.1007/s11427-017-9358-7

View online: http://engine.scichina.com/doi/10.1007/s11427-017-9358-7

Published by the Science China Press

SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-017-9358-7

Comparison of diploid and triploid *Carassius auratus* provides insights into adaptation to environmental change

Li Ren[1,2†], Xin Gao[1,2†], Conghui Yang[1,2†], Hui Tan[1,2], Jialin Cui[1,2], Shi Wang[1,2], Wuhui Li[1,2], Chun Zhang[1,2], Min Tao[1,2], Qinbo Qin[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 April 19, 2018; accepted May 26, 2018; published online August 30, 2018

Focusing on adaptation of aquatic organisms, especially fish, can help elucidate complex dynamics in freshwater ecology. The differences in genetic and epigenetic regulation between diploid and triploid *Carassius auratus* affect survival under eutrophication. To identify the underlying mechanisms that lead to better adaption of triploids than diploids, we compared mRNA and microRNA (miRNA) expressions in liver tissue of diploid and triploid individuals obtained from the Dongting lake water system in central China. Differential expression analysis revealed that 566 transcripts were significantly up-regulated, whereas 758 were down-regulated in triploids; of these differentially expressed transcripts, 33 transcripts including *cacna1d*, *nfkb2*, *hspa1* and *fgfr4* were involved in the MAPK signaling pathway, and eight transcripts were determined to be regulated by seven miRNAs. Additionally, four of 25 differential expressed (DE) transcripts (*mhc1*, *irf7*, *nfkb2* and *pik3c*) involving the viral carcinogenesis pathway were regulated by four miRNAs. Furthermore, genetic polymorphisms analysis showed that more heterozygous mutations were detected in triploids than diploids. The d*N*/d*S* results revealed that 21 genes were under positive selection (d*N*/d*S*>1) in *C. auratus* complex. We hypothesize that these changes related to genetic and epigenetic regulation may be caused by abiotic stresses, and facilitate adaptation to environmental changes.

triploid, adapting, positive selection, differential expression

Citation: Ren, L., Gao, X., Yang, C., Tan, H., Cui, J., Wang, S., Li, W., Zhang, C., Tao, M., Qin, Q., and Liu, S. (2018). Comparison of diploid and triploid *Carassius auratus* provides insights into adaptation to environmental change. Sci China Life Sci 61, https://doi.org/10.1007/s11427-017-9358-7

INTRODUCTION

Industrial waste, sewage effluent, agricultural run-off and decomposition in biological wastes may cause changes in water composition (Jin et al., 2017). Increasing concentrations of the total nitrogen and phosphorus in the Dongting lake water system and accompanying increase of phytoplankton abundance ultimately resulted in ecological modifications, including of the types and proportions of fish populations (Lu et al., 2007; Wang et al., 2014). *Carassius auratus* is a common economically important freshwater fish in China, and is widely used to monitor environmental quality (Qin et al., 2016; Xiao et al., 2011; Yoshida et al., 2005; Zhu et al., 2008). From 2005 to 2009, the ratio of diploid to triploid populations (1.00:1.56) gradually reduced (Xiao et al., 2011), reflecting that the triploid populations were more tolerant to environmental changes. However, the underlying mechanism is poorly known and needs to be elucidated.

The phenomenon of heterosis has been observed in triploid plants (Hochholdinger and Hoecker, 2007; Yao et al., 2013) and fishes (Arkhipchuk and Garanko, 2005; Chen et al., 2009; Tiwary et al., 2004). Heterosis leads to shorter trunks and longer tails in triploid *Gasterosteus aculeatus* (Swarup,

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

1959), smaller heads in triploid *Ictalurus punctatus* (Wolters et al., 1982); and affects a variety of traits, including body depth, predorsal length, width of gape, fin ray, and scale counts, in triploid *Ctenopharyngodon idella* (Bonar et al., 2010). However, diploid and triploid *C. auratus* do not differ significantly in their morphology (Xiao et al., 2011). Therefore, molecular level analysis was performed to investigate the differences between diploid and triploid *C. auratus*.

Triploid individuals can be obtained in different ways including cold-shock treatment (Swarup, 1956) and hybridization (Chen et al., 2009). Of these, the type of triploid (autotriploids and allotriploids, respectively) was inferred based on the heterozygosity and genetic polymorphism (White, 1970). Enzyme locus analysis revealed a higher frequency of heterozygotes in allotriploid compared with diploid in rainbow trout (*Salmo gairdneri*), which led to increasing of developmental stability (Stanley et al., 1984). A similar phenomenon was observed in *C. auratus* based on *tf* allele analysis (Liu et al., 2017). However, diploid *Lilium lancifolium* Thunb. has higher genetic diversity than autotriploids, which was determined based on SSRP and IRAP molecular markers (Lee et al., 2016). One study on the *C. auratus* complex concluded that triploid individuals may have undergone multiple independent polyploidy origins from sympatric diploids (Liu et al., 2017), whereas another study based on micro chromosome analysis revealed that triploid populations may have originated from hybridization or gynogenesis (Xiao et al., 2011).

Here, we conducted a novel integrated analysis of mRNA and miRNA expression to investigate triploid *C. auratus* adaptation to the constantly changing ecological environment of the Dongting lake. We identified numerous mRNAs and miRNAs that were differentially expressed (DE) between diploids and triploids based on the mRNA-seq, real-time quantitative PCR (qRT-PCR) and miRNA-seq. Additionally, we inferred regulatory networks between mRNAs and miRNAs for investigating the underlying molecular mechanism that facilitates adaptation to water change. Furthermore, we explored the distribution of heterozygous loci and positively selected genes in diploids and triploids. The differences we identified in genetic and epigenetic regulation may be responsible for the increased immunity and higher survival rate of triploid *C. auratus*.

RESULTS

Differential mRNA expression between diploids and triploids

To examine changes in the global transcriptome profiles between diploid and triploid *C. auratus*, six liver transcriptomes were obtained with paired-end sequencing (PE×125). The transcriptome data were submitted to NCBI (accession number: SRR6001365, SRR6001366, SRR6001367, SRR6001371, SRR6001437, and SRR6001438). After initial adapter trimming and quality filtering, 346.29 million cleaned reads from the six libraries were collected (Table S1 in Supporting Information). When the cleaned reads (46.11 Gb, 346.3 million reads) were mapped against the *C. auratus* red var. reference genome (http://rd.biocloud.org.cn/) using Tophat (Table S2 in Supporting Information), 29,380 and 29,386 transcripts were detected in diploid and triploid *C. auratus*, respectively (Figure S1A in Supporting Information). Further investigation of the DE data revealed 566 and 758 up-regulated transcripts in triploids and diploids, respectively, based on the parameters FDR≤0.05 and $\log_2$FC>2 (Figure S1B in Supporting Information).

To investigate the underlying function of DE transcripts in diploid and triploid *C. auratus*, we performed GO and KEGG functional enrichment analysis on the 1,324 DE transcripts (Figure S2 in Supporting Information). Of the 1,324 DE transcripts, eight transcripts were associated with negative regulation of transcription from RNA polymerase II promoter (GO: 0000122), whereas six transcripts were associated with signal transduction (GO: 0007165) (Figure S2A in Supporting Information). KEGG analysis revealed that these genes were distributed in 308 pathways. The top 20 pathways are shown in Figure S2B in Supporting Information. Thirty-three DE transcripts were enriched in the mitogen-activated protein kinase (MAPK) signaling pathway (ko04010), which elicit many of the cell responses by environmental changes (Chen and Thorner, 2007) (Figure 1, Table S3 in Supporting Information). In addition, 25 DE transcripts were distributed in the viral carcinogenesis pathway (ko05203), which is related to cellular defense and occurrence of carcinogenesis (Chen et al., 2014) (Table S4 in Supporting Information).

