PNAS
www.pnas.org

Supplementary Information for

The evolutionary origin and domestication history of goldfish (*Carassius auratus*)

Duo Chen, Qing Zhang, Weiqi Tang, Zhen Huang, Gang Wang, Yongjun Wang, Jiaxian Shi, Huimin Xu[b], Lianyu Lin[b], Zhen Li[b], Wenchao Chi, Likun Huang, Jing Xia, Xingtan Zhang, Lin Guo, Yuanyuan Wang, Panpan Ma, Juan Tang, Gang Zhou, Min Liu, Fuyan Liu, Xiuting Hua, Baiyu Wang, Qiaochu Shen, Qing Jiang, Jingxian Lin, Xuequn Chen, Hongbo Wang, Meijie Dou, Lei Liu, Haoran Pan, Yiying Qi, Bin Wu, Jingping Fang,Yitao Zhou, Wan Cen, Wenjin He, Qiujin Zhang, Ting Xue, Gang Lin, Wenchun Zhang, Zhongjian Liu, Liming Qu, Aiming Wang, Qichang Ye, Jianming Chen, Yanding Zhang, Ray Ming, Marc Van Montagu, Haibao Tang, Yves Van de Peer, Youqiang Chen, and Jisen Zhang

Zhen Huang, Marc Van Montagu, Haibao Tang, Yves Van de Peer, Youqiang Chen, and Jisen Zhang
E-mail: zhuang@fjnu.edu.cn; marc.vanmontagu@ugent.be; tanghaibao@gmail.com; yves.vandepeer@psb.ugent.be; yqchen@fjnu.edu.cn; zjisen@fafu.edu.cn

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Datasets S1 to S12

SUPPLEMENTARY TEXT

History of goldfish domestication

A recent study showed that common carp (*Cyprinus carpio*) aquaculture in Neolithic China dates back 8,000 years (1) because fish and rice produce protein and carbohydrates, respectively, to provide an overall balance in the ancient Chinese diet. Similar to common carp, *crucian carp* (*Carassius auratus*) belongs to the family *Cyprinidae* and is one of the most important farmed fish with a global aquaculture production of 2.91 million tons in 2015 (2). The fertile region of China has been called “the hometown of rice and fish” (Chinese: 鱼米之乡), with the fish mostly indicating common carp, crucian carp and grass carp. Therefore, crucian carp has long been a popular food fish in East China and has been taken from the wild as food for several millennia.

Goldfish (*Carassius auratus*) was domesticated in ancient China from crucian carp. According to ancient Chinese literature, goldfish with red or xanthic scales were reported in the Jin dynasty period (AD 265-420) in a lake near Mount Lu (Jiangxi Province, southwestern China) (3, 4). These records suggested that the mutation of xanthic scales was among the first steps of artificial selection for goldfish. Because goldfish have strong adaptability, the aquaculture system has already been well established for populations of goldfish. The natural mutant of fish showing color variation would be retained as ornamental fish, which would eventually evolve into the goldfish we know today, while the wild-type fish could still serve as food.

It is likely that the strength and efficiency of artificial selection were highly limited before the Tang Dynasty (AD 618-906). Buddhism was disseminated to China from India approximately 100 AD, and one of its more important tenets is respect for all forms of life. The Tang dynasty is the Golden Age of Chinese culture and civilization with prosperous Buddhism. Goldfish aquaculture in the 'ponds of mercy' at numerous Buddhist temples and “fish sanctuaries” were popular during that period (3). Therefore, the goldfish research community assumed that preliminary phases of goldfish domestication and systematic breeding efforts started during the Tang Dynasty.

In the Song Dynasty (AD 960-1279), the domestication of goldfish was considered an important aspect of the life of royal and noble families and rapidly ascended to popularity. The Southern Song Emperor Zhao Gao (AD 1107-1187) was a goldfish enthusiast and ordered the collection of gold and silver fish for the imperial palace in Hangzhou (Zhejiang Province). The favor of imperial family had a strong influence on the noble class, and the goldfish became the "Royal fish." The artificial selection and breeding for goldfish to meet fashionable demands of noble class became highly profitable.

The Ming dynasty was also a booming period of time for goldfish breeding due to technical development and innovation for ceramic manufacturing. The production of ceramics of different sizes makes the control of breeding easier. The goldfish were found

to be painted in ceramic Vase during the Jiajing period (1521–1567) of the Ming dynasty (https://en.wikipedia.org/wiki/Chinese_ceramics). The famous new variants, eggfish (or 'dorsal-less' goldfish), were depicted in Chinese paintings dated 1429 (5), and the twin-tail goldfish were first documented in 1,596 (6). It is likely that the introduction of goldfish to Japan from China via Korea occurred in 1502 or 1620 (7), and the introduction to Europe via Macao by the Portuguese occurred in the 17th century (8).

The period from the beginning of the Qing (Ch'ing) dynasty (1644 -1912) to 1954 was called "The Era of Conscious Artificial Selection" by Chen (6). The famous Sauvigny scroll in 1772 was painted by Chinese artists, and 92 goldfish variants, including telescope eye and tumbler goldfish, were represented (9). After the Opium War, Chen recorded approximately ten new varieties, including the hooded goldfish lionhead, tigerhead and goosehead, from 1848-1925 (3).

During the twentieth century, the breeding of goldfish has attracted little attention in China. However, the accumulation of genetic mutations in goldfish served as an enlightenment for modern genetic studies in China by Shisan Chen and Japan by Yoshiichi Matsui dating back to the beginning of the 1900s. In these two decades, goldfish has revived popularity following the fast-growing economy in China. In southern China, Fuzhou city has declared to become the "hometown of goldfish" in China. At least two hundred goldfish breeding farms were located in the outskirts of Fuzhou, which produced millions of goldfish that are sold and delivered to all over the world.

Goldfish varieties and key morphological traits

Based on the phenotype, goldfish are classified into four main types according to the Chinese tradition. 1) Common goldfish, also called grass goldfish (crucian, Chinese pronunciation "Cao"-草), is the original goldfish that has no unusual anatomical features, which include the common goldfish, Shubunkin and comet goldfish. 2) Wen goldfish has a fancy tail with dorsal fin, the classic type contained Fantails and Veiltails ("Wen"-文 is also the name of the characteristic head growth on such strains as oranda and lionhead). 3) Dragon eye goldfish. These goldfish typically display extended eyes, e.g., black moor, bubble eye, and telescope eye. 4) Egg goldfish have no dorsal fin with an 'egg-shaped' body, e.g., lionhead.

In this study, we collected 185 goldfish varieties covering the major phenotypes of modern goldfish varieties from all over China, including 9 common goldfish, 87 Wen goldfish and 89 egg goldfish. Eighty-seven Wen goldfish included 53 Chinese-ranchu, 23 eggfish-lionhead and 13 egg-pompon, and 89 egg goldfish contained 23 telescope, 16 Wen-pompon and 46 Wenyu-Oranda.

