Genetic responses in sexual diploid and asexual triploid goldfish (Carassius auratus) introduced into a high-altitude environment

Xiu Feng[1], Shenglin Liu[2], Xiaoyun Sui[1], Yifeng Chen[1], Ren Zhu[1], Yintao Jia[1], Jingou Tong[1], Xiaomu Yu[1], Chunlong Liu[1], and Michael Hansen[2]

[1]Institute of Hydrobiology Chinese Academy of Sciences
[2]Aarhus University

November 16, 2022

Abstract

Anthropogenic biological invasions represent major concerns but enable us to investigate rapid evolutionary changes and adaptation to novel environments. The goldfish Carassius auratus with sexual diploids and asexual triploids coexisting in natural waters, is one of the most widespread invasive fishes in Tibet, providing an ideal model to study evolutionary processes during invasion in different reproductive forms from the same vertebrate. Here, using whole-genome resequencing data of 151 C. auratus individuals from invasive and native ranges, we found different patterns of genomic responses between diploid and triploid populations during their invasion to Tibet. For diploids, although invasive individuals derived from two different genetically distinct sources and had a relative higher diversity (π) at the population level, their individual genetic diversity (genome-wide observed heterozygosity) was significantly lower (21.4%) than that of source individuals. Population structure analysis revealed that the invasive individuals formed a specific genetic cluster distinct from the source populations. Runs of homozygosity analysis showed low inbreeding only in invasive individuals, and only the invasive population experienced a recent decline in effective population size reflecting founder events. For triploids, however, invasive populations showed no loss of individual genetic diversity and no genetic differentiation relative to source populations. Regions of putative selective sweeps between invasive and source populations of diploids mainly involved genes associated with mannosidase activity and embryo development. Our results suggest invasive diploids deriving from distinct sources still lost individual genetic diversity resulting from recent inbreeding and founder events and selective sweeps, and invasive triploids experienced no genetic change owing to their reproduction mode of gynogenesis that precludes inbreeding and founder effects and may make them more powerful invaders.

Genetic responses in sexual diploid and asexual triploid goldfish (*Carassius auratus*) introduced into a high-altitude environment

Running title: Genetic responses in invasive goldfish

Xiu Feng[1], Shenglin Liu[2], Xiaoyun Sui[1*], Yifeng Chen[1*], Ren Zhu[1], Yintao Jia[1], Jingou Tong[1], Xiaomu Yu[1], Chunlong Liu[1], Michael M. Hansen[2]

[1] Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

[2] Department of Biology, Aarhus University, Aarhus C, Denmark

* Authors for correspondence:

Xiaoyun Sui: xiaoyunsui@ihb.ac.cn

Yifeng Chen: chenyf@ihb.ac.cn

Abstract

Anthropogenic biological invasions represent major concerns but enable us to investigate rapid evolutionary changes and adaptation to novel environments. The goldfish *Carassius auratus* with sexual diploids and asexual triploids coexisting in natural waters, is one of the most widespread invasive fishes in Tibet, providing an ideal model to study evolutionary processes during invasion in different reproductive forms from the same vertebrate. Here, using whole-genome resequencing data of 151 *C. auratus*individuals from invasive and native ranges, we found different patterns of genomic responses between diploid and triploid populations during their invasion to Tibet. For diploids, although invasive individuals derived from two different genetically distinct sources and had a relative higher diversity (π) at the population level, their individual genetic diversity (genome-wide observed heterozygosity) was significantly lower (21.4%) than that of source individuals. Population structure analysis revealed that the invasive individuals formed a specific genetic cluster distinct from the source populations. Runs of homozygosity analysis showed low inbreeding only in invasive individuals, and only the invasive population experienced a recent decline in effective population size reflecting founder events. For triploids, however, invasive populations showed no loss of individual genetic diversity and no genetic differentiation relative to source populations. Regions of putative selective sweeps between invasive and source populations of diploids mainly involved genes associated with mannosidase activity and embryo development. Our results suggest invasive diploids deriving from distinct sources still lost individual genetic diversity resulting from recent inbreeding and founder events and selective sweeps, and invasive triploids experienced no genetic change owing to their reproduction mode of gynogenesis that precludes inbreeding and founder effects and may make them more powerful invaders.

Keywords : Genetic response, biological invasion,*Carassius auratus* , ploidy, Tibet, selective sweep

1 Introduction

Biological invasion denotes the phenomenon that an alien species spreads outside its natural range and proliferates and persists in new habitats. This is increasingly occurring as a result of human activities and causes negative impacts on native ecosystems (Mack et al., 2000; Reid et al., 2005). Nevertheless, invasive populations are also interesting models to investigate rapid genetic response and adaptation to novel environments, thus providing valuable insights into basic biological processes (Prentis et al., 2008; Sakai et al., 2001). In general, an introduced population tends to lose its genetic diversity because of bottlenecks which can reduce the fitness and evolutionary potential (Lee, 2002). The genetic paradox of invasion appears when a bottlenecked introduced population becomes invasive (Estoup et al., 2016). However, recent studies provide evidence that the genetic paradox of invasion is to a large extent overrated. In some cases, no paradox exists, as introduced populations shows no loss of genetic diversity possibly owing to multiple introductions, or no adaptive challenge in the introduced habitat (Blumenfeld et al., 2021; Facon et al., 2006; Rius et al., 2015). In other cases, an apparent paradox is spurious, when the loss in genetic diversity of introduced populations mainly occurs at neutral genetic markers which are not reflected in adaptive traits, or the diversity loss is a consequence of a successful response to strong selection (Dlugosch & Parker, 2008; Estoup et al., 2016; Gonzalez et al., 2013). Genetic analysis of the genetic architecture and composition of introduced populations would provide insights into the evolutionary mechanisms facilitating the invasion success.

High-throughput sequencing technologies, such as whole genome resequencing and reduced representation sequencing, provide huge number of molecular markers for genome-wide population genetic studies in non-model organisms (Ellegren, 2014). Genome-wide markers, such as single nucleotide polymorphisms (SNPs), can be used for accurately estimating genome-wide genetic diversity both within individuals and at population levels and provide powerful tools to uncover population genetic structure and reconstruct invasion history (Austin et al., 2006; Baltazar-Soares et al., 2020; Le Moan et al., 2021; Liu et al., 2018). Moreover, genome scans allow detecting possible footprints of selection associated with local adaptation, albeit not without challenges (Haasl & Payseur, 2016). Finally, significant progress has been made in using genomic data for reconstructing long-term demographic history and estimating genomic inbreeding level and recent demographic history which affect genetic diversity (Beichman et al., 2018; Ceballos et al., 2018; Dong et al., 2021).

The Tibetan Plateau is one of the largest and highest plateaus on Earth. The fish diversity of the Tibetan Plateau is very sensitive and vulnerable to biological invasion due to the fragile ecosystem and unique fish fauna (Favre et al., 2015; He et al., 2020; Jia et al., 2019; Tao et al., 2018). One of the most widespread invasive species in the region is the goldfish *Carassius auratus* which is originally distributed throughout the East Asian region except for the Tibetan Plateau (Luo & Yue, 2000). It represents a remarkable species complex containing individuals with different ploidy levels in natural waters (mainly diploid and triploid, and in rare cases tetraploid) (Liu et al., 2017b; Xiao et al., 2011). They have different reproduction modes, with diploids exhibiting sexual reproduction and triploids exhibiting unisexual gynogenesis (Gui & Zhou, 2010; Zhou et al., 2000). Triploids produce chromosome number-unreduced eggs by suppression of the first meiotic division. Eggs are subsequently activated by the sperm of sympatric sexual species to initiate embryogenesis, resulting in clonal offspring from their mothers (Wang et al., 2022). Such parthenogenesis in vertebrates is also observed in a few other fishes, such as*Poecilia formosa* (Warren et al., 2018) and *Phoxinus eos-neogaeus* (Angers & Schlosser, 2007). Unisexual vertebrates are generally predicted to have low evolutionary potentials due to lack of meiotic recombination which thus result in the accumulation of deleterious mutations and hindrance of the creation of genetic diversity (Butlin, 2002). On the other hand, asexual organisms may have advantages to be better colonizers than sexual organisms, because they can populate new habitats without mates thus avoiding inbreeding (Avise, 2008). The unisexual triploid *C. auratus* derives from sympatric diploids by multiple independent polyploidization events and possesses a comparable or slightly higher genetic diversity compared with diploids (Liu et al., 2017c; Luo et al., 2014; Ren et al., 2018). If both sexual diploids and unisexual triploids have been introduced into the Tibetan Plateau, this raises an interesting possibility to compare the genetic responses of the two forms.