Differential miRNA expression profiles

To detect the miRNA expression profile differences between diploid and triploid *C. auratus*, we evaluated the miRNA liver transcriptome data. After adaptor removal and quality control, clean reads (37.04 million) were mapped to miRBase (version 21.0) (Table S5 in Supporting Information). Among these, 1,868 known miRNAs were identified from the 6,500,318 clean reads in diploids, whereas 1,889 known miRNAs were detected from 10,314,077 clean reads in triploids (Table S6 in Supporting Information). miRNA comparison between diploids and triploids revealed 1,632 shared miRNAs (Figure S3A in Supporting Information). We predicted 428 and 410 novel miRNAs in diploids and triploids, respectively. After clean reads were mapped to miRBase and the reference genome, a total of 34.06 million (92.0%) reads were used to calculate the miRNA expression level of di-

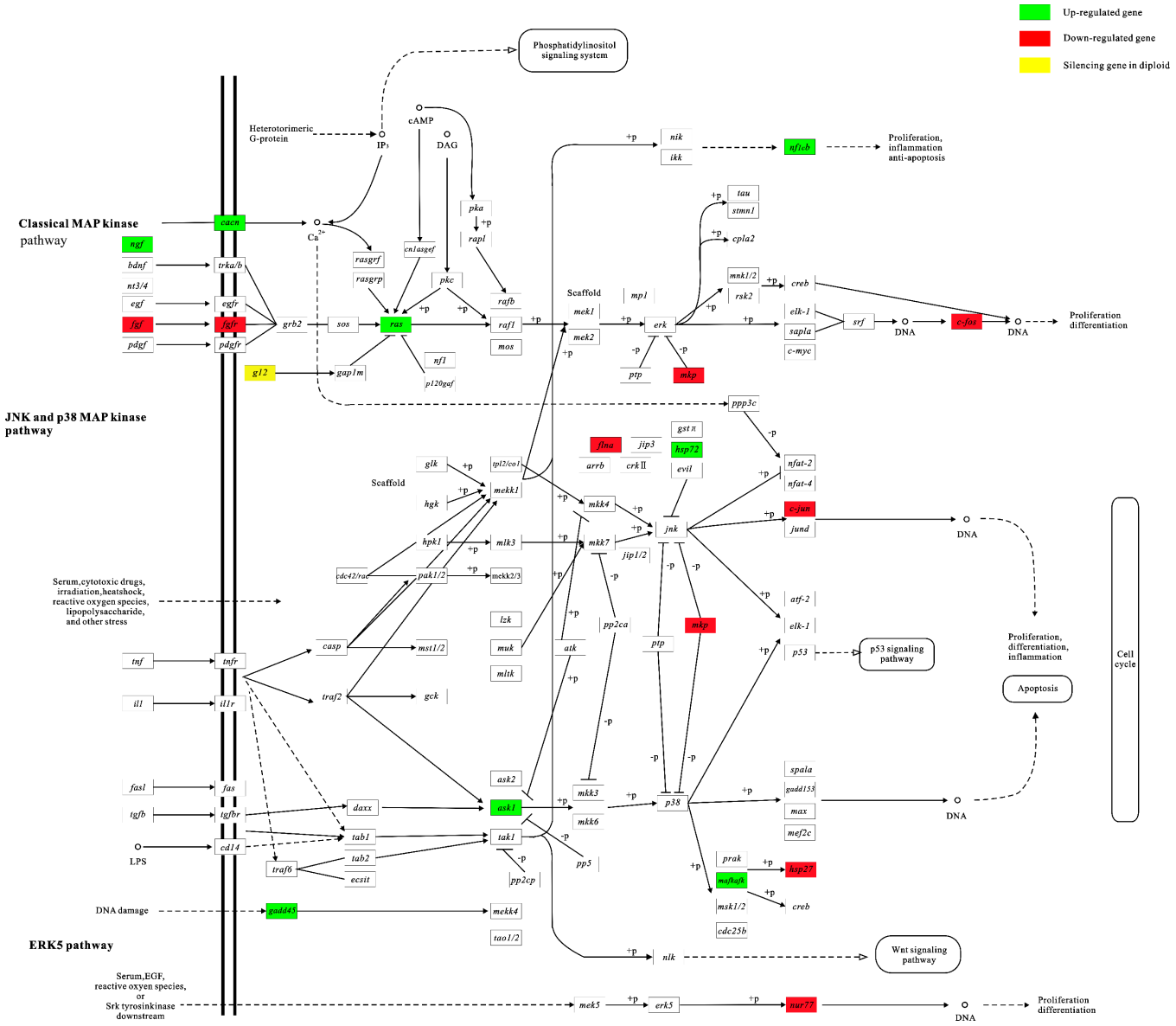

Figure 1 The DE genes in MAPK signaling pathway in comparison of diploid and triploid *C. auratus*. The red symbol represents down-regulated expression genes in triploids. The green symbol represents up-regulated expression genes in triploids. The yellow symbol describes the silencing genes in diploids.

ploids and triploids (Figure S3B in Supporting Information). Then, differential expression analysis showed that 52 and 347 up-regulated miRNAs in triploids and diploids, respectively.

Functional analysis of miRNA-mRNA co-expression

Based on DE miRNA analysis, 267 miRNAs regulated differential expression of 5,977 target transcripts, and 110 of the transcripts were regulated by two miRNAs and exhibited dual-direction regulation. Moreover, 7,615 DE pairs (miRNA-mRNA) were obtained by diploid and triploid comparison. miRNA target prediction showed that 52 up-regulated miRNAs were associated with 478 expression changes transcripts in triploids, whereas 347 down-regulated miRNAs regulated expression of 7,137 transcripts. Correlation analysis showed more positively correlated miRNA-mRNA pairs (122) than negatively correlated miRNA-mRNA pairs (86).

GO enrichment analysis showed that the target mRNAs regulated by DE miRNAs were mainly associated with the molecular functions of ATP binding (122 genes), metal ion binding (114 genes), and DNA binding (97 genes). Additionally, 33 DE transcripts were detected in the MAPK signaling pathway; among these, eight transcripts related to six genes (*cacna1d*, *map3k5*, *nr4a1*, *nfkb2*, *hspa1* and *fgfr4*) were affected by differential expression of seven miRNAs (Figure 2, Table 1). In the viral carcinogenesis pathway, four

of the 25 DE transcripts (*mhc1*, *irf7*, *nfkb2* and *pik3c*) were regulated by four miRNAs, including miR-466i-5p, miR-2898, novel miRNAs of scaffold114_3546, and scaffold721_8985 (Figure 2, Table 1). Additionally, the expression values of 4 DE transcripts (*cacna1d*, *nr4a*, *hspa1* and *fgfr4*) were validated by RT-qPCR based on the primers in Table S7 in Supporting Information. And the similar expression trends in three genes (*nr4a*, *hspa1* and *fgfr4*) were detected in our results (Figure 3). However, the opposite expression trend of *cacna1d* was detected (Figure 3). The sequencing preferences of illumina may result in no reads detected of *cacna1d* in diploids, which led to an error result similar to gene silencing. Then, we validated the expression levels of three miRNAs (scaffold1720_12188, miR-187-3p and miR-133c-3p) by RT-qPCR. Compared to the miRNA-seq result, there was a similar relative expression between diploids and triploids in miR-187-3p and miR-133c-3p, and a difference was found in expression of miRNA scaffold1720_12188 (obtained from *C. auratus* red var. genome) (Figure 3).