The goldfish mainly has three color components: melanin, xanthic pigments and guanine. Melanin produces darker components of color, including black, brown (chocolate), and gray; xanthic pigment cells generate red, orange, and yellow shades; and

the absence of guanine in the other two pigment cell types results in a white or silvery color. We scrutinized the coloration of 185 goldfish: 32 red/yellow, 24 red/white, 9 white, 19 black, 12 black/red, 13 chocolates, 2 blue/chocolate, 6 blue, 1 blue/white, and 2 lilacs. Based on the observation of body conformation of the 185 stains, 9 were normal long-bodied goldfishes, and 176 were short-bodied and egg-shaped goldfishes. In addition, 43 of them were observed in the "dragon" eye, including 31 telescope, 6 celestial, and 6 bubble-eye, and 102 of them had a characteristic hood in the head region. More details of the distribution of morphological traits of the collected samples are provided in the *SI Appendix* (Table S13 and S14; Dataset S2).

Collection of wild crucian carp

Jiangxi Province was first recorded as the original region of the red crucian carp, Zhejiang Province was the location of the political center of the Southern Song Dynasty, Jiangsu Province was the prosperous area in China since the Tang Dynasty, and Fujian Province was considered to be the hometown for modern goldfish breeding. Therefore, we collected 16 wild crucian carps from the enclosed waters of the potential area of goldfish origin. Of these samples, two were collected from each of the provinces Zhejiang, Jiangxi, and Henan, the likely area of the goldfish origin, one each from of the provinces Jiangshu, Shangdong, and Heilongjiang, and seven from the Fujian Province. More details of the geographical distribution of the collected samples are provided in *SI Appendix (*Fig. S26, Table S13, Datasets S1 and S2).

Characterization of the subgenomes in goldfish

In a recent study, phylogeny based on mitochondrial and nuclear DNA sequences revealed that the Barbinae species could be divided into two subgroups and were the closest diploid of the goldfish (10). To identify the potential diploid progenitors of *C. auratus*, we performed whole genome sequencing for six representative Barbinae species, including *P. semifasciolatus*, *H. vernayi*, *M. arginatus*, *H. macrolepidota*, *B. schwanenfdi* and *B. melanopterus* (*SI Appendix*, Fig. S9 and S10). A total of 78.0~105.4 million reads from the six species were aligned to the goldfish genome, and 40% on average were mapped to the goldfish genomes. Of the mapped reads, 70.11%~84.52% were mapped to one set of homoeologous chromosomes, which were defined as subgenome A, and another set of homoeologous chromosomes was defined as subgenome B. We also aligned the Illumina short reads from two relative species, *D. rerio* and *C. idella*, and the results showed that the two subgenomes presented similar mapping patterns, i.e., the Illumina short reads. with no preference towards either subgenome A or B (*SI Appendix*, Fig. S10, S17, and S18). These results supported that the Barbinae species is likely the diploid progenitor of subgenome A.

RepeatMasker (11) output was used to analyze the distribution differences of each repeat family on the two subgenomes. We therefore defined that a repeat family was enriched in a subgenome by the criterion of A/(A+B)-B/(A+B)≥0.8 or B/(A+B)-A/(A+B)≥0.8, resulting in one and six A-specific and B-specific repeat families, respectively, which are the hAT-Ac and TcMar-Tc1 elements, respectively. hAT-Ac was significantly more abundant on the A than on the B subgenome (Fig. S11). To predict the activity time of the A-specific repeat family, subfamilies were computed manually as follows. MAFFT (12) was used to align all sequences in the A-specific family, whereafter phylogenies were constructed with FastTree. In total, three A-specific subfamilies with at least 10 copies of 150 bp or longer were identified (Fig. S12). We aligned the filtered subfamily sequences with MAFTT (12), obtained conservative sequences based on these sequences, and calculated the distance from each sequence. The distances were then adjusted for multiple substitutions according to the Jukes-Cantor formula and a density plot was plotted for the A-specific repeat families (Fig. S13). The burst of the A- and B-specific TE activities occurred independently in the progenitors during that period.

The gene cluster analysis of the two subgenomes showed that 11,876 gene clusters (accounting for 31,572 genes) were shared in subgenomes A and B, and 1,810 (accounting for 3,295 genes) gene clusters were specific in subgenome A and 2,240 (accounting for 1,354 genes) in subgenome B (*SI Appendix*, Fig. S16). The duplicate retention rates are ~89% and ~83% for subgenomes A and B, respectively. To characterize the expression features of *C. auratus*, we performed RNA-seq profiling for ten tissues, including spleen, scale, muscle, mid-kidney, head kidney, gut, gill, dorsal fin, atrium and testis. We carried out comparative analysis of the genome-wide transcriptional levels for subgenomes A and B based on the gene expression levels of 16,362 homoeologous gene pairs in these 10 tissues to investigate the homoeologous gene expression patterns (*SI Appendix*, Fig. S19; Dataset S10). To minimize noise from low-expression genes, we only retained the homoeologous gene pairs that have at least one gene with expression of FPKM>=10 **(**Fig. 1c; *SI Appendix*, Fig. S20). We defined homoeologous gene pairs with greater than twofold change within a pair with the higher-expressed gene as the dominant genes and others as recessive genes. The results showed that 7,951 (48.6%) were dominant genes in at least one of these ten tissues. Of the 7,951 homoeologous gene pairs, 5,010 presented dominant expression in subgenome A, and 3,828 were dominant expressed in subgenome B (Fig. 1*c*); 887 homoeologs were swinging dominant bias either in subgenome A or subgenome B. Further analysis revealed that 4,123 and 2,941 homoeologs were exclusively dominant in subgenomes A and B in the 10 tissues, respectively, and 170 and 121 homoeologs were dominantly expressed in subgenomes A and B in all 10 tissues (Fig. 1*c*; *SI Appendix*, Fig. S19 and S20), respectively.

The homoeologous expression bias exhibited asymmetric expression patterns between subgenomes A and B, and the genome-wide expression level dominance was

biased towards subgenome A in the allotetraploid *C. auratus* genome. GO term enrichment showed that the dominantly expressed homoeologs of subgenome A were primarily enriched in the "oxidation-reduction process", "acid-amino ligase activity", "actin binding", "glycolytic process" and "oxidoreductase activity" terms, while those of subgenome B were enriched in the "Protein folding", "Oxidation-reduction process" and "Cellular response to xenobiotic stimulus" terms (*SI Appendix*, Fig. S21). KEGG pathway analysis indicated that dominantly expressed homoeologs of subgenome A mainly participate in "Insulin signaling pathway", "Leukocyte transendothelial migration", "Neurotrophin signaling pathway" and "HIF-1 signaling pathway", while the homoeologs dominantly expressed in subgenome B primarily participate in "Tight junction", "Chemokine signaling pathway" and "Complement and coagulation cascades" pathways (*SI Appendix*, Fig. S22 and Dataset S12). These results suggested the functional divergence between the two subgenomes in the goldfish.