C. auratus is one of the most popular fishes in Tibet markets and are bought and released into local waters for religious reasons (Jia et al., 2019), resulting in repeated introductions of the species into the waters of Tibet. Based on our preliminary market survey (data not published), *C. auratus* in Tibet is directly imported from the aquatic product markets in Ningxia (NX: located on the upper Yellow River) and Sichuan (SC: located on the upper Yangtze River) Provinces mainly by the two main transportation routes (Qinghai-Tibet Highway and Sichuan-Tibet Highway) connecting Tibet and its east (Figure 1). In addition, *C. auratus* in Ningxia was also imported from the central and eastern regions along the Yangtze River and Huai River. Hence, invasive populations of *C. auratus* in Tibet came from diverse sources which may be genetically distinct. Previous studies of multiple invasive species have shown that multiple introductions from the same or distinct source populations may have prevented or even reversed loss of genetic diversity within the invasive range owing to founder events (Blumenfeld et al., 2021; Kelager et al., 2013; van Boheemen et al., 2017). However, given the extreme environmental conditions on the Tibetan Plateau (Feng et al., 2019), this begs the question to which extent genetic diversity of *C. auratus* has been shaped by the extreme environments that it has been transplanted into during multiple introductions.

In this study, by analyzing whole-genome sequence data of 151 goldfish individuals from two invasive and 11 native populations, we investigated how genomic changes occurred in invasive populations in Tibet derived from diverse sources. We specifically tested the hypothesis that different genetic imprints of the colonization process would be detected between invasive diploid and triploid populations, with loss of variation and inbreeding occurring in diploids, but with no loss of variation in triploids. To address this, we first determined the ploidy levels of each individual and called SNPs for diploids and triploids separately. Subsequently, genetic diversity and population genetic structure were analyzed in both invasive and native populations, with a specific focus on detecting inbreeding and footprints of recent founder events in invasive populations. Finally, putative signatures of selection in invasive populations were analyzed, with the expectation that such patterns would be evident only in diploids. Our results provide insights into evolutionary processes facilitating the invasion success in Tibet and also insights into different genetic responses to novel habitats between sexual and asexual forms of vertebrate species.

2 Materials and Methods

2.1 Sample collection and whole-genome sequencing

In the invasive range, a total of 62 *C. auratus* juveniles with an average weight of 2.9 g were collected in 2019 from the Lala (LL) and Chabalang (CBL) wetlands within the Yarlung Zangbo River basin in Tibet (Table 1). In the native range, 73 *C. auratus* adults were collected in 2015-2021 from 11 sites belonging to five river systems, including the Yangtze River (SC, Sichuan; SS, Shishou; ZDL, Zhangdu lake; CL, Chao lake; TH, Tai lake), Yellow River (NX, Ningxia), Huai River (RR, Ru River; HZL, Hongze lake), Songhua River (DQ, Daqing; JPL, Jingpo lake) and Pearl River (WR, You River). As the Ningxia and Sichuan regions assumed to be the sources of goldfish introduced into Tibet, samples were collected from these two regions across a wide geographical range; the coordinate of each sample can be found in Table S1. Here populations and individuals sampled from invasive and native ranges are denoted "invasive" and "native" populations and individuals, respectively. A small piece of fin tissue was sampled from each fish individual and stored in absolute ethanol until DNA extraction.

Genomic DNA was extracted from fin clips using a phenol-chloroform method (Taggart et al., 1992). At least 0.5 μg of genomic DNA from each sample was used to construct the resequencing library with an insert size of ˜350 bp using the NEBNext Ultra DNA Library Prep Kit for 150 bp paired-end sequencing on the Illumina X Ten platform (Illumina USA). Sequencing was outsourced to Biomarker Technologies Corporation (Beijing, China). In addition, resequencing data of 16 individuals from the native range were downloaded from a previous study (Chen et al., 2020). A total of 151 goldfish genomes were included in this study. Additional information of each sample is shown in Table S1.

2.2 Read mapping, ploidy identification and variant calling

The raw reads were filtered using fastp 0.20.0 (Chen et al., 2018) (with parameters -3 -W 4 -M 20 -q 20 -u 40 -c -l 50) to eliminate reads with low quality or short length. The clean reads were aligned to the recent chromosome-level goldfish reference genome (Luo et al., 2020) using the BWA-MEM algorithm from BWA (Li, 2012).The resulting SAM file was converted to a coordinate-sorted and indexed BAM file using Samtools (Li et al., 2009). Duplicated read pairs were removed using Sambamba (Tarasov et al., 2015).

The ploidy level of each sample was determined by the nQuire software (Weiß et al., 2018) using the BAM file. The denoised input of base frequencies at biallelic sites generated with default parameters was utilized to estimate ploidy levels with a Gaussian Mixture Model which describes the histogram as a mixture of Gaussians distributions between 0 and 1. The expected distributions of read frequencies at biallelic sites are one Gaussian with mean 0.5 for diploid, two Gaussians with means 0.33 and 0.67 for triploid. The likelihood of certain assumptions of di, tri- and tetraploidy based on this model given the empirical data is maximized using an Expectation-Maximization (EM) algorithm.

Variant calling was carried out for diploids and triploids separately using the Genome Analysis Toolkit (GATK version 4.2.2.0) (McKenna et al., 2010), with the parameter of -ploidy being 2 for diploids and 3 for triploids. Briefly, the genomic variants for each sample (in GVCF format) were identified with the HaplotypeCaller module. The GVCF files of all samples were merged into a single GVCF with the CombineGVCFs module, followed by variant calling to generate a genotype file (in VCF format) using the GenotypeGVCFs module based on a joint calling approach. The SNPs were selected using the SelectVariants module and were initially filtered with the VariantFiltration module using the parameters 'QUAL $<$ 30.0 || QD $<$ 2.0 || FS $>$ 60.0 || MQ $<$ 40.0'. The SNPs were further filtered using the VCFtools software (Danecek et al., 2011) for diploids and BCFtools software (Danecek et al., 2021) combined with a custom script for triploids using the following criteria: (1) SNPs with more than two alleles, (2) genotype calls with a depth $<$ 4 or $>$ 24, (3) minor allele frequency (MAF) $<$ 0.05, (4) SNPs with missing rates $>$ 0.1.

2.3 Assembly of mitogenomes and phylogenetic analysis

To assess the phylogenetic relationships among individuals, maximum-likelihood phylogenetic trees were constructed for diploids and triploids separately, as well as all individuals together, using the IQ-TREE software version 2.0.3 (Nguyen et al., 2015), based on their complete mitochondrial genome sequences. The mitogenome sequences were*de novo* assembled from the resequencing data using procedures described in Liu et al. (2016). Briefly, the first 5 M paired-end reads were selected for further analysis to fragmentize the assembly

of nuclear sequences. *De novo* assembly of relatively short sequences (e.g. mitogenomes) was conducted using Spades v3.13.0 (Bankevich et al., 2012). The mitogenome sequence was then captured from the assembly using a probe sequence which is conserved in *C. auratus*(GCTAGCGTAGCTTAATACAAAGC).

2.4 Population structure analysis

As linkage disequilibrium (LD) may affect the inference of population structure, the diploid SNPs were firstly filtered with LD (r^2) < 0.2 using PLINK 1.9 (Chang et al., 2015) with the parameters '–indep-pairwise 100 10 0.2'. To analyze data of diploids and triploids together, the pruned diploid SNPs were then compared with the triploid SNPs using the isec function in BCFtools (Danecek et al., 2021), and the intersection of SNPs was used for population structure analysis. Three methods were used to infer population genetic structure including principal component analysis (PCA), structure analysis and genetic distance analysis. The first method was used for diploids and triploids separately as well as all samples together, while the other two methods were used for diploids and triploids separately.

For PCA, the genotype data at each locus was firstly converted into the frequency of the reference allele, that is, 0/0.5/1 for diploids and 0/0.33/0.67/1 for triploids. The PCA was then performed using the R built-in function *prcomp* with default parameters. The STRUCTURE software Version 2.3.4 (Pritchard et al., 2000) was used for genetic structure analysis with five run times for each K value ranging from 1 to 12. The optimal K , which indicates the most likely number of genetic clusters, was determined according to the method described in Evanno et al. (2005) (Evanno et al., 2005). For genetic distance analysis, the identity-by-state (IBS) which describes the genetic relationship among individuals was calculated using a custom R script. The minimum evolution phylogeny trees were constructed based on the genetic distance matrix of 1-IBS values using the FastME program (Lefort et al., 2015) and visualized using the online tool iTOL (http://itol.embl.de) (Letunic & Bork, 2019).

Pairwise genetic differentiation (F_{ST}) and tests for significance were estimated for invasive populations and regionally defined genetic clusters of native populations using the R package*StAMPP* (Pembleton et al., 2013). Additionally, genetic differentiation among populations was analyzed by an Analysis of Molecular Variance (AMOVA) using 100 permutations, where the variance components were partitioned between regions (invasive and source ranges), among populations within regions and within populations.

2.5 Individual and population genetic diversity

Genome-wide genetic diversity was measured at both individual and population level. For the individual level, genome-wide observed heterozygosity (H_O) was calculated by averaging the values of both variant and non-variant sites. The genotype data that include both variant and non-variant sites were re-analysed, and mono/bi-allelic SNPs with read depth ranging from 4 to 24 were retained for estimation of genetic diversity. H_O at each locus was calculated as the generalized gametic heterozygosity which estimates the probability of randomly sampling two different alleles from an individual (Meirmans et al., 2018; Moody et al., 1993). Hence, a bi-allelic SNP had a H_O value of 1 for diploids, and 2/3 for triploids because of three homologous chromosomes. For population level, genome-wide nucleotide diversity (π) was calculated using the R package *PopGenome* (Pfeifer et al., 2014) with a 10 kb window size.