Genetic differences between diploids and triploids

Compared with the reference genome, we detected 1,128,104 and 1,102,052 mutations in diploids and triploids, respectively (Table S8 in Supporting Information). Among these, only 3,451 of 1,128,104 loci were heterozygous mutations (HMs) in diploids, whereas 3,544 of 1,102,052 loci were HMs in triploids. HMs were distributed in 2,547 and 2,594 transcripts in diploids and triploids, respectively (Figure 4A and C). The GO enrichment analysis exhibited that the differences in distribution (biological process level) were mainly associated to developmental process (GO: 0032502) and metabolic process (GO: 0008152) (Figure S4 in Supporting Information). Compared with diploids, we found novel HMs in triploids that were distributed in 1,783 transcripts (Table S9 in Supporting Information). Of these, the more than 2 HMs were detected in 27 genes including *pr1*, *mhc2*, and *man1* (Table S10 in Supporting Information). Additionally, 131,156 and 125,379 InDels were detected in diploids and triploids, respectively (Table S11 in Supporting Information), and the common InDels of diploids and triploids were distributed in 4,584 scaffolds (Figure 4B and C). In addition, a differential number of InDels (≥10) between diploids and triploids were detected in 426 transcripts (Table S12 in Supporting Information).

Positive selection detected in diploid and triploid *C. auratus*

After screening the 4,090 single-copy orthologs shared in three fish, we have obtained the *d*N/*d*S information from only 3,141 gen. Of these, 21 genes (ω>1), including *il21*, *fam213a*, and *itch* (ω>2), were under positive selection in diploid and triploid *C. auratus* (Figure 5, Table 2). These genes were related to biological processes, including immune response (GO: 0006955 in *il21*) and cell death (GO: 0008219 in *fam213a*) (Table 2). Compared with diploid *C. auratus* red var., we also found 17 and 26 positively selected genes in diploid and triploid *C. auratus*, respectively.

DISCUSSION

A recent study showed that changes in the population ratio of diploid and triploid *C. auratus* could be affected by environmental change (Xiao et al., 2011). Compared with the diploid population, the triploid population has a wider geographic distribution and more easily adapts to environmental change (Liu et al., 2017), including eutrophication that is caused by increased nitrogen and phosphorus, in the Dongting lake water system (Wang et al., 2014). This may be related to higher nucleotide diversity in triploids compared with diploids (Liu et al., 2017). However, only slight genetic structure differences between diploids and triploids were detected by gene flow analysis (Luo et al., 2014). These findings demonstrate molecular genetic mechanisms can play an important role in ecological adaptation. Therefore, we performed mRNA-seq, qRT-PCR and miRNA-seq to investigate underlying regulatory mechanisms.

Differential adaptation to harmful environmental changes is commonly observed in fish with different ploidy levels. A study on rainbow trout (*Oncorhynchus mykiss*) showed that triploids differed from diploids in their stress response and the changes induced by acute stress (Benfey and Biron, 2000). Under high-temperature conditions, distinct traits differ between diploids and triploids, including ovarian development, muscle growth and survival ratios were further described in both rainbow trout (Ojolick et al., 1995; Sumpter et al., 1991) and Atlantic salmon (*Salmo salar* L.) (Johnston et al., 1999). These differences could be affected by underlying mechanisms, including differential expression of mRNAs, miRNAs and proteins related to specific functions (Marie et al., 2006). Comparison of diploid and triploid *C. auratus*, revealed that 33 out of 1,324 significantly DE transcripts were associated with the MAPK signaling pathway, which is a key pathway of cellular proliferation and apoptosis regulation (Table S3 in Supporting Information). Moreover, we detected 25 DE transcripts in the viral carcinogenesis pathway (Table S4 in Supporting Information), which induced reprogramming and reduced genomic instability, including DNA damage accumulation that results from environmental agents (Chen et al., 2014). In addition, miRNAs are a class of small RNAs that regulate the mRNA expression, by influencing stability and translational efficiency of target mRNAs (Ambros, 2004). In this study, a

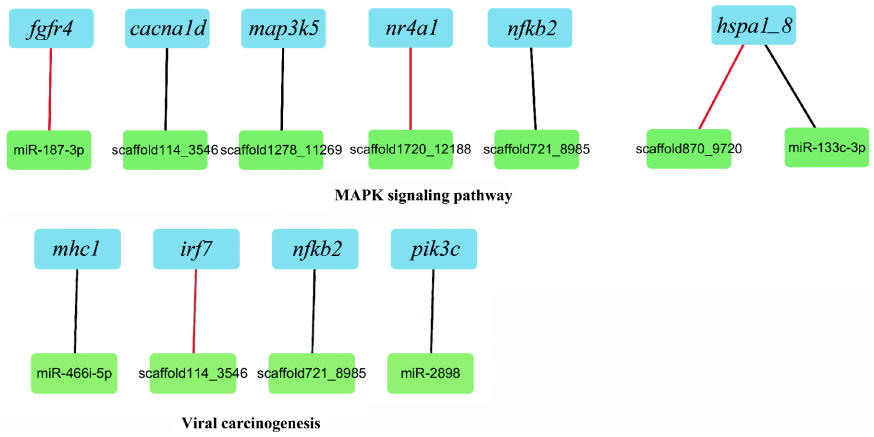

Figure 2 miRNA-mRNA interaction in MAPK signaling pathway and viral carcinogenesis pathway. A, Regulatory relationship in MAPK signaling pathway. B, Regulatory relationship in viral carcinogenesis pathway. Red line denotes positive correlation between expression of miRNA and mRNA, and black line indicates negative correlation between them. The expression level of mRNA (blue square) was regulated by corresponding miRNA (green square).

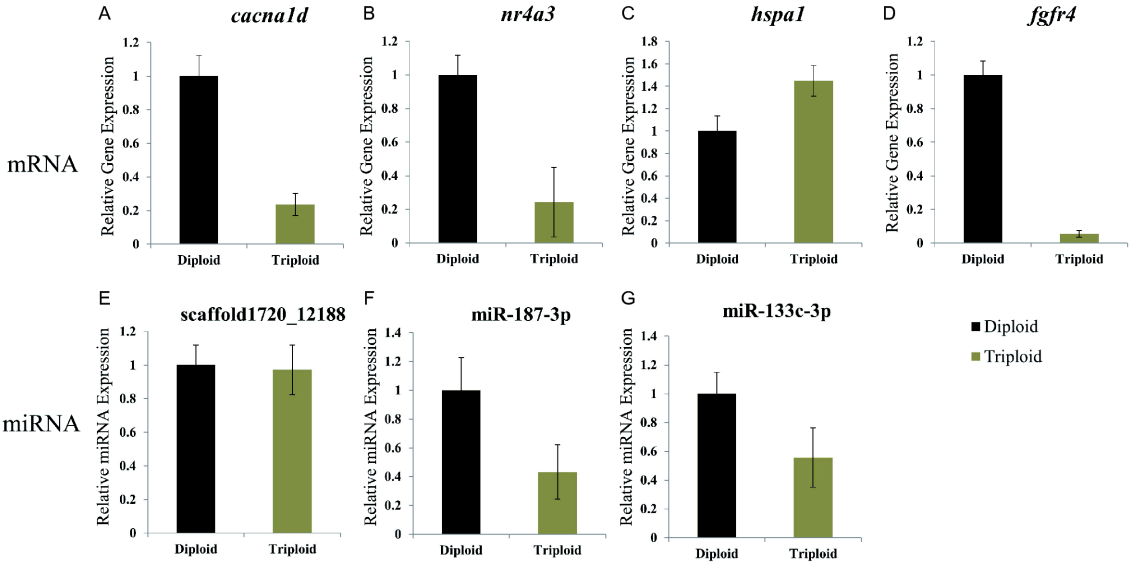

Figure 3 qRT-PCR analysis of four DE transcripts and three DE miRNAs. A, *cacna1d*, voltage-dependent calcium channel L type alpha-1D. B, *nr4a*, nuclear receptor subfamily 4 group A. C, *hspa1*, heat shock 70 kD protein 1. D, *fgfr4*, fibroblast growth factor receptor 4. E–G, the expression level of three miRNAs including scaffold1720_12188, miR-187-3p and miR-133c-3p. Comparative analysis reveals significant differences in gene expression (P<0.05) (n=3 for each group).

total of 12 DE transcripts were regulated by the expression of 11 miRNAs in the MAPK signaling and viral carcinogenesis pathways (Figure 2).