Sex chromosome

In *Cyprinidae*, the mechanism of sex determination has not been clearly studied. The high-density meiotic map SATmap was used to identify the sex determination region. Most of chromosome 16 carried a strong signal with a broad peak around the centromere, but no specific region can be identified as a sex-determination region (13). Coincidentally, the identity of the sex determination chromosome(s) in grass carp is still unclear, although the genetic map was used (14). In zebrafish, the long arm of zChr4 was suggested to be the sex-associated region in zebrafish zChr4 (13, 15), and very recently, a relatively large nonrecombining region (~11.7 Mb) on the goldfish linkage group 22 (LG22) that contained a high density of male-biased genetic polymorphisms was also detected on ChrA22 (~19.9 Mb) (16). We found that *Amh* considered a putative male sex-determining gene occurred also in ChrA22 (Cau.A22G0011390) of goldfish based on the homology analysis. In our study, we used a female fish (XX) for the whole-genome sequencing and assembly. The analysis of the sex chromosomes was preliminary.

Goldfish were suggested to have a sex determination system with male heterogametic (XX/XY) (17). In addition to genetic factors, the sex determination system was also controlled by the environmental factors in goldfish, for example, female-to-male sex reversal occurred under the high temperatures (17). In this study, in order to identify the sex-determination candidate region, we resequenced 3 male and 3 female goldfish individuals. In theory, if the goldfish have mature sex chromosomes, using the reference male genome that we generated, the whole genome shotgun (WGS) reads of males are supposed to have higher depth of coverage (approximately twice) for the X sex-chromosome (regions) than the reads from females (X/Y). A total 323.56 million reads from the male fish and 403.60 million reads from female fish were respectively mapped the reads to the reference genome using Bowtie2 (18) with default parameters, respectively. We scrutinized the potential sex-related genomic regions in the goldfish genome which

have reported in zebrafish, grass fish and goldfish (13, 14, 19). However, we did not observe any chromosome with expected sex chromosomal patterns (Fig. S44). In the future, the *de novo* assembly of the female genome combined with the genetic map and the male/female goldfish resequencing could provide further evidence for determining the sex chromosome in goldfish.

Remarkably, the long arm of the homeologous chromosome pair, ChrA4/ChrB04 (20.49~32.10 Mb in ChrA04, 15.44~41.40 Mb in ChrB04), was found to lack synteny both between the chromosome pair (*SI Appendix*, Fig. S15, S45, and S46), as well as showing only slight collinearity with the presumed homologous zebrafish chromosome 4 (Fig. S47). Despite the slight collinearity between the long arms of ChrA04/ChrB04 in goldfish and zChr04, these chromosomes in zebrafish and goldfish are very similar to each other in terms of genomic content including NLRP genes, TEs (13, 15), indicating this highly repetitive chromosomal regions originated before the divergence of zebrafish and Cyprininae subfamily. Since the long arm of chromosome 4 in zebrafish was suggested to be a sex chromosome, ChrA04/ChrB04 are thus presumed to be the sex chromosomes in goldfish due to this implied orthology, despite a general lack of collinearity. Of note, as the allotetrapolyploid species, the pairs of homoeologs potential sex chromosomes indicated a different genetic system from the diploid species for sex determination and segregation in goldfish. However, it is possible that either ChrA04 or ChrB04 is the sex chromosome, and thus goldfish could have similar genetic system as diploid.

In these regions, 202 and 404 genes were annotated in ChrA04 and ChrB04, respectively, of which 38 (18.8%) and 184 (45.5%) genes in ChrA04 and ChrB04 were duplicated genes, respectively. Among those genes, a total of 10 and 44 are duplicated NLRP genes, respectively, which have been considered to have putative functions in innate immunity (20), indicating disease resistance associated with these genomic regions. Furthermore, high levels of TEs (52.1% in ChrA04 and 50.2% in ChrB04, compared with the genome average of 37.22%) exist in these regions with a majority of Class II type, including a large number of DNA/hAT and DNA/CMC-EnSpm elements (Fig. 2; *SI Appendix*, Fig. S45). Class II type transposons are typically associated with heterochromatin, with mutations rapidly accumulating in these regions leading to noncollinearity to closely related chromosomes (Fig. S47). However, the enrichment of genes responsible for innate immunity suggests that this rapidly evolving region nonetheless carries important biological functions (Fig. S48 and S49). Since immunity-related genes are often under strong positive selection in various taxa (21), for example, due to the ‘arms-race’ with infectious pathogens, their location in a rapidly evolving chromosomal segment may not be coincidental and might be considered a selective advantage. We hypothesize that the highly divergent chromosome segments on ChrA04 and ChrB04 may serve an important role that might have facilitated the adaptation of Cyprinidae species to different natural habitats that often contain a variety of pathogens.

Evolutionary history of goldfish

The availability of the reference genomes for Cyprinus carpio (22), grass carp (23), and zebrafish (13), as well as the current genome, enabled us to reconstruct the history of chromosomal evolution of goldfish within the Cyprinidae family. Based on Ks value estimates, zebrafish diverged from the last common ancestor (LCA) of *Cyprinus carpio*, *Carassius carassius*, and *Ctenopharyngodon idellus* ~32.8–36.8 MYA, and segmental inversions in ChrA04\ChrA14\ChrA19 occurred after the divergence of zebrafish and the lineage leading to the subfamily Cyprininae, as the orientation of this chromosomal fragment was conserved among the extant *Cyprininae* genomes examined to date (Fig. 2c and d; *SI Appendix,* Fig. S15, Fig. S23, S24). The grass carp lineage split off from the LCA of *Cyprinus carpio* and crucian carp ~21.1-23.7 MYA followed by the fusion of corresponding ancestral zebrafish Chr10 and Chr22 to Chr24 (*SI Appendix,* Fig. S24 and S25), resulting in a basic chromosome number of n = 24 (2).

Goldfish had originally been classified as common carp (*Cyprinus auratus*) by Linnaeus in 1758 and, later, reclassified as *Carassius auratus*, considering that the phenotype was closer to crucian carp than to common carp (24). Based on these results, we can infer that goldfish originated more probably from crucian carp than from common carp. After exclusion of common carp, phylogenetic analysis of 201 individual crucian carp and goldfish samples by means of a set of 4,842,951 biallelic single-nucleotide polymorphisms (SNPs) revealed that all the wild crucian carp samples clustered together and were clearly separated from goldfish (Fig. 2*a-c*; *SI Appendix,* Fig. S29 and S30). The principal component analysis (PCA) of the total SNP variation also confirmed the results; because both of the largest principal components (PC)-PC1 (11.29%) and PC2 (7.87%) distinguished crucian carp from goldfish (Fig. 2*a*).