2.6 Detection ofruns of homozygosity (ROH)

To assess whether inbreeding occurred in invasive populations, the genome-wide inbreeding level of each individual was estimated by detecting the total length and number of ROHs using the R package*detectRUNs* (Biscarini et al., 2018). As ROHs were detected only based on the genotype state (homozygous/heterozygous), the triploid VCF file was converted into diploid format which could be analyzed with*detectRUNs* . The following constraints were applied to the sliding-window-based run detection: (1) The minimum number of SNPs in a ROH was 138 for diploids and 116 for triploids which was determined by the method described in Purfield et al. (2012) (Purfield et al., 2012). (2) One heterozygote and two missing genotypes were allowed in a window. (3) The maximum gap between consecutive SNPs was 300 kb. (4) The minimum length of a run was 500 kb. (5) The minimum SNP density was set to one SNP per 100 kb. The genomic inbreeding

coefficient based on ROHs (F_{ROH}) was calculated following the method described in McQuillan et al. (2008) (McQuillan et al., 2008):$F_{\text{ROH},i} = L_{\text{ROH}}/L_{\text{AUTO}}$, where$L_{\text{ROH}}$ is the total length of ROHs in individual$i$, and L_{AUTO} is the length of autosomal genome, here assumed to be 1.54 Gb. The ROH segment length (E ($L_{IBD\text{-}H}$ $|gcA$)) represents the number of generations from the common ancestor (gcA), estimated as$E$ ($L_{IBD\text{-}H}$ $|gcA$)= $41/(2gcA$) (Curik et al., 2014) assuming that 1 cM is roughly equivalent to 0.41 Mb (Liu et al., 2017a).

2.7 Demographic history reconstruction

Demographic history was reconstructed, both to assess ancient population history and detect possible recent population declines associated with founder events in invasive populations. As triploids exhibit unisexual gynogenesis, the demographic history was only constructed for diploid populations. The fluctuation of effective population size was estimated using SMC++ (version 1.15.4) which combines the computational efficiency of site frequency spectra (SFS) and the advantage of using LD information in coalescent HMMs and provides a high resolution of recent demography (Terhorst et al., 2017). The VCF file for each population was converted to the SMC input format using the vcf2smc module. The population size was estimated using the estimate module with the option –knots 20 and a per-generation mutation rate of 1e-8. The plot module was used for plotting fitted size histories using a generation length of one year.

2.8 Scanning for putative signatures of selection

Regions in the genome showing signatures of selective sweeps were detected by comparing the invasive and source populations using two different approaches. The first method estimated the Fixation index (F_{ST}) values between populations by using the R package *PopGenome* (Pfeifer et al., 2014). Loci under selection should exhibit large F_{ST} values. The second method calculated the absolute XP-EHH (Cross Population Extended Haplotype Homozygosity) scores to detect selective sweeps in which the selected allele has approached or achieved fixation in one population but remains polymorphic across both populations (Sabeti et al., 2007). Phasing of each individual's genotype data was inferred using WhatsHap (Patterson et al., 2015) and Shapeit4 (Delaneau et al., 2019) and XP-EHH scores were calculated with the R package *rehh* (Gautier & Vitalis, 2012). A sliding window of 10 kb was applied for both of the methods, and windows with values above the 99th percentile of the empirical distribution and with more than 10 SNPs were retained for further analyses (Pujolar et al., 2022). Overlapping windows between the two methods were considered as potential regions under selection. The candidate genes were obtained by the SNP annotation within the candidate regions using the ANNOVAR software (Wang et al., 2010). Gene Ontology (GO) enrichment analysis of these genes were performed using the hypergeometric Fisher exact test in an online tool (OmicShare, www.omicshare.com/tools).

3 Results

3.1 Ploidy identification and variant calling

Whole genome resequencing of the 151 individuals generated a total of ˜ 10.6 G of clean paired-end reads (70.2 M per individual), giving an average sequencing depth of 12.2× (Table S1). Among them, 99.14% of reads were on average mapped to the reference genome of goldfish. The analysis of allele frequency distributions at biallelic variants clearly showed one Gaussian distribution with mean 0.5 for diploids and bimodal distributions with means 0.33 and 0.67 for triploid (Figure S1). Among all individuals, 73 and 78 were identified as diploids and triploids, respectively. In invasive populations, the number of diploids and triploids were 29 and 9 in the LL population, and 2 and 22 in the CBL population, respectively (Table 1). In direct source regions, samples from Ningxia and Sichuan were almost all triploids with only one diploid detected from Sichuan. In indirect source regions, 29 diploids and 6 triploids were detected. After variant calling and filtering, a total of 16,888,283 and 17,954,020 SNPs were identified from diploids and triploids, respectively, with 10,719,815 SNPs being common to both ploidy categories.

3.2 Phylogenetic relationships based on mitogenome sequences

In the phylogenetic trees constructed based on mitogenome sequences, four distinct lineages were represented

among diploids (Figure 2a). They followed a pattern of geographical origin for the native samples, with the invasive individuals from Tibet belonging to two of the lineages. The first lineage (A) mainly contained 24 invasive individuals and source individuals from the lower Yangzte River (TH and CL) and Huai River (RR and HZL). The single individual from SC and one individual from DQ also belonged to lineage A. The second lineage (B) was composed of the remaining seven invasive individuals and source individuals from the mid Yangzte River (SS and ZDL). The third lineage (C) contained eight native individuals from the Songhua River (DQ and JPL), and the single individual from Fujian separately formed the fourth lineage (D).

For triploids, four lineages E, F, G and H were observed with 59, 15, 3 and 1 individuals being included, respectively (Figure 2b). However, the phylogenetic pattern of triploids was not related to geographical origin, with each large lineage containing individuals from different river systems and including invasive individuals. By combining diploids and triploids together, four lineages were found with A, B, E and H being clustered into one super lineage, C and F into another big lineage, D and G into single lineages (Figure S2). Hence, the phylogenetic pattern was not associated with the ploidy level.

3.3 Population genetic structure and differentiation

After LD filtering, 1,545,006 SNPs common to di- and triploids were selected and used for population structure analysis. The three methods including PCA, structure analysis and genetic distance analysis all showed the same patterns of population genetic structure (Figure 3 and 4, and Figure S3). For diploids, two clear genetic clusters were shaped and consisted of individuals from invasive and native populations, respectively. The genetic cluster owned by native individuals split into three subclusters and followed a pattern of geographical separation: individuals from the mid Yangzte River (SS and ZDL); individuals from the lower Yangzte River (TH and CL) and Huai River (RR and HZL) and individuals from the Songhua River (DQ and JPL). For triploids, although several genetic clusters were observed, the pattern was not related to the geographical origin, with each cluster containing individuals from different rivers. The invasive triploids were assigned into different clusters. Results of PCA for all individuals revealed that both PCA1 and PCA2 did not clearly separate diploids and triploids, indicating patterns of population structure were not linked to the ploidy level.

For diploids, highly significant genetic differentiation (F_{ST}) was detected for all pairwise comparisons among the invasive population (LL) and three genetic subclusters from the native range (Table 2). The F_{ST} values between LL and two source genetic subclusters (i.e., the mid Yangzte River, the lower Yangzte River and Huai River) were 0.1045 and 0.0926, respectively, which were both higher than that between the two subclusters (0.0449). AMOVA analysis revealed that genetic variation between regions (invasive and source), among populations within regions and within populations were all significant and accounted for 6.29%, 8.65% and 85.06% of the total variation ($P < 0.01$), respectively. For triploids, low F_{ST} values ranging from 0.0073 to 0.0281 ($P > 0.05$) were detected for pairwise comparisons between invasive (LL and CBL) and source (SC and NX) populations and a relative higherF_{ST} (0.0345) was obtained between the two source populations ($P < 0.01$). The variation between regions, among populations within regions and within populations accounted for -2.93%, 4.41% and 98.52%, respectively. Overall, theF_{ST} values between invasive and source populations were 0.0866 in diploids and 0.0027 in triploids.

3.4 Genome-wide genetic diversity

At the individual level for diploids (Figure 5a), the genome-wide observed heterozygosity (H_O) values of invasive individuals (mean: 0.00861) was significantly lower than those of native individuals (mean: 0.01062) (paired-sample t -test, $p = 2.2e^{-16}$). For triploids, $H_O$values of invasive individuals (mean: 0.00922) was comparable to those of the native individuals (mean: 0.00914). Compared with triploids, invasive diploids showed significantly lower observed heterozygosity than the invasive ($p = 5.1e^{-5}$) and native triploids ($p = 2.5e^{-7}$), while native diploids had significant higher genetic diversity than the invasive ($p = 5.4e^{-10}$) and native triploids ($p = 3.2e^{-12}$).