In triploid *C. auratus*, epigenetic regulation differences in parts of genes could facilitate adaptation to environmental changes. For example, expression changes of *nr4a* were regulated by miRNAs (scaffold1720_12188) (Table 1), and this change was related to environmental change in the available nutrition. The up-regulated of *nr4a* could improve food intake by regulating glucose and lipid metabolism (Oita et al., 2009). Our results revealed that *hspa1* and *hspa8* were regulated by two miRNAs (scaffold870_9720 and miR-133c-3p, respectively) (Figure 2), which could facilitate stabilization or degradation of mutant proteins and DNA repair (Duan et al., 2014; Mayer and Bukau, 2005). In comparison of diploids and triploids, the different expression trends were detected in two transcripts of the gene *hspa*. However, the transcript (CauG_02917, *hspa*) plays a leading role because of its higher expression value. Up-regulated of it in triploids could be beneficial for maintaining genetic stabilization. In addition, miR-187-3p, which is a miRNA that is less expressed in triploids, produces down regulation of the *fgfr4* transcript, and *fgfr4* is always overexpressed in gynecological tumor samples (Ye et al., 2011). Moreover, the *pik3c* is an immune gene associated with ovarian cancer regulation (Shayesteh et al., 1999), and its up-regulation

Table 1 The summary of differentially expressed miRNA in pathway of a MAPK signaling and viral carcinogenesis

Pathway name	miRNA ID	TPM* in 2*n*	TPM* in 3*n*	Fold change in miRNA	*P*-value	Differential expression in miRNA	miRNA sequences (5′→3′)	Length	Transcripts ID	Gene symbol	mRNA expression in 2*n*	mRNA expression in 3*n*	Fold change in mRNA	*P*-value	Differential expression in mRNA
MAPK signaling pathway	scaffold114_3546	1.35	0.28	0.2104	1.18×10^{-2}	Down	CTGCAGACTCG-GACTCTGGGTGC	23	CauG_02356	*cacna1d*	0.00	16.46	Inf	3.28×10^{-3}	Up
	scaf-fold1278_11269	3.01	0.38	0.1262	8.81×10^{-6}	Down	CAGGACGTGTG-TGGAGGAGTTC	22	CauG_14621	*map3k5*, *ask1*	16.50	160.02	9.6960	1.90×10^{-2}	Up
	scaf-fold1720_12188	13.08	1.99	0.1524	6.37×10^{-19}	Down	GTCAGGGGAGA-CAGAGGACTGT	22	CauG_24193	*nr4a1*, *hmr*	571.15	48.23	0.0844	2.06×10^{-2}	Down
	scaf-fold1720_12188	13.08	1.99	0.1524	6.37×10^{-19}	Down	GTCAGGGGAGA-CAGAGGACTGT	22	CauG_36905	*nr4a1*, *hmr*	4844.04	722.94	0.1492	4.40×10^{-2}	Down
	scaffold721_8985	5.86	0.95	0.1618	5.82×10^{-9}	Down	AGTCCCGCAG-CACACCCT-CAGTCCGTG	27	CauG_02493	*nfkb2*	181.57	519.58	2.8616	7.74×10^{-3}	Up
	scaffold870_9720	17.29	6.45	0.3732	4.33×10^{-11}	Down	CGCGGCTCCAG-AGACGCTCTTCT	23	CauG_08165	*hspa1_8*	10.52	0.64	0.0604	4.67×10^{-2}	Down
	miR-187-3p	0.75	0.01	0.0017	6.71×10^{-3}	Down	TCGTGTCTTGTG-TTGCAGCCGG	22	CauG_00857	*fgfr4*	248.05	45.66	0.1841	1.49×10^{-2}	Down
	miR-133c-3p	1.5	0.19	0.1262	1.90×10^{-3}	Down	TTGGTCCCCTTC-AACCAGCTGC	22	CauG_02917	*hspa1_8*	72.33	233.36	3.2263	5.11×10^{-3}	Up
	mmu-miR-466i-5p	0.6	0.01	0.0013	1.73×10^{-2}	Down	TGTGTGTGTGT-GTGTGTGTG	20	CauG_14683	*mhc1*	60.05	135.49	2.2564	4.99×10^{-2}	Up
Viral carcinogenesis	scaffold114_3546	1.35	0.28	0.21	1.18×10^{-2}	Down	CTGCAGACTCG-GACTCTGGGTGC	23	CauG_16862	*irf7*	133.42	38.98	0.2922	6.25×10^{-3}	Down
	scaffold721_8985	5.86	0.95	0.16	5.82×10^{-9}	Down	AGTCCCGCAG-CACACCCT-CAGTCCGTG	27	CauG_02493	*nfkb2*	181.57	519.58	2.8616	7.74×10^{-3}	Up
	bta-miR-2898	709.94	260.82	0.37	0.00	Down	TGGTGGA-GATGCCGGGGA	18	CauG_38594	*pik3c*	98.45	389.62	3.9577	4.66×10^{-3}	Up