Long noncoding transcripts

FEELnc tool v0.1.1 was employed for classifying non-coding RNAs with the default parameters, which compares the genomic location and transcript orientation with respect to the protein-coding RNA (25).The long noncoding transcripts were classified into 4,628 with intergenic locations and 4,713 with genic locations and further divided into multiple subtypes according to recently proposed definitions (25) (*SI Appendix,* Table S8).

Genes associated with selective sweeps in teleosts or other organisms

Several other genes associated with selective sweeps have reported functions in teleosts or other organisms, such as genes involved in embryonic development (*eif3f*) in zebrafish (26) and mouse (27); colorectal cancer in humans (*NAV3*) (28); regulation of autophagy in cone photoreceptors in zebrafish (*Arf6a*) (29); regulation of zebrafish primitive and definitive hematopoiesis (Ak2) (30); EXTL3 mutations causing skeletal dysplasia, immune deficiency, and developmental delay (31); causing T cell immune deficiency in zebrafish (*zap70*) (32), mechanotransduction in sensory hair cells (*pcdh15*) (33), and a planar cell polarity effector (*Nhsl1b*) in zebrafish facial branchiomotor neurons (34).

Gene flow of goldfish population

Gene flow from the Chinese Ranchu to common goldfish (weight = 0.187) was due to the modern Chinese Ranchu, which originated from Japanese Ranchu containing the genetic backgrounds of Japanese Wakin and Jikin in the common goldfish groups. These results demonstrated the significant contributions of Japanese goldfish genetic stocks to modern goldfish breeding due to Japanese goldfish breeding initiatives in the past century.

GWAS of eye and color variants in domesticated goldfish

Telescope eyes have been a symbolic feature of some goldfish in the western world since de Sauvigny published goldfish scrolls in 1780 in Paris (9). Based on 31 cases and 142 controls for telescope eyes, GWAS with a threshold of the top 0.1% revealed 112 genetic loci spanning 11.08 Mb in multiple genomic regions (*SI Appendix,* Fig. S50-52 and Dataset S7). In these regions, 7 genes associated with telescope eyes overlapped with genes associated with dorsal fin phenotypes, but none of them were common to the 393 domesticated candidate genes within the domestication sweeps, indicating that telescope eyes likely emerged after the original domestication of goldfish (*SI Appendix,* Fig. S53). A total of 20 goldfish genes corresponded to 17 zebrafish orthologs whose gene knockout lines display mutant phenotypes associated with eye development, supporting the reliability of these GWAS results in goldfish (Dataset S7). Telescope eyes were thought to be controlled by one or two loci, since a cross with wild goldfish (Funa) gave a segregation ratio of 15:1 (wild-type eyes: telescope eyes) (35). However, in the present study, three genes, i.e., Cau.A01G0001330, Cau.B20G0012610, and Cau.B22G0008190, underlying a strong association peak, were orthologous to the eye development-related genes *ncapg*, *rps29* and *sec13* in zebrafish (*SI Appendix,* Fig. S50). The eye variants in the current study included spherical, ovoid, a truncated cone, and a segmented eye, according to the classifications of Hervey & Hems (36). The observed phenotypes could be explained by telescope-eye phenotypes being under the control of one major genetic locus in addition to several other eye-associated genes with minor effects.

Color mutations are almost a commonplace among ornamental fish, and color represents the earliest trait to have undergone artificial selection in goldfish and has been a target of selection for over 1000 years, resulting in ‘an almost infinite diversity of color’ in goldfish (37). To explore color-associated genes in goldfish, we performed GWAS with three paired case/control datasets including blue coloration compared with red (24 to 32), red compared with partial white and white (33 to 32) and black compared with red (19 to 32), for which 598, 504, and 807 genes, respectively, were in genomic regions associated with the threshold of top 0.1% (*SI Appendix,* Fig. S54-S62 and Datasets S8-S10). At least 60 of these candidate genes have previously been related to pigmentation in zebrafish and vertebrates (Datasets S8-S10). For example, a series of pigment-related genes were located

in the strong association peaks, including an SNW domain-containing protein 1 (Cau.A20G0001720) that regulates the BMP signaling pathway controlling hair pigmentation in humans (38), the endothelin-1 receptor (Cau.B01G0008200) related to pigmentation in cats (39), a cadherin-associated protein (Cau.B01G0009880) involved in melanocyte differentiation, a point mutation in optineurin (Cau.B04G0002560) related to pigment cells in zebrafish (13), a cytochrome c1 (Cau.B06G001201) and an adenylate cyclase related to melanophore development (40-42).

Discussion

As ornamental fish, the domestication history of goldfish has been impacted by three historical lines, including the contemporary social, religious, and political environments. Goldfish is unique among domesticated animals since it was largely used as a 'luxury' item, historically petted by the nobles, riches and royal families, thus subject to shifting aesthetics and tastes over time. Consequently, goldfish breeding differed from the breeding of herd livestock, such as goats (43), sheep (44), cattle (45), and pigs (46), because these animals were under direct selection for common goals, such as yield and quality, that would occur independently in different geographical areas. Goldfish appears to have developed with extensive admixture of different genetic backgrounds, because breeders often wished to develop new lines with novel combinations of anatomical features. Numerous mutations, including the xanthic mutation, the depigmented mutation (*bifactorial mutant*), albinism, eye mutant, narial bouquet mutant, hood mutant, scale variant, and body conformation mutants, have been selected in goldfish populations and spread via admixture.

The selection of the xanthic mutation had been the initial step in goldfish domestication, resulting in a bottleneck for genetic diversity in the goldfish population and in an initial reduction in diversity in some regions through selective sweeps. Afterward, artificial selections based on mutations in large breeding populations, accompanied by hybridization, have since then maintained the morphological diversity via high degrees of polymorphism at many loci (Fig. 4 *a* and *b*).

Several morphological variants in goldfish are inherited in a Mendelian fashion, including transparent scale (47), blue/brown coloration (7), bifactorial mutants (48), albinism (48, 49), recessive transparency (48), and telescope eye mutants (48, 50). These morphological variants are in concordance with the diverse genetic backgrounds of goldfish, as shown by the transparent scale mutant in our study. Such morphological variation in, for example, body conformation and fin size, has evolved under directly strong selection that has consequently facilitated genetic divergence, as shown in the case of the dorsal fin in Wen and Egg goldfish. The diverse selection schemes had a large impact on the genetic landscape of goldfish, with changes affecting a larger number of loci than

expected for other farm animals that had primarily been used for herding and meat consumption.