As population structure inferred three genetic subclusters among diploid native samples, the population genetic diversity was estimated for these subclusters. For diploids, the genome-wide nucleotide diversity (π)

value of the invasive population LL was 0.00297, which was slightly higher than those of native subclusters (0.00266 for mid Yangzte River, 0.00279 for lower Yangzte River and Huai River, and 0.00251 for Songhua River) (Figure 5b and Table 3). For triploids, π values of the invasive populations LL and CBL were 0.00286 and 0.00284, respectively, which were also slightly higher than those of source populations (0.00273 for SC and 0.00277 for NX).

3.5 Estimation of inbreeding level

A total of 78 ROHs were detected, with almost all of them (76) found in 25 invasive diploid individuals from the LL population (Figure 6 and Table S1). The number of ROHs per individual ranged from one to eight (mean: 3.04), and the total length ranged from 0.518 to 5.989 Mb (mean: 2.129 Mb). The genomic inbreeding coefficient (F_{ROH}) was low, ranging from 0.0003 to 0.0038 across individuals, with an average value of 0.0014. The ROHs of a length from 0.518 to 5.989 Mb suggested inbreeding events that have occurred 3-39 generations ago.

3.6 Recent demographic history

The results of SMC++ analysis revealed a difference in recent demographic history between invasive and native populations (Figure 7). Among all diploid populations, only the invasive population (LL) experienced a recent decline in effective population size reflecting founder events, while all native populations experienced a recent expansion during the same time period. For more ancient time periods, the population trajectories were almost similar across all populations, suggesting an expansion since about 150 ka followed by a peak at about 60 ka and a subsequent decline, assuming mutation rate of 1e-8 and generation length of one year.

3.7 Putative signatures of selection

The 99th percentile threshold values of the F_{ST} and |XP-EHH| were used to detect the putative signatures of selective sweeps between invasive and source populations. These thresholds were 0.300 and 2.399 for diploids, and 0.021 and 2.021 for triploids, respectively (Figure 8). For diploids, 324 overlapping outlier windows between the two methods were detected on 37 chromosomes, with most outliers on Chr24 (66) followed by Chr13 (38), Chr9 (34) and Chr30 (29) (Table S2). A total of 16,754 SNPs were observed within these outlier regions, among which 4.16% and 52.75% were annotated to the exonic and intronic regions, respectively, with the percentages being higher than those found in the whole genome (2.51% and 49.56%). The nucleotide diversity within these outliers in the source range was more than two times higher than that in the invasive range. The SNPs within exonic and intronic regions were annotated to 191 genes (Table S3). GO enrichment of these genes revealed one enriched cluster namely 'mannosyl-oligosaccharide mannosidase activity' (Pvalue < 0.01 and Qvalue < 0.05) involving in three genes (*Endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase* ,*Mannosyl-oligosaccharide 1,2-alpha-mannosidase IB* and*Alpha-mannosidase 2*). Other enriched clusters with Pvalue < 0.01 but Qvalue > 0.05 were mainly involved in mannosidase and embryo development (Table S4).

For triploids, 244 overlapping outliers between the two methods were observed on 40 chromosomes. However, these outliers only formed visual peaks on Chr14, Chr15, Chr18 and Chr35. Among 2,543 SNPs within these outliers, 3.54% and 33.23% were annotated to exonic and intronic regions, respectively, with the percentages being lower than those found in the whole genome (3.85% and 49.31%). These SNPs were annotated to 60 genes with enriched GO terms related to metabolic process and immune response (Table S4). Interestingly, one enriched GO term was 'development of secondary sexual characteristics' with two involved genes (*Beta-1,4-galactosyltransferase 1* and *Signal transducer and activator of transcription 5B*). When comparing outliers for diploids and triploids, only four were overlapped on Chr2 and Chr18 with all SNPs being annotated to intergenic regions.

4 Discussion

The results of our study document a highly different genetic and phylogeographic structure of diploid and triploid goldfish, particularly reflected in a clear geographic structure in diploids and no obvious association between structure and geographical origin in triploids. Also, and in accordance with our main hypothesis,

we found a major difference in patterns of genetic variation between invasive populations of diploids and triploids in Tibet, with decreased variation in diploids and evidence for recent population decline reflecting founder events, but in contrast no detectable loss of variation in triploids. Nevertheless, inbreeding was low in diploids suggesting that multiple introductions from diverse sources have buffered bottlenecks associated with founding. Finally, we found evidence for selective sweeps in invasive diploid goldfish, with genes within outlier regions enriched for functions related to mannosidase activity and embryo development. We discuss these findings in more detail below.

4.1 Comparison between diploids and triploids in the native range

Our analyses of genetic diversity and structure between diploid and triploid *C. auratus* showed that overall nucleotide diversity of diploids was comparable to that of triploids based on both mitogenome and nuclear genome (Table 3). This is inconsistent with results of previous studies showing a slightly higher genetic diversity in triploids than diploids (Liu et al., 2017b; Luo et al., 2014). This may reflect the fact that our sampling work mainly focused on the native range of source populations for introduction into Tibet, with a higher percentage of diploids than triploids, while triploids actually have wider geographic distribution than diploids (Liu et al., 2017b; Liu et al., 2017c). By using the same procedures of variant calling, total number of detected SNPs in triploids was slightly higher than that in diploids, which is consistent with a previous study based on SNPs from transcriptome data (Ren et al., 2018). Furthermore, at the individual level, although the mean observed heterozygosity of diploids was significantly higher than that of triploids when considering ploidy level (Figure 5), the mean frequency of heterozygous loci across genomes in diploids was significantly lower than that in triploids. Similarly in other unisexual organisms that are derived from auto-polyploidization (uneven ploidy in general), they show levels of genetic diversity close to that of their sexual diploid counterparts (Joly & Bruneau, 2004; Lee et al., 2016; Liu et al., 2015b). However, unisexual organisms with interspecific hybrid origins, such as vertebrate *P. formosa* and*P. eos-neogaeus* and some plants, usually have a much higher genetic diversity than their sexual parental species (Angers & Schlosser, 2007; Robertson et al., 2010; Vallejo-Marin & Lye, 2013; Warren et al., 2018). Hence, genetic diversity of unisexual organisms depends on their specific origin relative to related sexual species.

Previous studies revealed that triploid *C. auratus* have undergone multiple independent polyploidy origins from sympatric diploids (Liu et al., 2017b; Luo et al., 2014). Our results also supported this point by showing that native diploids were genetically clustered with triploids within some lineages. Based on mitogenomes, diploids and triploids coexist in almost each lineages, indicating the same origin within the species. However, by using genomic SNPs, diploids were clustered with triploids only within two clusters (Figure 3), and triploids shaped three more clusters. Given that the genetic variation of triploids remains almost frozen across generations because of unisexual gynogenesis (Wang et al., 2022), we assumed that the two shared clusters reflected recent triploidization events and the three triploid-specific clusters originated from more ancient triploidization events. Genetic structure of contemporary diploids linked to a geographical separation pattern should be shaped by local adaptation or other factors. However, the genetic structure of triploids is accumulated by previous triploidization from diploids. Based on results of this study, the triploid *C. auratus* has at least five polyploidy origins and might possess more origins when investigating more samples from a broader range of locations.

4.2 Genetic changes in invasive diploid populations

The loss of genetic diversity within invasive populations resulting from genetic bottlenecks can be prevented or even reversed by multiple introductions from the same or distinct sources, which have been reported in many animals and plants, such as *Pseudorasbora parva* ,*Coptotermes formosanus* , *Aedes albopictus* , *Ambrosia artemisiifolia* and *Rosa rugosa* (Baltazar-Soares et al., 2020; Blumenfeld et al., 2021; Kelager et al., 2013; Sherpa et al., 2019; van Boheemen et al., 2017). The invasive *C. auratus* in Tibet was introduced from different sources, which was also evidenced by phylogenetic analysis of diploid mitogenomes in this study, showing that invasive individuals were clustered into two regionally defined lineages of native individuals (i.e., the mid Yangzte River, the lower Yangzte River and Huai River). The two lineages showed a low but significant genetic differentiation (F_{ST}: 0.0449) estimated based on genomic SNPs. In general, recent hybrid

populations from diverse intraspecific source populations should possess higher genetic diversity than that of individual source populations (Baltazar-Soares et al., 2020; Simon et al., 2011; Smith et al., 2020). However, the genetic diversity of invasive diploid population of *C. auratus* was only slightly (9.0%) higher than that of source populations. Interestingly, at the individual level, the genome-wide observed heterozygosity of invasive individuals was significantly (21.4%) lower than that of source individuals (Figure 5a). The discordance between genetic diversity at the population and individual levels suggests that the total invasive population derives from genetically distinct sources, which may mask loss of variation at the population level during invasion. Nevertheless, the observed low levels of inbreeding suggest that founder events have still not had a significant biological impact. In that sense, the results provide an illustration of factors that may cause biological invasions to be associated with lower loss of genetic variation than expected (Blumenfeld et al., 2021; Facon et al., 2006). On a longer time scale, triploid *C. auratus* now dominate the market (Liu, 2010), and the LL wetland is padlocked since becoming a natural reserve zone since 1999 (Chen et al. 2018). Therefore, the diploid invasive population is expected to receive very low supplements of new invaders recently and in the future, which could accelerate loss of genetic diversity.