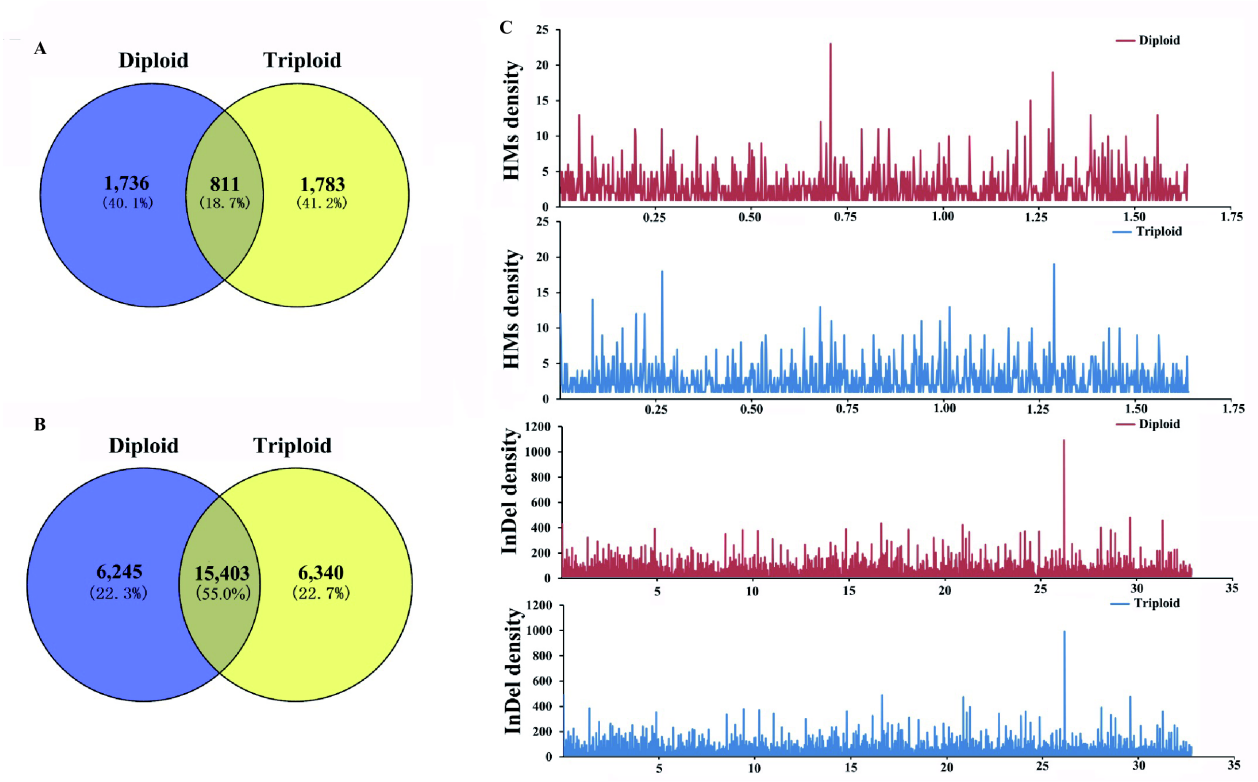

Figure 4 The distribution of heterozygotic mutations and InDels in diploid and triploid *C. auratus*. A, The number of HMs in diploid and triploid transcriptome. B, The number of InDels in diploid and triploid transcriptome. C, The distributed of HMs and InDels were exhibited with the diploids (red line) and triploids (blue line). The *x* axis represents the nucleotide position in Mb.

could promote cell growth (Samuels et al., 2005). Consequently, modified expression of these genes reflects the potential health of triploids as comparison with diploids, and could affect adaption to hazardous environments.

Rapid water change is likely to impose strong selective pressures on traits crucial for fitness, therefore, genetic changes in response to eutrophication are potential mechanisms that mitigate negative consequences of environmental changes. In the MAPK signaling pathway (ko04010), we found that three novel HMs were produced in *pr1* (CauG_09167) of triploids compared with diploids (Table S8 in Supporting Information). This finding reveals high genetic diversity in triploids that was consistent with the results of a previous population genetic study (Luo et al., 2014). A similar phenomenon was also described in other fishes, including triploid salmonid species and cyprinids (Allenjr and Stanley, 1983; Arai and Wilkins, 1987), and the higher levels of polymorphism could increase genetic vigor and promote further environmental adaptation. The higher neutral mutation rate in teleosts offers a higher chance for selection to produce and retain favorable mutations (Ravi and Venkatesh, 2008). Although the high d*N*/d*S* ratio may be related to the lack of information in full-length genes, the rapid evolution of 21 genes also indirectly reflected the advantage of triploids when adapting to a new environment (Table 3). Of these genes, the non-synonymous substitution in *il21* could help improve resistance ability and enhance adaptation to harmful environments.

Overall, epigenetic regulation and genetic changes were observed in the diploid and triploid populations that may have led to adaptation advantages in the triploids compared with sympatric diploids. However, because of the lack of transcriptome data for other members of the *C. auratus* species complex in different water systems, we could not explore the evolutionary history of polyploidy throughout the *C. auratus* species complex and confirm whether there are shared epigenetic regulation and genetic changes. Regardless, this analysis on the *C. auratus* complex in the Dongting Lake system helps elucidate the underlying molecular mechanisms that affect polyploid vertebrate ecology, genetics, and evolutionary biology.

MATERIALS AND METHODS

Sample preparation and RNA extraction

In our study, all experiments were approved by Animal Care Committee of Hunan Normal University and followed guidelines statement of the Administration of Affairs Concerning Animal Experimentation of China. In August 2015,

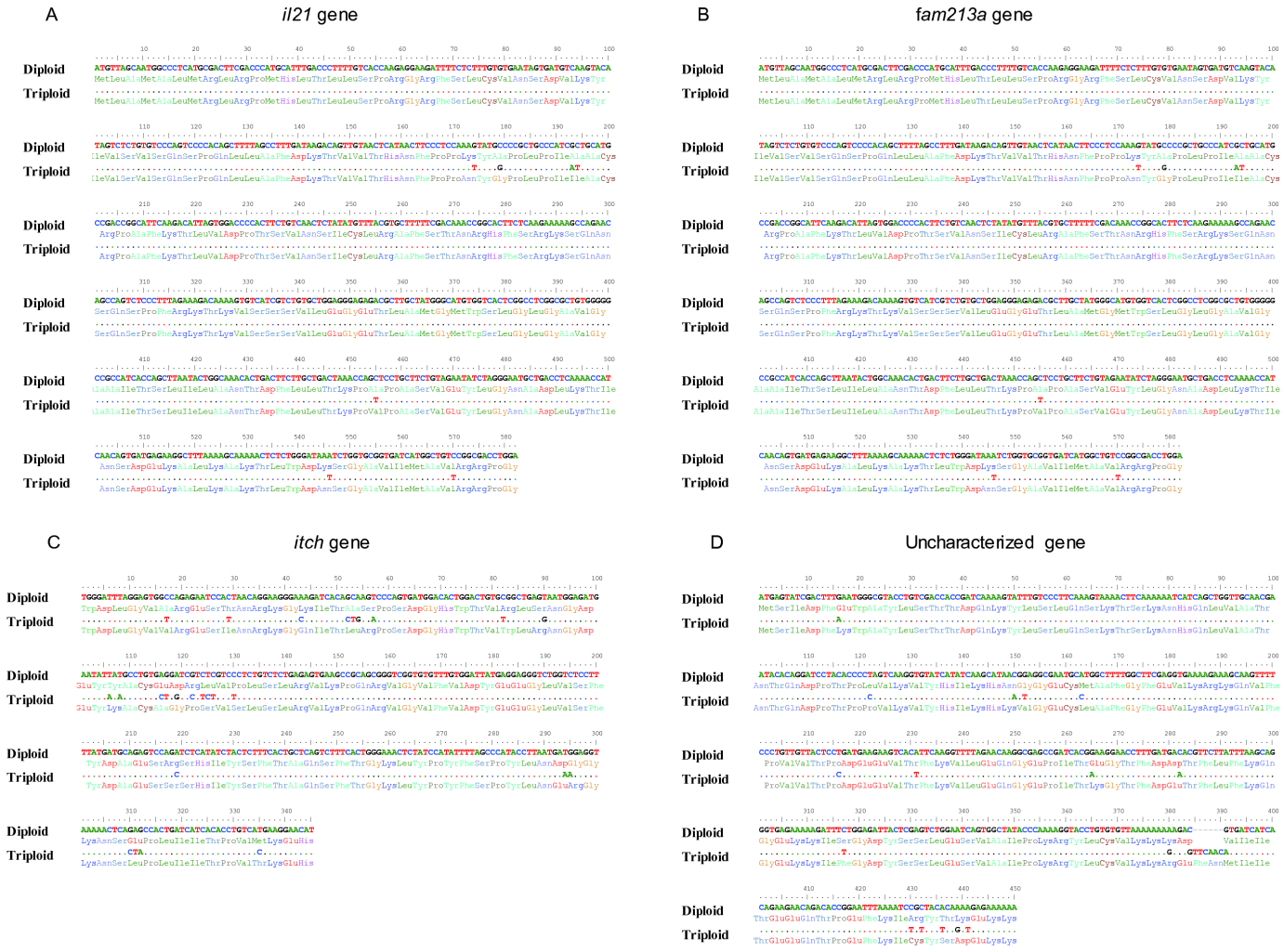

Figure 5 The amino sequences of four positive selective genes (ω>2) detected between diploid and triploid *C. auratus*. A, The sequence information of *il21* gene. B, The sequence information of *fam213a* gene. C, The sequence information of *itch* gene. D, The sequence information of unknown gene.

three individuals each of diploid and triploid *C. auratus* (200 g) were collected from the same region in Dongting lake water system, Hunan Province. The fish with similar morphological traits were obtained based on the description by Jun Xiao et al. (Xiao et al., 2011). The ploidy level of each sample was confirmed by a metaphase chromosome assay of cultured blood cells. Then, a flow cytometer was used to measure the DNA content of liver cells. The DNA content of diploid *C. auratus* red var. was used as a control. Then, the DNA of diploid and triploid *C. auratus* was obtained from blood samples based on the DNA extraction protocol (Asahida et al., 1996). The alignment of mitochondrial DNA sequences between obtained samples and GenBank database (diploid: GU086395, triploid: GU086396) was used to identification of the genetic background (Table S13 in Supporting Information).