SUPPLEMENTARY MATERIALS AND METHODS

Goldfish materials. A 12-generation inbred line of a female common goldfish (G-12) was collected from the Chunyuanli Ecological Farm, Minhou County Nantong, Fuzhou, Fujian, China (E119°13'46.55", N25°54'44.56") and used for whole-genome sequencing. The collection time was March 10, 2016 with the sample number GF0001.

Illumina short read sequencing. Genomic DNA was extracted from blood of goldfish using the DNeasy Blood and Tissue Kit (Qiagen) and used to construct libraries using the NEBNext Ultra DNA Library Prep Kit for 150 bp paired-end sequencing on the Illumina X Ten platform. The raw reads were trimmed using Trimmomatic (51) (v0.36) with default parameters.

Genome annotation. Tandem repeats were identified using the Tandem Repeat Finder (TRF) package (52), and unknown TEs were classified using Teclass (53). We also integrated results from LTR_FINDER (54) and LTRharvest (55) and removed false positives from the initial predictions using the LTR_retriever pipeline (56). The goldfish genome assembly was annotated by means of homology-based, transcriptome-based, and *ab initio* annotations. For the homology prediction, four representative species were selected, including *C. auratus* (17), *Cyprinus carpio* (57), *C. idellus* (55), and *D. rerio* (23). The protein sequences from these genomes were aligned to the goldfish genome assembly with TBLASTN with an E-value cutoff of 1e-05. GeneWise was used to predict the exact gene structure of the corresponding genomic regions for each BLAST hit. For the transcriptome-based prediction, RNA-Seq and ISO-Seq data were sampled from female and male goldfish. Transcriptome assembly was carried out with both *de novo* and reference-guided Trinity with RNA-Seq (58) data. A high-quality complete transcriptome was obtained through the PacBio pipeline with ISO-Seq data. *Ab initio* genes were predicted by SNAP (59), GENEMARK (60), and AUGUSTUS (61) with repeat-masked genome sequences. We integrated all genes predicted from the three annotation procedures with two rounds of MAKER (62). In the second MAKER (62) run, the predicted gene models with annotation edit distance (AED) scores of 0.25 were extracted for retraining again with SNAP (59), GENEMARK (60), and AUGUSTUS (61), where after the TE-related genes were removed. Finally, the MAKER (62) pipeline generated 56,251 gene models (Table 1). Although different annotation methods may have subtle variations, they generally do not affect the comparative genomics analysis, because in general the different annotation results should be consistent.

Identification of homoeologous genes and syntenic blocks. The conserved paralogs of the protein sequences were obtained with BLASTP (E-value ≤ 1E-5). Homoeologous gene and pairwise collinearity analyses between species were applied by means of MCScanX (http://chibba.pgml.uga.edu/mcscab2). At least five genes and ≥70% identity scores were required to identify the syntenic blocks. Alignments were used as an edge to group genes into clusters according to whether it was before or after the carp WGD event. Two genes or gene clusters were merged as described (1).

Analysis of chromosomal synteny and structure. Structural comparisons were performed using minimap2 (63) alignment tools, and visual interpretations of dot plots created using D-Genies (64). Gene density and NLRP gene distribution plots were created by RIdeogram package (65) of R program.

Discrimination of distinct subgenomes. To discriminate the subgenomes of goldfish, DNA was extracted from the recent diploid Cyprinidae species, *Puntius semifasciolatus*, *Hypsibarbus vernayi, Mystacoleucus marginatus, Hampala macrolepidota, Barbonymus schwanenfdi* and *Balantiochelos melanopterus*, for Illumina sequencing on the HiSeq™ X Ten sequencing platform. In addition, 202.9 million reads of *C. idella, D. rerio* and *C. molitorella* were download from NCBI. The Illumina reads for the six species (~12 Gbp) *C. idella*, and *D. rerio* were mapped onto the goldfish genome using Bowtie2 (18, 66), and mapping results were standardized using BEDTtools (67) and PYTHON scripts.

RNA-Seq of *Carassius auratus* tissues. Total RNA was extracted from *C. auratus* gill, atrium, head kidney, mid-kidney, testis, gut, scale, dorsal fin, muscle, and spleen using TRIzol (Invitrogen, California, USA) reagent following the manufacturer's instructions. RNase-Free DNase I (Takara, China) was used to remove genomic DNA according to the manufacturer's instructions. Approximately 120 Gb transcript data were generated for *C. auratus* using the Illumina HiSeq™ X Ten sequencing platform (Illumina, USA) and processed using the Trimmomatic (v0.36) pipeline (51). FPKM(fragments per kilobase of transcript per million mapped reads) for filtered reads (after removing low-quality reads(Q-value < 20) obtained from all samples were calculated using RESM in Trinity (Dataset S11)(68).

Analysis of genome expression dominance. The clean RNA-Seq reads after quality control were mapped onto the *C. auratus* genome using TopHat2 (69). Transcript expression levels of individual genes were quantified using FPKM values using RESM in Trinity (68). Analysis of homoeolog expression dominance was performed between syntenic gene pairs compared to that of homoeolog gene pairs (Dataset S10). To avoid the false positive results, we only selected homoeologous gene pairs having at least one gene with FPKM>=10. Differentially expressed gene pairs with expression differences greater

than twofold were defined as dominant gene pairs. Dominant genes were those with relatively higher transcript expression between dominant gene pairs, and those with lower transcript expression were considered subordinate genes. The rest of the syntenic gene pairs that showed no dominance relationship were classified as codominant genes.

Read mapping, variant calling, and filtering. NGS reads of 201 samples were processed for quality control using BBTools (https://jgi.doe.gov/data-and-tools/bbtools) with the default parameters, and the clean reads were then imported to map them against our goldfish genome assembly using the Burrows-Wheeler Aligner (BWA) MEM (70) with the parameter -M -k 30. The alignments were then sorted using SAMtools (71) (v1.9), and putative PCR-generated duplicates were removed using Samblaster (72). Variants were called for each sample using the GATK (73) (v3.8) HaplotypeCaller and subsequently joint-called across all samples using the GATK GenotypeGVCFs tool.

A subset of 5,202,428 biallelic SNPs in these 201 samples were defined using the following tools and parameters: vcfutils.pl varFilter (72) with the parameters -w 5 -W 10; GATK Variant Filtration with the parameters --filterExpression "QUAL < 30 || QD < 2.0 || FS > 60.0 || MQ < 40.0" --clusterWindowSize 5 --clusterSize 2; BCFtools View (74) with the parameters -m2 -M2; and then VCFtools (75) with the parameters --maf 0.05 --max-missing 0.8.