4.3 No genetic change in invasive triploid populations

Triploids are considered an evolutionary 'dead end' because they lack meiotic recombination and thus accumulate deleterious mutations (Butlin, 2002). Such 'disadvantage' in triploids may become an advantage in terms of avoiding inbreeding and thus being better colonizers than diploids (Avise, 2008). In the present study, the lack of change of genetic diversity and structure in invasive triploid populations is to be expected, given the pattern of asexual gynogenesis which produces clonal offspring from mothers (Wang et al., 2022). This reproductive mode may make triploid *C. auratus* a more powerful invader in novel environments compared with the sexual diploid form. Firstly, parthenogenesis enables triploids to avoid negative impacts of bottlenecks and inbreeding on genetic diversity during invasion (Avise, 2008; Roman & Darling, 2007). Secondly, eggs of triploids can be activated by the sperm of both conspecific and non-conspecific fishes, providing more chances for reproduction. Concerning adaptability, the pros and cons of invaders being triploid are less clear-cut. On one hand, triploids possess slightly higher genetic diversity and frequency of heterozygous loci (Liu et al., 2017b; Luo et al., 2014) which would be an advantage in the case of overdominance. On the other hand, the lack of meiosis and sexual reproduction precludes reshuffling of allelic variation at loci influencing polygenic traits, hence precluding adaptive responses. As a whole, experiences from different organisms do suggest stronger invasiveness of asexual forms as compared to sexual forms (Donne et al., 2020; Pandit et al., 2011).

4.4 Candidate genes within genomic regions of putative sweeps

The regions with potential selective sweeps for invasive diploid populations of *C. auratus* were widely detected across the genome. The SNPs within these regions were annotated to 191 candidate genes with enriched GO terms mainly related to mannosidase activity and embryo development (Table S4). The biological significance of selection in genes related to mannosidase are not immediately obvious, although in humans mannosidase deficiency can cause disease, such as congenital dyserythropoietic anemia (Moremen, 2002). The selective sweeps related to genes involved in embryo development may be attributed to an advancement of spawning time in Tibet (Liu et al., 2015a) and subsequent exposure of embryos to the cold waters. The water temperature during reproduction period is about 20 °C in the native range (e.g. the mid and lower Yangtze River) and is below 10 °C in the invasive range (Huo et al., 2013; Zhang et al., 2016). Low temperature exposure can increase the mortality and frequency of abnormal larvae in *C. auratus*(Wiegand et al., 1989), resulting in individual selection during embryonic development.

For triploids, selection occurs at the whole individual level owing to the asexual gynogenesis and would be expected to be low, as numbers of mitogenome haplotypes and genetic diversity in invasive populations were not reduced as compared to the source populations. Hence, it is doubtful whether signals of selective sweeps found in invasive triploids were actual targets of selection; more likely these signals would derive from the process of establishment of triploids from diploid ancestors. It is nevertheless interesting that one enriched GO term found in triploids is 'development of secondary sexual characteristics'. A previous study has shown

that the ratio of males among invasive triploid *C. auratus*in Croatia was only 2.3% which was lower than that reported in the native range (Gui & Zhou, 2010; Jakovlic & Gui, 2011). It would be of interest to further investigate the ratio of males in Tibet.

5 Conclusion

Our study represents a unique case of diploid and triploid individuals from the same species invading an extreme high-altitude environment. The results demonstrated profoundly different outcomes of the invasion process at the genomic level for diploids and triploids. For diploids, in spite of introductions from two genetically distinct sources, the invasive individuals possessed significantly lower heterozygosity than source individuals and formed a genetic cluster distinct from the source populations. Demographic history reconstruction also documented recent population declines coinciding with invasion. Nevertheless, inbreeding was still low, presumably reflecting buffering effects of multiple continuous events of introduction. For triploids, no change in genetic diversity and structure was found in invasive populations, which is owing to their reproduction mode of gynogenesis that precludes inbreeding and reduces founder effects. Finally, there evidence for selective sweeps in invasive diploid goldfish, pointing to a role of mannosidase activity that nevertheless requires further clarification and embryo development, the latter of which may play important roles for adaptive processes of individuals introduced into this extreme environment.

Acknowledgments

This study was supported by the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (2019QZKK0501 and 2019QZKK0304), the National Natural Science Foundation of China (31601844), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA20050204) and the China Scholarship Council (CSC, 202004910172). We would like to thank Prof. Yaping Zhang from Chinese academy science, Prof. Shaojun Liu from Hunan Normal University, Prof. Jing Luo from Yunnan University and Dr. Zhen Huang from Fujian Normal University for their assistance in data collection and analysis.

Data Availability Statement

The genome resequencing data have been deposited in the Genome Sequence Archive (GSA; http://gsa.big.ac.cn/) under accession number PRJCA010043.

Author Contributions

YC, XS and XF conceived and designed the study. XF, RZ, YJ, XY, and CL collected fish samples and conducted experiments. XF, SL, JT and MMH performed the data analysis. XF drafted the manuscript, and SL, XS, YC and MMH revised the manuscript. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Reference

Angers, B., & Schlosser, I. J. (2007). The origin of Phoxinus eos-neogaeus unisexual hybrids. *Molecular Ecology, 16* (21), 4562-4571.*https://doi.org/10.1111/j.1365-294X.2007.03511.x*

Austin, J. W., Szalanski, A. L., Scheffrahn, R. H., Messenger, M. T., Mckern, J. A., & Gold, R. E. (2006). Genetic evidence for two introductions of the Formosan Subterranean Termite, *Coptotermes formosanus* (Isoptera : Rhinotermitidae), to the United States.*Florida Entomologist, 89* (2), 183-193.*https://doi.org/10.1653/0015-4040(2006)89[183:GEFTIO]2.0.CO;2*

Avise, J. C. (2008). *Clonality : the genetics, ecology, and evolution of sexual abstinence in vertebrate animals* . Oxford: Oxford University Press.

Baltazar-Soares, M., Blanchet, S., Cote, J., Tarkan, A. S., Zahorska, E., Gozlan, R. E., & Eizaguirre, C. (2020). Genomic footprints of a biological invasion: Introduction from Asia and dispersal in Europe of the topmouth gudgeon (*Pseudorasbora parva*). *Molecular Ecology, 29* (1), 71-85.*https://doi.org/10.1111/mec.15313*

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology, 19* (5), 455-477.*https://doi.org/10.1089/cmb.2012.0021*

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics, 49* , 433-456.*https://doi.org/10.1146/annurev-ecolsys-110617-062431*

Biscarini, F., Cozzi, P., Gaspa, G., & Marras, G. (2018). detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes. *R package version 0.9.6* ,*https://CRAN.R-project.org/package=detectRUNS*.

Blumenfeld, A. J., Eyer, P. A., Husseneder, C., Mo, J. C., Johnson, L. N. L., Wang, C. L., Grace, J. K., Chouvenc, T., Wang, S. C., & Vargo, E. L. (2021). Bridgehead effect and multiple introductions shape the global invasion history of a termite. *Communications Biology, 4* (1), 196.*https://doi.org/10.1038/s42003-021-01725-x*

Butlin, R. (2002). The costs and benefits of sex: new insights from old asexual lineages. *Nature Reviews Genetics, 3* (4), 311-317.*https://doi.org/10.1038/nrg749*

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics, 19* (4), 220-+.*https://doi.org/10.1038/nrg.2017.109*

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4* , 7.*https://doi.org/10.1186/s13742-015-0047-8*

Chen, D., Zhang, Q., Tang, W. Q., Huang, Z., Wang, G., Wang, Y. J., Shi, J. X., Xu, H. M., Lin, L. Y., Li, Z., Chi, W. C., Huang, L. K., Xia, J., Zhang, X. T., Guo, L., Wang, Y. Y., Ma, P. P., Tang, J., Zhou, G., Liu, M., Liu, F. Y., Hua, X. T., Wang, B. Y., Shen, Q. C., Jiang, Q., Lin, J. X., Chen, X. Q., Wang, H. B., Dou, M. J., Liu, L., Pan, H. R., Qi, Y. Y., Wu, B., Fang, J. P., Zhou, Y. T., Cen, W., He, W. J., Zhang, Q. J., Xue, T., Lin, G., Zhang, W. C., Liu, Z. J., Qu, L. M., Wang, A. M., Ye, Q. C., Chen, J. M., Zhang, Y. D., Ming, R., Van Montagu, M., Tang, H. B., Van de Peer, Y., Chen, Y. Q., & Zhang, J. S. (2020). The evolutionary origin and domestication history of goldfish (*Carassius auratus*). *Proceedings of the National Academy of Sciences of the United States of America, 117* (47), 29775-29785.*https://doi.org/10.1073/pnas.2005545117*

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics, 34* (17), 884-890.*https://doi.org/10.1093/bioinformatics/bty560*

Curik, I., Ferencakovic, M., & Solkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livestock Science, 166* , 26-34.*https://doi.org/10.1016/j.livsci.2014.05.034*

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., McVean, G., Durbin, R., & Grp, G. P. A. (2011). The variant call format and VCFtools. *Bioinformatics, 27* (15), 2156-2158.*https://doi.org/10.1093/bioinformatics/btr330*