After anesthetizing the fish with 2-phenoxyethanol, liver tissue was excised carefully to avoid gut contamination. Tissues were cut into small pieces and immediately pulled into RNALater (Ambion, AM7021, USA) at −80°C following the manufacturer's instructions. Total RNA was extracted from liver tissue according to a standard Trizol protocol (Invitrogen, USA) after RNALater removal. Total RNA was obtained with a DNA-free™ DNA Removal Kit (Ambion) to remove any contaminating genomic DNA. The purified RNA was quantified using a 2100 Bioanalyzer system (Agilent, USA).

mRNA sequencing

The six libraries (125-bp paired-end) were constructed following the Illumina Paired-End Sequencing Library protocol. Library quality and concentration were determined using an Agilent 2100 Bioanalyzer (Agilent). The Illumina HiSeq 2500 platform was used to detect the transcriptomes of three

biological replicates.

After removing the read adapters and low-quality reads, clean reads from each library were examined using FastQC (version 0.11.3). Then, mRNA-Seq reads from each sample were mapped against the reference genome (*C. auratus* red var., http://rd.biocloud.org.cn/) using Tophat with the parameter (Trapnell et al., 2012). We performed the analysis for removing the negative effects of background expression noise in all biological replicates. To analyze differential expression between diploid and triploid *C. auratus*, fragments per kilobase of transcript per million mapped reads (FPKM) (Mortazavi et al., 2008) values were calculated using Cufflinks (version 2.1.0) (Trapnell et al., 2012). Meanwhile, transcripts with false discovery rate (FDR)≤0.05 and log2 transformed fold change) ($\log_2$FC)>2 were defined as differentially expressed. Data were analyzed with the DEGseq package of R (version 2.13) (R Foundation for Statistical Computing, Austria) (Wang et al., 2010).

Small RNA sequencing

Approximately 2.5 μg of total RNA were obtained to prepare small RNA library according to the protocol of TruSeq Small RNA Sample Prep Kits (Illumina, USA). Then the libraries were sequenced by Illumina Hiseq 2500 (single-end) following the vendor's recommended protocol. Adaptor removal and quality control were performed by cutadapt (version 1.7.1), fastx_toolkit (version 0.0.13) and NGSQCToolkit (version 2.3.2). Subsequently, clean reads 15–26 bp in length were mapped to specific species precursors in miRBase 21.0 by Bowtie search to identify known miRNAs. The reads that did not align to above known miRNAs were used to perform novel miRNA discovery, which was described in the Creighton et al. (Creighton et al., 2010).

miRNA-mRNA integrated analysis

To process paired-end sequencing, reads were separately aligned to cover the mature miRNA on both forward and reverse reads, and the obtained number of matches was averaged. After normalizing match counts, miRNA expression profiling was performed. Using the edgeR package with default settings, $\log_2$FC>2 and P<0.05 were used as cutoffs to identify DE miRNA between diploids and triploids. miRNA target gene prediction was performed with the algorithm miRanda (Enright et al., 2003; John et al., 2004), and the parameters S≥150, $\Delta G \leq -30$ kcal mol^{-1} and strict 5′ seed pairing were used for above analysis. PCA and hierarchical clustering of the DE transcripts were performed and visualized with R (version 3.0.2), and transcript networks from the same gene ontology (GO) category were visualized with Cytoscape (version 3.6.0) (Kohl et al., 2011).

Determination of DE transcripts using qRT-RCR

Alignment of the transcripts sequences of diploid and triploid *C. auratus* were performed using Bioedit ver. 7.0.9. The qRT-PCR primers of four genes (*cacna1d*, *nr4a*, *hspa1* and *fgfr4*) were designed in conserved region of alignment sequences. Then, these primers had been used to detect expression using the ABI Prism 7500 Sequence Detection System (Applied Biosystems, USA). Amplification conditions were as follows: 50°C for 5 min, 95°C for 10 min, and 36 cycles at 95°C for 15 s and 60°C for 45 s. Each test was performed three times to improve the accuracy of the results. Finally, relative quantification was performed and melting curve analysis was used to verify the generation of a single product at the end of the assay. Triplicates of each sample were used both for standard curve generation and during experimental assays. The relative expression of each gene was calibrated with *β-actin*, and the relative mRNA expression data were analysis using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001). The expression level of the reference gene *β-actin* in triploid *C. auratus* was estimated using the ratio of the transcript abundance to the gene copy using PCR and qRT-PCR of co-extracted DNA and RNA from the liver of diploid and triploid individuals. A similar expression value of *β-actin* was detected between diploids and triploids.

To validate the expression level of three miRNAs (scaffold1720_12188, miR-187-3p and miR-133c-3p) obtained from the miRNA-seq result, the relative expressions of miRNAs were evaluated by quantifying the miRNA stem-loop RT-qPCR. The cDNA was obtained from total RNA using a reverse transcription kit (Invitrogen, USA). The qPCR reaction was performed using the same method using the above conditions with the miRNA-specific primers (Table S7 in Supporting Information), which were designed in the laboratory and synthesized by Generay Biotech (Generay, PRC) based on the miRNA sequences (Table 1). The relative expression of each miRNAs was calibrated with miR-22a (primer: 5′-AAGCTGCCAGCTGAAGAACTGT-3′), and the $2^{-\Delta\Delta C_t}$ method was used in next miRNA expression analysis.

Heterozygous mutation (HM) analysis, and insertion and deletion (InDel) variation

To detect HM and InDel differences in transcripts between diploid and triploid *C. auratus*, we performed mapping analysis using BWA (version 0.7.5a) (Li and Durbin, 2009). The clean reads of six samples were mapped to the *C. auratus* red var. reference genome. The mapping data were indexed, and HM and InDel information was obtained using samtool (version 1.3.1) (Li et al., 2009) and bedtools (version 1.3.1) (Quinlan and Hall, 2010). Then, the snpEff (Cingolani et al., 2012) was used to calculate HM and InDel distribu-

Table 2 The summary of positively selected genes between diploid and triploid crucian carp