The marker set was phased by beagle (v5.1) (76) and annotated using SnpEff (77) (v4.3) with our gene annotation for common goldfish. The density of SNP markers was 317 bp per marker, and the Ts/Tv ratio was 1.05. Most of the markers (approximately 40%) are distributed in intergenic regions. We also performed the above procedures and used the same filtering criteria to generate an additional SNP marker set including the previously sequenced 33 *Cyprinus carpio* (78) and our 201 resequenced samples. We used this marker set to clarify the relationships among *Cyprinus carpio*, crucian carp, and goldfish.

Population structure analysis. A subset of SNP markers in only 185 goldfish were generated using PLINK to remove the 16 crucian carp samples and to filter SNPs using the parameters --maf 0.05 --geno 0.2. Both sets of the above SNP markers were used to conduct PCA and identity-by-state (IBS) calculations in PLINK (79). We used the first four eigenvectors to draw scatterplots. A genetic distance matrix of 1-IBS values was used to construct a minimum evolution phylogeny tree using the FastME program (80) that was visualized with the web service iTOL (81). Analysis of population structure revealed that these 201 samples could be partitioned into four subpopulations (CC, CG, WG, and EG). To perform TreeMix (82) analysis, the per-site allele count (AC) of each group was calculated using PLINK (79) with the --within command, and the results were then converted into TreeMix (82) input format using the plink2treemix.py script. We then ran TreeMix (82) to construct a maximum-likelihood population tree.

Selective sweep detection. Whole genome detection of selective sweeps was conducted using Fst between wild (C16) and domestic (G33) populations, the composite likelihood ratio (CLR) test, and Shannon entropy analysis within G33. The per-site Fst for C16 compared with G33 was estimated using PLINK by scanning with an average 50-kb block size across all chromosomes. The CLR test was implemented in SweepFinder2 (83) for multiple markers across a contiguous grid space. To measure genetic diversity, which could also be interpreted as the degree of chaos per site for the two subpopulations, we used custom R scripts to calculate the relative Shannon entropy of biallelic loci as denoted by S'(A) in the following formula:

$$S'(\mathrm{A}) = \frac{-[p \cdot \ln(p) + q \cdot \ln(q)]}{ln2}$$

The top 1% of CLR scores were selected as sweep regions. GO term enrichment analysis was performed for candidate genes located within these sweep regions. TopGO (http://www.bioconductor.org/packages/release/bioc/html/topGO.html) was used for GO term enrichment analysis, the REVIGO (84) tool was used to summarize nonredundant GO terms, ClusterProfiler (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) was used to perform KEGG pathway enrichment analysis, and Cytoscape (85) was used to visualize networks of terms based on Jaccard similarity coefficients. Extended haplotype homozygosity (EHH) analysis was performed by the R package rehh (86) around the selected SNPs of representative candidate genes.

Population parameter estimation. Genetic diversity (π) was calculated with the parameters --window-pi 50000 --window-pi-step 50000. Tajima's D was estimated with the parameters --TajimaD 50000. Both π and Tajima's D also used the parameter --keep to choose samples from each group. Fst was calculated with the settings --fst-window-size 50000 --fst-window-step 50000 and the parameter --weir-fst-pop to set groups.

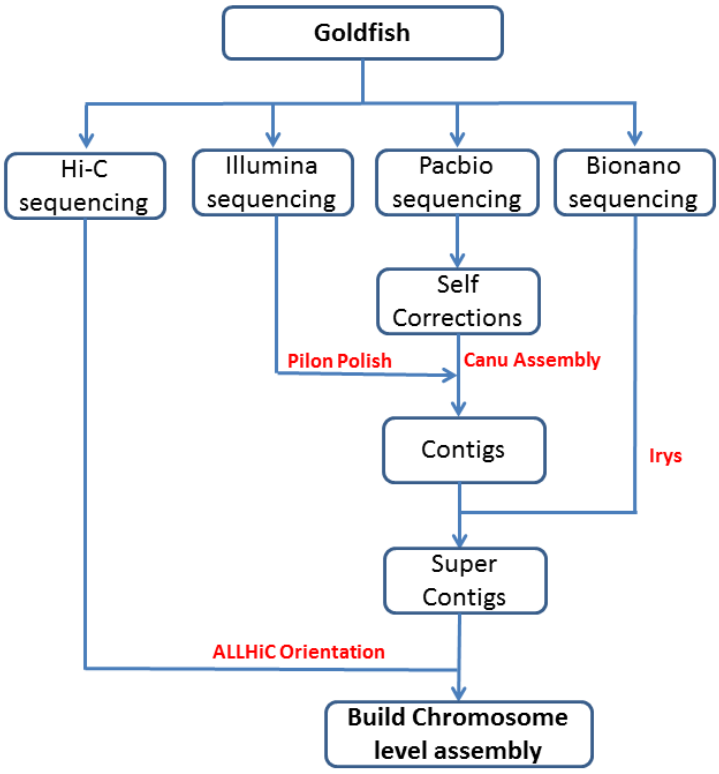


Fig. S1. Assembly strategy of goldfish genome.