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools.*Gigascience, 10* (2), giab008.*https://doi.org/10.1093/gigascience/giab008*

Delaneau, O., Zagury, J. F., Robinson, M. R., Marchini, J. L., & Dermitzakis, E. T. (2019). Accurate, scalable and integrative haplotype estimation. *Nature Communications, 10* , 5436.*https://doi.org/10.1038/s41467-019-13225-y*

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology, 17* (1), 431-

449.*https://doi.org/10.1111/j.1365-294X.2007.03538.x*

Dong, F., Kuo, H. C., Chen, G. L., Wu, F., Shan, P. F., Wang, J., Chen, D., Lei, F. M., Hung, C. M., Liu, Y., & Yang, X. J. (2021). Population genomic, climatic and anthropogenic evidence suggest the role of human forces in endangerment of green peafowl (*Pavo muticus*).*Proceedings of the Royal Society B-Biological Sciences, 288* (1948), 20210073.*https://doi.org/10.1098/rspb.2021.0073*

Donne, C., Neiman, M., Woodell, J. D., Haase, M., & Verhaegen, G. (2020). A layover in Europe: Reconstructing the invasion route of asexual lineages of a New Zealand snail to North America.*Molecular Ecology, 29* (18), 3446-3465.*https://doi.org/10.1111/mec.15569*

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution, 29* (1), 51-63.*https://doi.org/10.1016/j.tree.2013.09.008*

Estoup, A., Ravigne, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is There a Genetic Paradox of Biological Invasion?*Annual Review of Ecology, Evolution, and Systematics, 47* , 51-72.*https://doi.org/10.1146/annurev-ecolsys-121415-032116*

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14* (8), 2611-2620.*https://doi.org/10.1111/j.1365-294X.2005.02553.x*

Facon, B., Genton, B. J., Shykoff, J., Jarne, P., Estoup, A., & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution, 21* (3), 130-135.*https://doi.org/10.1016/j.tree.2005.10.012*

Favre, A., Packert, M., Pauls, S. U., Jahnig, S. C., Uhl, D., Michalak, I., & Muellner-Riehl, A. N. (2015). The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas.*Biological Reviews, 90* (1), 236-253.*https://doi.org/10.1111/brv.12107*

Feng, X., Jia, Y. T., Zhu, R., Chen, K., & Chen, Y. F. (2019). Characterization and analysis of the transcriptome in *Gymnocypris selincuoensis* on the Qinghai-Tibetan Plateau using single-molecule long-read sequencing and RNA-seq. *DNA Research, 26* (4), 353-363.*https://doi.org/10.1093/dnares/dsz014*

Gautier, M., & Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics, 28* (8), 1176-1177.*https://doi.org/10.1093/bioinformatics/bts115*

Gonzalez, A., Ronce, O., Ferriere, R., & Hochberg, M. E. (2013). Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Philosophical Transactions of the Royal Society B-Biological Sciences, 368* (1610), 20120404.*https://doi.org/10.1098/rstb.2012.0404*

Gui, J. F., & Zhou, L. (2010). Genetic basis and breeding application of clonal diversity and dual reproduction modes in polyploid*Carassius auratus gibelio* . *Science China-Life Sciences, 53* (4), 409-415.*https://doi.org/10.1007/s11427-010-0092-6*

Haasl, R. J., & Payseur, B. A. (2016). Fifteen years of genomewide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology, 25* (1), 5-23.*https://doi.org/10.1111/mec.13339*

He, D. K., Sui, X. Y., Sun, H. Y., Tao, J., Ding, C. Z., Chen, Y. F., & Chen, Y. Y. (2020). Diversity, pattern and ecological drivers of freshwater fish in China and adjacent areas. *Reviews in Fish Biology and Fisheries, 30* (2), 387-404.*https://doi.org/10.1007/s11160-020-09600-4*

Huo, B., Xie, C. X., Ma, B. S., Yang, X. F., & Huang, H. P. (2013). Reproductive biology of *Oxygymnocypris stewartii* in the Yarlung Zangbo River in Tibet, China. *Environmental Biology of Fishes, 96* (4), 481-493.*https://doi.org/10.1007/s10641-012-0031-4*

Jakovlic, I., & Gui, J. F. (2011). Recent invasion and low level of divergence between diploid and triploid forms of *Carassius auratus* complex in Croatia. *Genetica, 139* (6), 789-804.*https://doi.org/10.1007/s10709-011-9584-y*

Jia, Y., Liu, Y., Chen, K., Sun, H., & Chen, Y. (2019). Climate, habitat and human disturbance driving the variation of life-history traits of the invasive goldfish *Carassius auratus* (Linnaeus, 1758) in a Tibetan Plateau river. *Aquatic Invasions, 14* (4), 724-737.

Joly, S., & Bruneau, A. (2004). Evolution of triploidy in *Apios americana* (Leguminosae) revealed by genealogical analysis of the histone H3-D gene. *Evolution, 58* (2), 284-295.*https://doi.org/10.1111/j.0014-3820.2004.tb01645.x*

Kelager, A., Pedersen, J. S., & Bruun, H. H. (2013). Multiple introductions and no loss of genetic diversity: invasion history of Japanese Rose, *Rosa rugosa* , in Europe. *Biological Invasions, 15* (5), 1125-1141.*https://doi.org/10.1007/s10530-012-0356-0*

Le Moan, A., Roby, C., Fraisse, C., Daguin-Thiebaut, C., Bierne, N., & Viard, F. (2021). An introgression breakthrough left by an anthropogenic contact between two ascidians. *Molecular Ecology, 30* (24), 6718-6732.*https://doi.org/10.1111/mec.16189*

Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution, 17* (8), 386-391.*https://doi.org/10.1016/S0169-5347(02)02554-5*

Lee, S. I., Nguyen, X. T., Kim, J. H., & 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 & Genomics, 38* (5), 467-477.*https://doi.org/10.1007/s13258-016-0398-2*

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. *Molecular Biology and Evolution, 32* (10), 2798-2800.*https://doi.org/10.1093/molbev/msv150*

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research, 47* (W1), W256-W259.*https://doi.org/10.1093/nar/gkz239*

Li, H. (2012). Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics, 28* (14), 1838-1844.*https://doi.org/10.1093/bioinformatics/bts280*

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Proc, G. P. D. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25* (16), 2078-2079.*https://doi.org/10.1093/bioinformatics/btp352*

Liu, C. L., Chen, Y. F., Olden, J. D., He, D. K., Sui, X. Y., & Ding, C. Z. (2015a). Phenotypic Shifts in Life History Traits Influence Invasion Success of Goldfish in the Yarlung Tsangpo River, Tibet. *Transactions of the American Fisheries Society, 144* (3), 602-609.*https://doi.org/10.1080/00028487.2014.996668*

Liu, H. Y., Fu, B. D., Pang, M. X., Feng, X., Yu, X. M., & Tong, J. G. (2017a). A high-density genetic linkage map and QTL fine mapping for body weight in crucian carp (*Carassius auratus*) using 2b-RAD sequencing. *G3-Genes Genomes Genetics, 7* (8), 2473-2487.*https://doi.org/10.1534/g3.117.041376*

Liu, S. J. (2010). Distant hybridization leads to different ploidy fishes. *Science China-Life Sciences, 53* (4), 416-425.*https://doi.org/10.1007/s11427-010-0057-9*

Liu, S. L., Ferchaud, A. L., Gronkjaer, P., Nygaard, R., & Hansen, M. M. (2018). Genomic parallelism and lack thereof in contrasting systems of three-spined sticklebacks. *Molecular Ecology, 27* (23), 4725-4743.*https://doi.org/10.1111/mec.14782*

Liu, S. L., Hansen, M. M., & Jacobsen, M. W. (2016). Region-wide and ecotype-specific differences in demographic histories of threespine stickleback populations, estimated from whole genome sequences. *Molecular Ecology, 25* (20), 5187-5202.*https://doi.org/10.1111/mec.13827*

Liu, X. L., Jiang, F. F., Wang, Z. W., Li, X. Y., Li, Z., Zhang, X. J., Chen, F., Mao, J. F., Zhou, L., & Gui, J. F. (2017b). Wider geographic distribution and higher diversity of hexaploids than tetraploids in*Carassius* species complex reveal recurrent polyploidy effects on adaptive evolution. *Scientific Reports, 7* , 5395.*https://doi.org/10.1038/s41598-017-05731-0*

Liu, X. L., Li, X. Y., Jiang, F. F., Wang, Z. W., Li, Z., Zhang, X. J., Zhou, L., & Gui, J. F. (2017c). Numerous mitochondrial DNA haplotypes reveal multiple independent polyploidy origins of hexaploids in*Carassius* species complex. *Ecology and Evolution, 7* (24), 10604-10615.*https://doi.org/10.1002/ece3.3462*

Liu, Y. X., Xia, T., Zheng, Y. H., Zhi, Y. Q., & Zhou, J. (2015b). Genetic diversity and the population structure at two ploidy levels of Lycoris radiata as revealed by SCoT analysis. *Biochemical Systematics and Ecology, 62* , 106-114.*https://doi.org/10.1016/j.bse.2015.08.003*