Gene name	d*N*	d*S*	d*N*/d*S*	*P*-value (fisher)	Length (bp)	*S*-sites	*N*-sites	Substitutions	*S*-substitutions	*N*-substitutions
Interleukin	0.02	0.01	2.12	0.61	426	98.46	327.54	8	1	7
MACRO domain-containing 1	0.01	0.00	1.74	0.54	1,071	323.77	747.23	5	1	4
Ufm1-specific protease 2	0.00	0.00	1.23	0.84	1,398	407.29	990.71	4	1	3
Hypothetical protein cypCar_00005244	0.00	0.00	1.06	0.81	1,074	280.18	793.82	4	1	3
SUMO/sentrin peptidase family member, NEDD8-specific	0.02	0.01	1.12	0.95	486	150.40	335.60	7	2	5
E3 ubiquitin-protein ligase TRIM39-like	0.10	0.03	2.83	0.07	345	95.11	249.89	26	3	23
Uncharacterized protein LOC109101439 (*Cyprinus carpio*)	0.05	0.02	2.47	0.20	447	145.72	301.28	18	3	15
Redox-regulatory protein FAM213A-like	0.02	0.01	2.73	0.40	582	181.05	400.95	7	1	6
Fas (tnfrsf6)-associated via death domain	0.01	0.01	1.02	0.80	582	147.66	434.34	4	1	3
Smad nuclear interacting protein	0.00	0.00	1.03	0.80	1,113	283.54	829.46	4	1	3
E2F transcription factor 4	0.00	0.00	1.17	0.83	1,203	338.03	864.97	4	1	3
Cyclic GMP-AMP synthase-like isoform X2	0.00	0.00	1.26	0.84	1,254	371.27	882.73	4	1	3
E3 ubiquitin-protein ligase RNF182-like	0.01	0.01	1.17	0.83	471	106.79	364.21	5	1	4
Hypothetical protein cypCar_00046847 (*Cyprinus carpio*)	0.02	0.02	1.20	0.96	471	120.27	350.74	9	2	7
CD3e molecule, epsilon-associated protein	0.01	0.01	1.14	0.95	843	263.69	579.31	7	2	5
UV-stimulated scaffold protein A	0.00	0.00	1.12	0.95	2,175	674.02	1,500.98	7	2	5
G patch domain-containing 8	0.03	0.03	1.13	1.00	804	173.46	630.54	23	5	18
GDNF-inducible zinc finger protein 1 isoform X1	0.01	0.01	1.02	0.80	513	129.65	383.35	4	1	3
Immunoglobulin superfamily member 5-like isoform X1	0.01	0.01	1.11	0.95	1,062	285.89	776.11	8	2	6
PHD finger protein 13-like isoform X3	0.01	0.01	1.01	0.94	924	264.91	659.09	7	2	5
Myb-like protein X isoform X3	0.05	0.03	1.82	0.11	1,191	287.413	903.587	55	8	47

tions based on threshold of quality value>20.

Nonsynonymous/synonymous substitution rate identification

To detect whether positive selection occurred among diploid *C. auratus* red var., and diploid and triploid *C. auratus*, *de novo* assembly of each them was performed using Trinity (Grabherr et al., 2011). Then, to obtain putative single-copy orthologs, reciprocal BLASTX search (*e*-value=1×10^{-20}) and OrthoMCL analysis were performed on the three fishes (Blanc and Wolfe, 2004; Li et al., 2003). Of these, two sequences were defined as orthologs if each of them was the best hit and the sequences were aligned over 300 bp. Then, ortholog alignment was performed using ClustalW (version 2.1) (Li, 2003). After deletion of unmatched regions of aligned orthologs, nonsynonymous and synonymous substitution rates (ω=d*N*/d*S*) were estimated using the Yang-Neilsen method in KaKs_Calculator (version 2.0) (Zhang et al., 2006).

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

Acknowledgements *This work was supported by the National Natural Science Foundation of China (317023343, 1430088, 31730098), the earmarked fund for China Agriculture Research System (CARS-45), the Key Research and Development Project of Hunan Province (2016NK2128), the Key Research and Development Project of Hunan Province (2016NK2128), Hunan Provincial Natural Science and Technology Major Project (2017NK1031), 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, Natural Science Foundation of Hunan Province (14JJ2148) and the Scientific Research Fund of Hunan Provincial Education Department (16C0974).*

Allenjr, S.K., and Stanley, J.G. (1983). Ploidy of hybrid grass carp×bighead carp determined by flow cytometry. Trans Am Fisheries Soc 112, 431–435.

Ambros, V. (2004). The functions of animal microRNAs. Nature 431, 350–355.

Arai, K., and Wilkins, N.P. (1987). Triploidization of brown trout (*Salmo trutta*) by heat shocks. Aquaculture 64, 97–103.

Arkhipchuk, V.V., and Garanko, N.N. (2005). Using the nucleolar biomarker and the micronucleus test on *in vivo* fish fin cells. Ecotoxicol Environ Saf 62, 42–52.

Asahida, T., Kobayashi, T., Saitoh, K., and Nakayama, I. (1996). Tissue preservation and total DNA extraction form fish stored at ambient temperature using buffers containing high concentration of urea. Fisheries Sci 62, 727–730.

Benfey, T.J., and Biron, M. (2000). Acute stress response in triploid rainbow trout (*Oncorhynchus mykiss*) and brook trout (*Salvelinus fontinalis*). Aquaculture 184, 167–176.

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.

Bonar, S.A., Thomas, G.L., and Pauley, G.B. (2010). Evaluation of the separation of triploid and diploid grass carp, *Ctenopharyngodon idella* (*Valenciennes*), by external morphology. J Fish Biol 33, 895–898.

Chen, R.E., and Thorner, J. (2007). Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta 1773, 1311–1340.

Chen, S., Wang, J., Liu, S.J., Qin, Q.B., Xiao, J., Duan, W., Luo, K.K., Liu, J.H., and Liu, Y. (2009). Biological characteristics of an improved triploid crucian carp. Sci China Ser C 52, 733–738.

Chen, Y., Williams, V., Filippova, M., Filippov, V., and Duerksen-Hughes, P. (2014). Viral carcinogenesis: factors inducing DNA damage and virus integration. Cancers 6, 2155–2186.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly 6, 80–92.

Creighton, C.J., Benham, A.L., Zhu, H., Khan, M.F., Reid, J.G., Nagaraja, A.K., Fountain, M.D., Dziadek, O., Han, D., Ma, L., et al. (2010). Discovery of novel microRNAs in female reproductive tract using next generation sequencing. PLoS ONE 5, e9637.

Duan, Y., Huang, S., Yang, J., Niu, P., Gong, Z., Liu, X., Xin, L., Currie, R. W., and Wu, T. (2014). HspA1A facilitates DNA repair in human bronchial epithelial cells exposed to Benzo[a]pyrene and interacts with casein kinase 2. Cell Stress Chaperones 19, 271–279.

Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D.S. (2003). microRNA targets in *Drosophila*. Genome Biol 5, R1.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29, 644–652.

Hochholdinger, F., and Hoecker, N. (2007). Towards the molecular basis of heterosis. Trends Plant Sci 12, 427–432.

Jin, J., Wang, Y., Wu, Z., Hergazy, A., Lan, J., Zhao, L., Liu, X., Chen, N., and Lin, L. (2017). Transcriptomic analysis of liver from grass carp (*Ctenopharyngodon idellus*) exposed to high environmental ammonia reveals the activation of antioxidant and apoptosis pathways. Fish Shellfish Immunol 63, 444–451.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human microRNA targets. PLoS Biol 2, e363.

Johnston, I.A., Strugnell, G., Mccracken, M.L., and Johnstone, R. (1999). Muscle growth and development in normal-sex-ratio and all-female diploid and triploid *Atlantic salmon*. J Exp Biol 202, 1991–2016.

Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. Methods Mol Biol 696, 291–303.

Lee, S.I., Nguyen, X.T., Kim, J.H., and Kim, N.S. (2016). Genetic diversity and structure analyses on the natural populations of diploids and triploids of tiger lily, *Lilium lancifolium* Thunb., from Korea, China, and Japan. Genes Genom 38, 467–477.