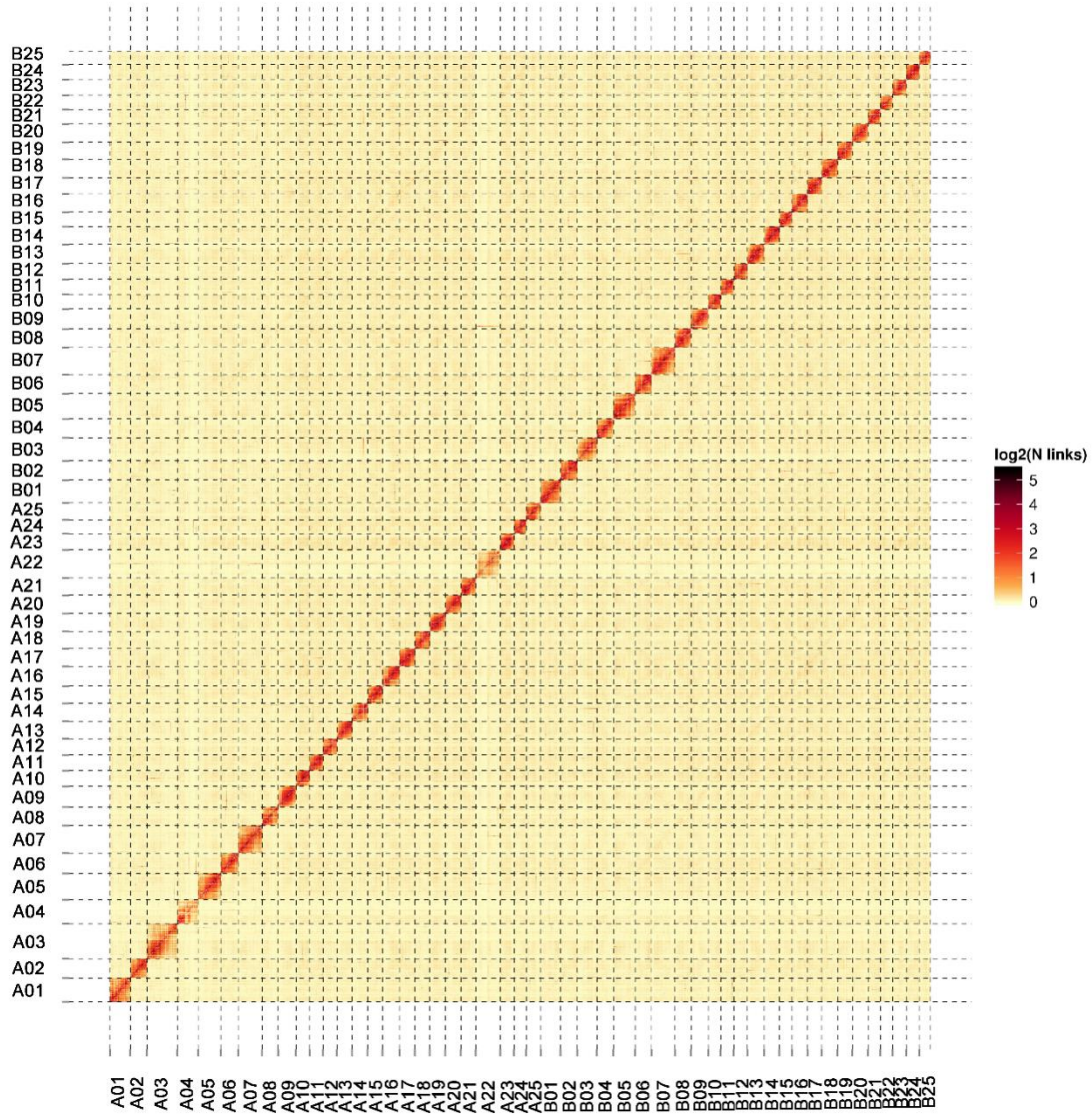
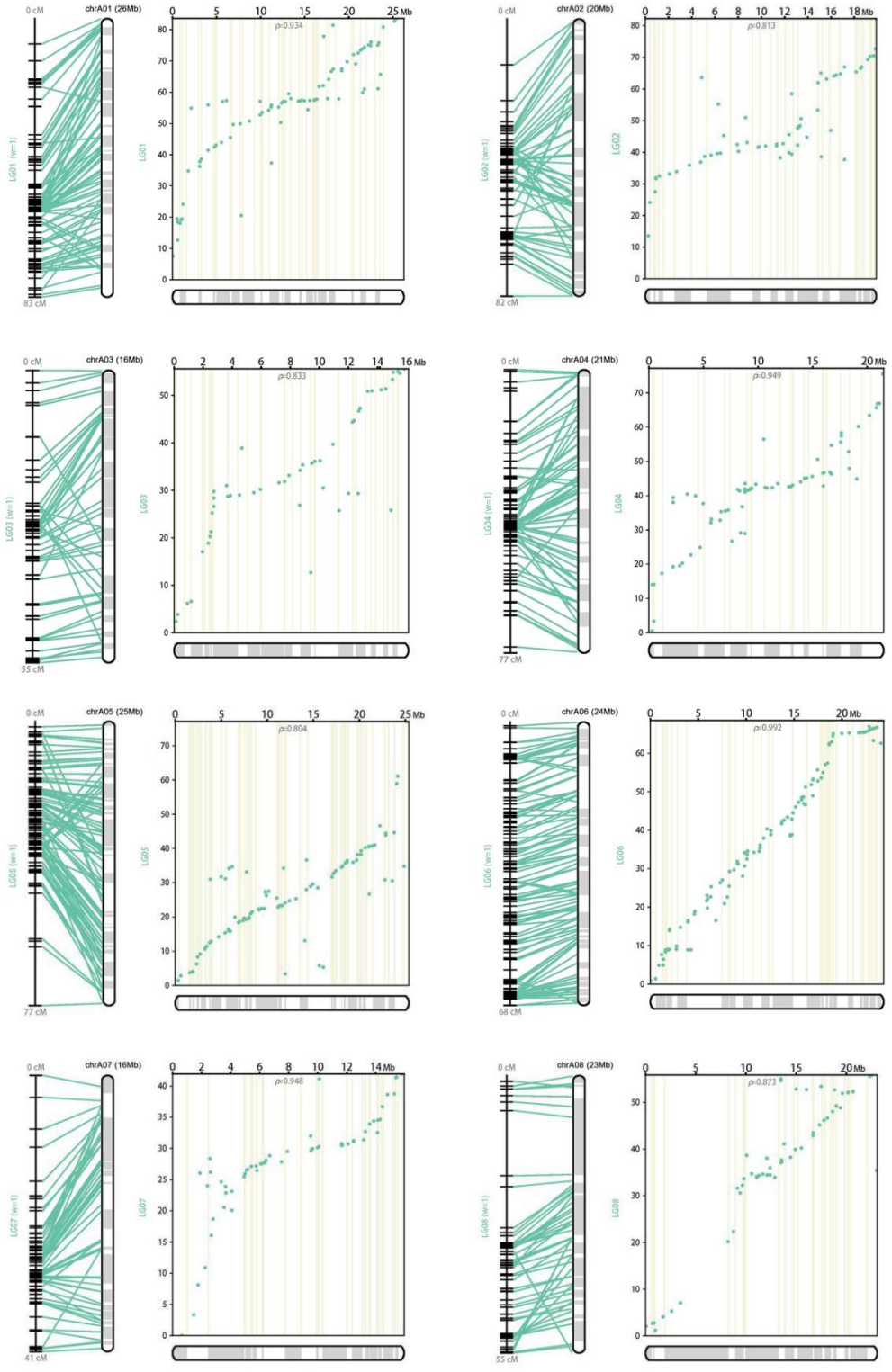