Luo, J., Chai, J., Wen, Y. L., Tao, M., Lin, G. L., Liu, X. C., Ren, L., Chen, Z. Y., Wu, S. G., Li, S. N., Wang, Y. D., Qin, Q. B., Wang, S., Gao, Y., Huang, F., Wang, L., Ai, C., Wang, X. B., Li, L. W., Ye, C. X., Yang, H. M., Luo, M., Chen, J., Hu, H., Yuan, L. J., Zhong, L., Wang, J., Xu, J., Du, Z. L., Ma, Z. S., Murphy, R. W., Meyer, A., Gui, J. F., Xu, P., Ruan, J., Chen, Z. J., Liu, S. J., Lu, X. M., & Zhang, Y. P. (2020). From asymmetrical to balanced genomic diversification during rediploidization: Subgenomic evolution in allotetraploid fish.*Science Advances, 6* (22), eaaz7677.*https://doi.org/10.1126/sciadv.aaz7677*

Luo, J., Gao, Y., Ma, W., Bi, X. Y., Wang, S. Y., Wang, J., Wang, Y. Q., Chai, J., Du, R., Wu, S. F., Meyer, A., Zan, R. G., Xiao, H., Murphy, R. W., & Zhang, Y. P. (2014). Tempo and mode of recurrent polyploidization in the *Carassius auratus* species complex (Cypriniformes, Cyprinidae). *Heredity, 112* (4), 415-427.*https://doi.org/10.1038/hdy.2013.121*

Luo, Y., & Yue, P. (2000). Cyprininae. In P. Yue (Ed.), *Fauna Sinica, Osteichthyes, Cypriniformes III* (pp. 429-433). Beijing: Science Press.

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., & Bazzaz, F. A. (2000). Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications, 10* (3), 689-710.*https://doi.org/10.2307/2641039*

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data.*Genome Research, 20* (9), 1297-1303.*https://doi.org/10.1101/gr.107524.110*

McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., Campbell, H., & Wilson, J. F. (2008). Runs of Homozygosity in European Populations. *American Journal of Human Genetics, 83* (5), 658-658.*https://doi.org/10.1016/j.ajhg.2008.10.009*

Meirmans, P. G., Liu, S. L., & van Tienderen, P. H. (2018). The Analysis of Polyploid Genetic Data. *Journal of Heredity, 109* (3), 283-296.*https://doi.org/10.1093/jhered/esy006*

Moody, M. E., Mueller, L. D., & Soltis, D. E. (1993). Genetic-Variation and Random Drift in Autotetraploid Populations. *Genetics, 134* (2), 649-657.

Moremen, K. W. (2002). Golgi alpha-mannosidase II deficiency in vertebrate systems: implications for asparagine-linked oligosaccharide processing in mammals. *Biochimica et Biophysica Acta, 1573* (3), 225-235.*https://doi.org/10.1016/s0304-4165(02)00388-4*

Pandit, M. K., Pocock, M. J. O., & Kunin, W. E. (2011). Ploidy influences rarity and invasiveness in plants. *Journal of Ecology, 99* (5), 1108-1115.*https://doi.org/10.1111/j.1365-2745.2011.01838.x*

Patterson, M., Marschall, T., Pisanti, N., Van Iersel, L., Stougie, L., Klau, G. W., & Schonhuth, A. (2015).

WHATSHAP: Weighted Haplotype Assembly for Future-Generation Sequencing Reads. *Journal of Computational Biology, 22* (6), 498-509.*https://doi.org/10.1089/cmb.2014.0157*

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources, 13* (5), 946-952.*https://doi.org/10.1111/1755-0998.12129*

Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Molecular Biology and Evolution, 31* (7), 1929-1936.*https://doi.org/10.1093/molbev/msu136*

Prentis, P. J., Wilson, J. R. U., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science, 13* (6), 288-294.*https://doi.org/10.1016/j.tplants.2008.03.004*

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155* (2), 945-959.

Pujolar, J. M., Jacobsen, M. W., & Bertolini, F. (2022). Comparative genomics and signatures of selection in North Atlantic eels.*Marine Genomics, 62* , 100933.*https://doi.org/10.1016/j.margen.2022.100933*

Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *Bmc Genetics, 13* , 70.*https://doi.org/10.1186/1471-2156-13-70*

Reid, W. V., Mooney, H. A., Cropper, A., Capistrano, D., Carpenter, S. R., Chopra, K., Dasgupta, P., Dietz, T., Duraiappah, A. K., & Hassan, R. (2005). *Ecosystems and human well-being-Synthesis: A report of the Millennium Ecosystem Assessment* : Island Press.

Ren, L., Gao, X., Yang, C. H., Tan, H., Cui, J. L., Wang, S., Li, W. H., Zhang, C., Tao, M., Qin, Q. B., & Liu, S. J. (2018). Comparison of diploid and triploid *Carassius auratus* provides insights into adaptation to environmental change. *Science China-Life Sciences, 61* (11), 1407-1419.*https://doi.org/10.1007/s11427-017-9358-7*

Rius, M., Turon, X., Bernardi, G., Volckaert, F. A. M., & Viard, F. (2015). Marine invasion genetics: from spatio-temporal patterns to evolutionary outcomes. *Biological Invasions, 17* (3), 869-885.*https://doi.org/10.1007/s10530-014-0792-0*

Robertson, A., Rich, T. C. G., Allen, A. M., Houston, L., Roberts, C., Bridle, J. R., Harris, S. A., & Hiscock, S. J. (2010). Hybridization and polyploidy as drivers of continuing evolution and speciation in Sorbus. *Molecular Ecology, 19* (8), 1675-1690.*https://doi.org/10.1111/j.1365-294X.2010.04585.x*

Roman, J., & Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution, 22* (9), 454-464.*https://doi.org/10.1016/j.tree.2007.07.002*

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X. H., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., & Consortium, I. H. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature, 449* (7164), 913-U912.*https://doi.org/10.1038/nature06250*

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., & Weller, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics, 32* , 305-332.*https://doi.org/10.1146/annurev.ecolsys.32.081501.114037*

Sherpa, S., Blum, M. G. B., Capblancq, T., Cumer, T., Rioux, D., & Despres, L. (2019). Unravelling the invasion history of the Asian tiger mosquito in Europe. *Molecular Ecology, 28* (9), 2360-2377.*https://doi.org/10.1111/mec.15071*

Simon, A., Britton, R., Gozlan, R., Oosterhout, C., Volckaert, F. A. M., & Hanfling, B. (2011). Invasive Cyprinid Fish in Europe Originate from the Single Introduction of an Admixed Source Population Followed

by a Complex Pattern of Spread. *Plos One, 6* (6), e18560.*https://doi.org/10.1371/journal.pone.0018560*

Smith, A. L., Hodkinson, T. R., Villellas, J., Catford, J. A., Csergo, A. M., Blomberg, S. P., Crone, E. E., Ehrlen, J., Garcia, M. B., Laine, A. L., Roach, D. A., Salguero-Gomez, R., Wardle, G. M., Childs, D. Z., Elderd, B. D., Finn, A., Munne-Bosch, S., Baudraz, M. E. A., Bodis, J., Brearley, F. Q., Bucharova, A., Caruso, C. M., Duncan, R. P., Dwyerh, J., Gooden, B., Groenteman, R., Hamre, L. N., Helm, A., Kelly, R., Laanisto, L., Lonati, M., Moore, J. L., Morales, M., Olsen, S. L., Partel, M., Petry, W. K., Ramula, S., Rasmussen, P. U., Enri, S. R., Roeder, A., Roscher, C., Saastamoinen, M., Tack, A. J. M., Topper, J. P., Vose, G. E., Wandrag, E. M., Wingler, A., & Buckley, Y. M. (2020). Global gene flow releases invasive plants from environmental constraints on genetic diversity. *Proceedings of the National Academy of Sciences of the United States of America, 117* (8), 4218-4227.*https://doi.org/10.1073/pnas.1915848117*

Taggart, J. B., Hynes, R. A., Prodohl, P. A., & Ferguson, A. (1992). A Simplified Protocol for Routine Total DNA Isolation from Salmonid Fishes. *Journal of Fish Biology, 40* (6), 963-965.*https://doi.org/10.1111/j.1095-8649.1992.tb02641.x*

Tao, J., Kennard, M. J., Jia, Y. T., & Chen, Y. F. (2018). Climate-driven synchrony in growth-increment chronologies of fish from the world's largest high-elevation river. *Science of the Total Environment, 645* , 339-346.*https://doi.org/10.1016/j.scitotenv.2018.07.108*

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: fast processing of NGS alignment formats.*Bioinformatics, 31* (12), 2032-2034.*https://doi.org/10.1093/bioinformatics/btv098*

Terhorst, J., Kamm, J. A., & Song, Y. S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes.*Nature Genetics, 49* (2), 303-309.*https://doi.org/10.1038/ng.3748*

Vallejo-Marin, M., & Lye, G. C. (2013). Hybridisation and genetic diversity in introduced Mimulus (Phrymaceae). *Heredity, 110* (2), 111-122.*https://doi.org/10.1038/hdy.2012.91*