Li, H., and Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. (Oxford: Oxford University Press).

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, K.B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics 19, 1585–1586.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13, 2178–2189.

Liu, X.L., Jiang, F.F., Wang, Z.W., Li, X.Y., Li, Z., Zhang, X.J., Chen, F., Mao, J.F., Zhou, L., and Gui, J.F. (2017). Wider geographic distribution and higher diversity of hexaploids than tetraploids in *Carassius* species complex reveal recurrent polyploidy effects on adaptive evolution. Sci Rep 7, 5395.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods 25, 402–408.

Lu, P., Su, Y., Niu, Z., and Wu, J. (2007). Geostatistical analysis and risk assessment on soil total nitrogen and total soil phosphorus in the

Dongting Lake plain area, China. J Environ Qual 36, 935.
Luo, J., Gao, Y., Ma, W., Bi, X., Wang, S., Wang, J., Wang, Y., Chai, J., Du, R., Wu, S., et al. (2014). Tempo and mode of recurrent polyploidization in the *Carassius auratus* species complex (*Cypriniformes*, *Cyprinidae*). Heredity 112, 415–427.
Marie, V., Gonzalez, P., Baudrimont, M., Boutet, I., Moraga, D., Bourdineaud, J.P., and Boudou, A. (2006). Metallothionein gene expression and protein levels in triploid and diploid oysters *Crassostrea gigas* after exposure to cadmium and zinc. Environ Toxicol Chem 25, 412.
Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci 62, 670–684.
Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5, 621–628.
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.
Oita, R.C., Mazzatti, D.J., Lim, F.L., Powell, J.R., and Merry, B.J. (2009). Whole-genome microarray analysis identifies up-regulation of Nr4a nuclear receptors in muscle and liver from diet-restricted rats. Mech Ageing Dev 130, 240–247.
Ojolick, E.J., Cusack, R., Benfey, T.J., and Kerr, S.R. (1995). Survival and growth of all-female diploid and triploid rainbow trout (*Oncorhynchus mykiss*) reared at chronic high temperature. Aquaculture 131, 177–187.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.
Ravi, V., and Venkatesh, B. (2008). Rapidly evolving fish genomes and teleost diversity. Curr Opin Genet Dev 18, 544–550.
Samuels, Y., Diaz Jr., L.A., Schmidt-Kittler, O., Cummins, J.M., DeLong, L., Cheong, I., Rago, C., Huso, D.L., Lengauer, C., Kinzler, K.W., et al. (2005). Mutant *PIK3CA* promotes cell growth and invasion of human cancer cells. Cancer Cell 7, 561–573.
Shayesteh, L., Lu, Y., Kuo, W.L., Baldocchi, R., Godfrey, T., Collins, C., Pinkel, D., Powell, B., Mills, G.B., and Gray, J.W. (1999). PIK3CA is implicated as an oncogene in ovarian cancer. Nat Genet 21, 99–102.
Stanley, J.G., Hidu, H., and Allen Jr., S.K. (1984). Growth of American oysters increased by polyploidy induced by blocking meiosis I but not meiosis II. Aquaculture 37, 147–155.
Sumpter, J.P., Lincoln, R.F., Bye, V.J., Carrgher, J.F., and Le Bail, P.Y. (1991). Plasma growth hormone levels during sexual maturation in diploid and triploid rainbow trout (*Oncorhynchus mykiss*). General Comp Endocrinol 83, 103–110.
Swarup, H. (1956). Production of heteroploidy in the three-spined stickleback, *Gasterosteus aculeatus* (L.). Nature 178, 1124–1125.
Swarup, H. (1959). Production of triploidy in *Gasterosteus aculeatus* (L). J Genet 56, 129–142.
Tiwary, B.K., Kirubagaran, R., and Ray, A.K. (2004). The biology of triploid fish. Rev Fish Biol Fisheries 14, 391–402.
Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562–578.
Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26, 136–138.
Wang, Y., Jiang, X., Li, Y.F., Wang, S.H., Wang, W.W., and Cheng, G.L. (2014). Spatial and temporal distribution of nitrogen and phosphorus and nutritional characteristics of water in Dongting Lake. Res Environ Sci 27, 484–491.
White, M.J.D. (1970). Heterozygosity and Genetic Polymorphism in Parthenogenetic Animals. (New York: Springer).
Wolters, W.R., Chrisman, C.L., and Libey, G.S. (1982). Erythrocyte nuclear measurements of diploid and triploid channel catfish, *Ictalurus punctatus* (*Rafinesque*). J Fish Biol 20, 253–258.
Xiao, J., Zou, T., Chen, Y., Chen, L., Liu, S., Tao, M., Zhang, C., Zhao, R., Zhou, Y., Long, Y., et al. (2011). Coexistence of diploid, triploid and tetraploid crucian carp (*Carassius auratus*) in natural waters. BMC Genet 12, 20.
Yao, H., Dogra Gray, A., Auger, D.L., and Birchler, J.A. (2013). Genomic dosage effects on heterosis in triploid maize. Proc Natl Acad Sci USA 110, 2665–2669.
Ye, Y.W., Zhou, Y., Yuan, L., Wang, C.M., Du, C.Y., Zhou, X.Y., Zheng, B.Q., Cao, X., Sun, M.H., Fu, H., et al. (2011). Fibroblast growth factor receptor 4 regulates proliferation and antiapoptosis during gastric cancer progression. Cancer 117, 5304–5313.
Yoshida, M., Nagamine, M., and Uematsu, K. (2005). Comparison of behavioral responses to a novel environment between three teleosts, bluegill *Lepomis macrochirus*, crucian carp *Carassius langsdorfii*, and goldfish *Carassius auratus*. Fisheries Sci 71, 314–319.
Zhang, Z., Li, J., Zhao, X Q., Wang, J., Wong, K.S., and Yu, J. (2006). KaKs_Calculator: calculating *K*a and *K*s through model selection and model averaging. GPB 4, 259–263.
Zhu, X., Zhu, L., Lang, Y., and Chen, Y. (2008). Oxidative stress and growth inhibition in the freshwater fish *Carassius auratus* induced by chronic exposure to sublethal fullerene aggregates. Environ Toxicol Chem 27, 1979–1985.

SUPPORTING INFORMATION

Figure S1 The distribution of DE mRNAs between diploid and triploid *C. auratus*.

Figure S2 The function distribution of DE mRNAs between diploid and triploid *C. auratus*.

Figure S3 The distribution of DE miRNAs between diploid and triploid *C. auratus*.

Figure S4 Gene ontology (GO) (level 2) assignments for transcripts related to HMs in diploid and triploid *C. auratus*.

Table S1 The basic information of transcriptome data

Table S2 The summary of mapping to reference genome

Table S3 The information of differential expression genes in MAPK signaling pathway

Table S4 The information of differential expression genes in viral carcinogenesis pathway

Table S5 The summary of small RNA data

Table S6 The annotated information of small RNA in diploid and triploid samples

Table S7 RT-qPCR primer of five DE transcripts (*cacna1d*, *nr4a*, *hspa1*, *fgfr4* and *pik3c*) and three DE miRNAs

Table S8 The distribution of mutation sites in diploid and triploid *C. auratus*

Table S9 The gain of heterozygous loci distributed in 1,783 transcripts in triploid *C. auratus*

Table S10 The mergeing of heterozygous loci (⩾3) in triploid as compared with diploid *C. auratus*

Table S11 The summary of INDEL information in diploid and triploid *C. auratus*

Table S12 The different number of InDel (⩾10) between diploid and triploid *C. auratus*

Table S13 The summary of blast results as compared with the diploid and triploid *C. auratus* mitochondrial DNA sequences

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.