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia.*Molecular Ecology, 26* (20), 5421-5434.*https://doi.org/10.1111/mec.14293*

Wang, K., Li, M. Y., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.*Nucleic acids research, 38* (16), e164. https://doi.org/10.1093/nar/gkq603

Wang, Y., Li, X.-Y., Xu, W.-J., Wang, K., Wu, B., Xu, M., Chen, Y., Miao, L.-J., Wang, Z.-W., Li, Z., Zhang, X.-J., Yin, Z., Zhou, B.-T., Yang, Y.-L., Zhu, C.-L., Hu, M.-L., Zheng, J.-M., Feng, C.-G., Qiu, Q., Tian, L.-T., Lu, M., Peng, F., Lu, W.-J., Tong, J.-F., Tong, J.-G., Fu, B.-D., Yu, P., Ding, M., Gan, R.-H., Zhang, Q.-Q., Jian, J.-B., Zhang, C., He, W.-M., Yang, W., Zhao, Z.-C., Zhang, Q.-Q., Gao, Q., Xu, J.-Y., Bai, M.-Z., Zhang, Y.-P., Yang, H.-M., Fang, X.-D., Wang, W., Zhou, L., & Gui, J.-F. (2022). Comparative genome anatomy reveals evolutionary insights into a unique amphitriploid fish. *Nature Ecology & Evolution, 6* , 1354-1366.*https://doi.org/10.1038/s41559-022-01813-z*

Warren, W. C., Garcia-Perez, R., Xu, S., Lampert, K. P., Chalopin, D., Stock, M., Loewe, L., Lu, Y., Kuderna, L., Minx, P., Montague, M. J., Tomlinson, C., Hillier, L. W., Murphy, D. N., Wang, J., Wang, Z. W., Garcia, C. M., Thomas, G. C. W., Volff, J. N., Farias, F., Aken, B., Walter, R. B., Pruitt, K. D., Marques-Bonet, T., Hahn, M. W., Kneitz, S., Lynch, M., & Schartl, M. (2018). Clonal polymorphism and high heterozygosity in the celibate genome of the Amazon molly. *Nature Ecology & Evolution, 2* (4), 669-679.*https://doi.org/10.1038/s41559-018-0473-y*

Weiß, C. L., Pais, M., Cano, L. M., Kamoun, S., & Burbano, H. A. (2018). nQuire: a statistical framework for ploidy estimation using next generation sequencing. *Bmc Bioinformatics, 19* , 122.*https://doi.org/10.1186/s12859-018-2128-z*

Wiegand, M. D., Hataley, J. M., Kitchen, C. L., & Buchanan, L. G. (1989). Induction of Developmental

Abnormalities in Larval Goldfish,*Carassius-Auratus* L, under Cool Incubation Conditions.*Journal of Fish Biology, 35* (1), 85-95.*https://doi.org/10.1111/j.1095-8649.1989.tb03395.x*

Xiao, J., Zou, T. M., Chen, Y. B., Chen, L., Liu, S. J., Tao, M., Zhang, C., Zhao, R. R., Zhou, Y., Long, Y., You, C. P., Yan, J. P., & Liu, Y. (2011). Coexistence of diploid, triploid and tetraploid crucian carp (*Carassius auratus*) in natural waters. *Bmc Genetics, 12* , 20.*https://doi.org/10.1186/1471-2156-12-20*

Zhang, H., Wu, J. M., Wang, C. Y., Du, H., Liu, Z. G., Shen, L., Chen, D., & Wei, Q. W. (2016). River Temperature Variations and Potential Effects on Fish in a Typical Yangtze River Reach: Implications for Management. *Applied Ecology and Environmental Research, 14* (4), 553-567.*https://doi.org/10.15666/aeer/1404_553567*

Zhou, L., Wang, Y., & Gui, J. F. (2000). Genetic evidence for gonochoristic reproduction in gynogenetic silver crucian carp (*Carassius auratus gibelio* bloch) as revealed by RAPD assays.*Journal of Molecular Evolution, 51* (5), 498-506.*https://doi.org/10.1007/s002390010113*

Tables and Figures

Tables

Table 1 Information of sampling sites and ploidy distribution

Location	Abbr.	River system	Range type	Elevation (m)	Number of samples	Number of s
					Diploid	Triploid
Chabalang wetland	CBL	Yarlung Zangbo River	Invasive	3591	2	22
Lala wetland	LL	Yarlung Zangbo River	Invasive	3648	29	9
Ningxia	NX	Yellow River	Native (source)	1131	0	15
Sichuan	SC	Upper Yangtze River	Native (source)	519	1	11
Shishou	SS	Mid Yangtze River	Native (source)	35	6	1
Zhangdu lake	ZDL	Mid Yangtze River	Native (source)	16	8	1
Chao lake	CL	Lower Yangtze River	Native (source)	11	1	2
Tai lake	TH	Lower Yangtze River	Native (source)	10	4	0
Hongze lake	HZL	Huai River	Native (source)	10	6	0
Ru River	RR	Huai River	Native (source)	86	4	2
Daqing	DQ	Songhua River	Native	141	5	0
Jingpo lake	JPL	Songhua River	Native	349	3	1
You River	WR	Pearl River	Native	117	0	2
Chen et al., 2020	-	-	Native	-	4	12
Overall	-	-	-	-	73	78

Table 2 Pairwise genetic differentiation (F_{ST}) among invasive and native populations or genetic clusters

Diploid	Diploid	Diploid	Diploid	Triploid	Triploid	Triploid	Triploid
Population/Genetic cluster	LL	LYR & HR	Mid Yangzte River	Population	CBL	LL	NX
LL				CBL			
LYR & HR	0.0926**			LL	0.0148		
Mid Yangzte River	0.1045**	0.0449**		NX	0.0076	0.0073	
Songhua River	0.1193**	0.0348**	0.0789**	SC	0.0177	0.0281	0.0345**

LL, Lalu wetland; LYR & HR, lower Yangzte River and Huai River; CBL, Chabalang wetland; NX, Ningxia; SC, Sichuan; **, P value < 0.01.

Table 3 Genome-wide genetic diversity of invasive and native populations

Population/Cluster/Range	Population/Cluster/Range	Range type	Number of samples	Number of mitogenome h
Diploid	LL	Invasive	29	10
	LYR & HR	Native (source)	15	15
	Mid Yangzte River	Native (source)	14	14
	Songhua River	Native	9	9
	Invasive range	-	31	11
	Source range	-	30	30
	Native range	-	42	42
	Overall	-	73	53
Triploid	LL	Invasive	9	7
	CBL	Invasive	22	13
	NX	Native (source)	11	9
	SC	Native (source)	15	11
	Invasive range	-	31	17
	Source range	-	32	23
	Native range	-	47	34
	Overall	-	78	42

LL, Lalu wetland; LYR & HR, lower Yangzte River and Huai River; CBL, Chabalang wetland; NX, Ningxia; SC, Sichuan.

Figure legends

Figure 1 Sampling sites and introduction pathways of *Carassius auratus* in this study. The figures in brackets below site names represents the number of diploid and triploid samples. The two dotted lines represents main transportation routes (Qinghai-Tibet Highway and Sichuan-Tibet Highway) connecting Tibet and its east.

Figure 2 The maximum-likelihood phylogenetic trees based on mitogenome sequences of diploids (a) and triploids (b).

Figure 3 Principal component analysis based on genomic LD-pruned SNPs for diploids (a), triploids (b) and all samples (c).

Figure 4 Population structure analysis based on genomic LD-pruned SNPs for diploids (a) and triploids (b).

Figure 5 Genome-wide genetic diversity of invasive and native *Carassius auratus* at individual (a) and population (b) levels. The p value of the paired-sample t -test in figure a was showed for each pair-wise comparison. The light red and light blue bars figure b represented the invasive and native populations, respectively.

Figure 6 The genome-wide distribution of ROHs detected in invasive diploids (above) triploids (below).

Figure 7 Reconstruction of the population size histories with the SMC++ method.

Figure 8 Detection of putative signatures of selective sweeps between invasive and source populations by sliding-window analysis of the F_{ST} and |XP-EHH| values in diploids (a and b) and triploids (c and d). Dashed lines indicated genome-wide 1% outlier cut-off.

Supplemental Tables & Figures

Table S1 The detailed information of each sample of goldfish (*Carassius auratus*), including summary for genome resequencing data, individual genetic diversity, ROH detection and mitogenome

Table S2 The putative selective sweep regions identi ed between invasive and source populations

Table S3 Candidate genes obtained by annotation of SNPs within selective sweep regions

Table S4 The top 20 of enriched GO terms for candidate genes under selection in gold sh(Carassius auratus) after invasion to the Tibet Plateau

Figure S1 The analysis of allele frequency distributions at biallelic variants clearly showed one Gaussian with 0.5 in diploids (a) and two Gaussians with 0.33 and 0.67 (b) in triploids.

Figure S2 The maximum-likelihood phylogenetic tree based on mitogenome sequences of all samples. The circles and triangles represented the diploid and triploid samples, respectively.

Figure S3 Neighbor-joining clustering of diploids (a) and triploids (b) based on pairwise genetic distances calculated from genomic LD-pruned SNPs.