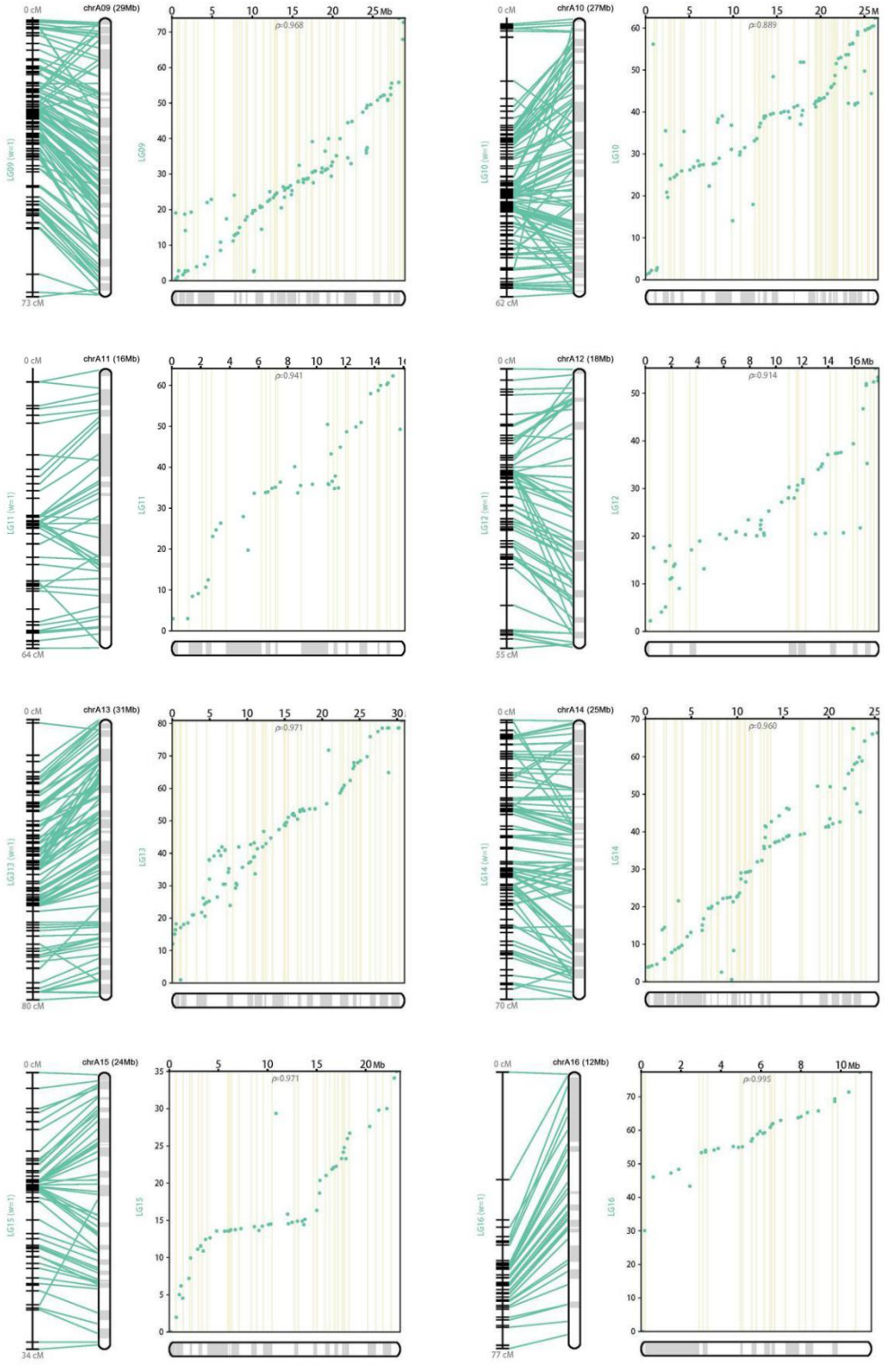
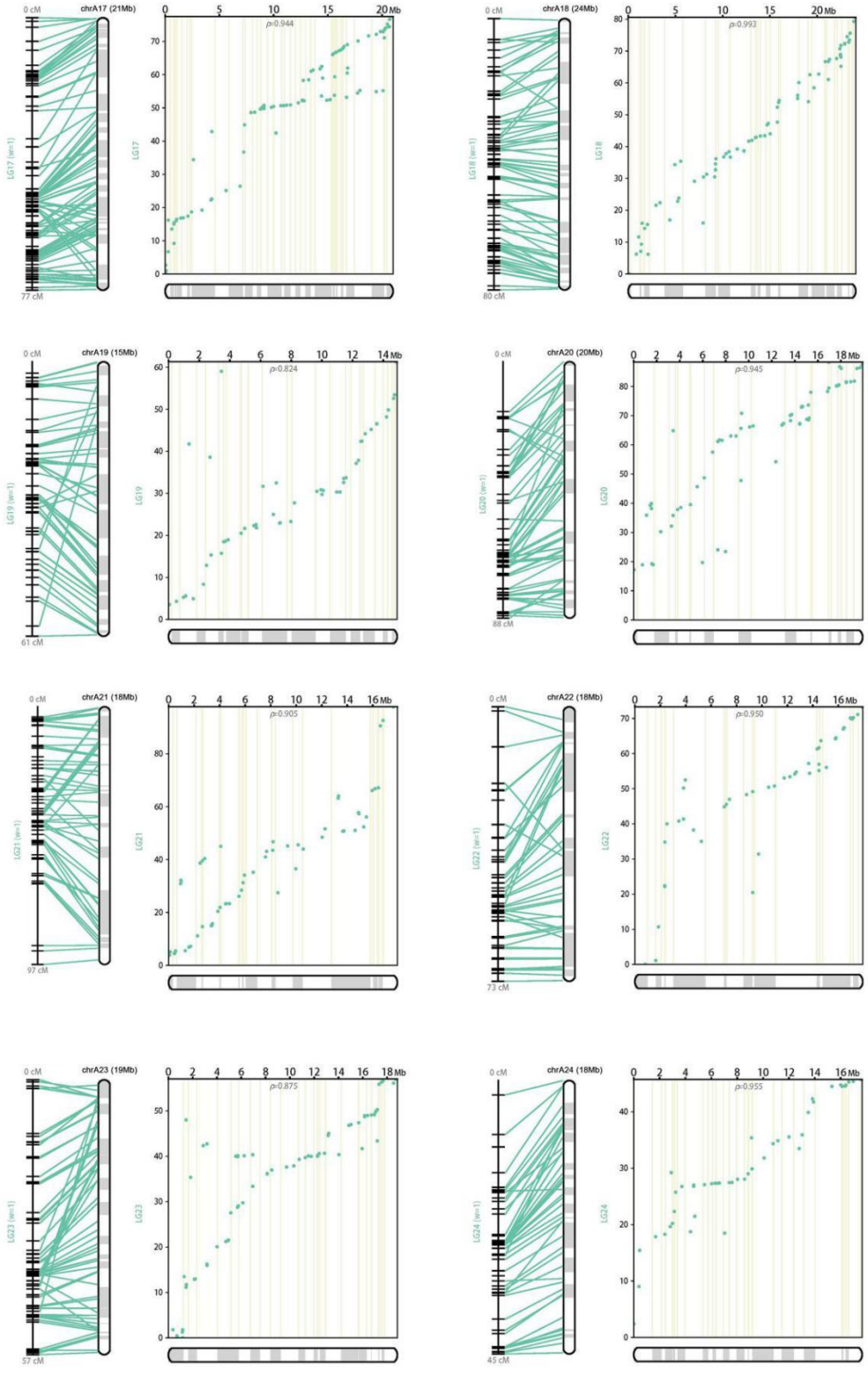
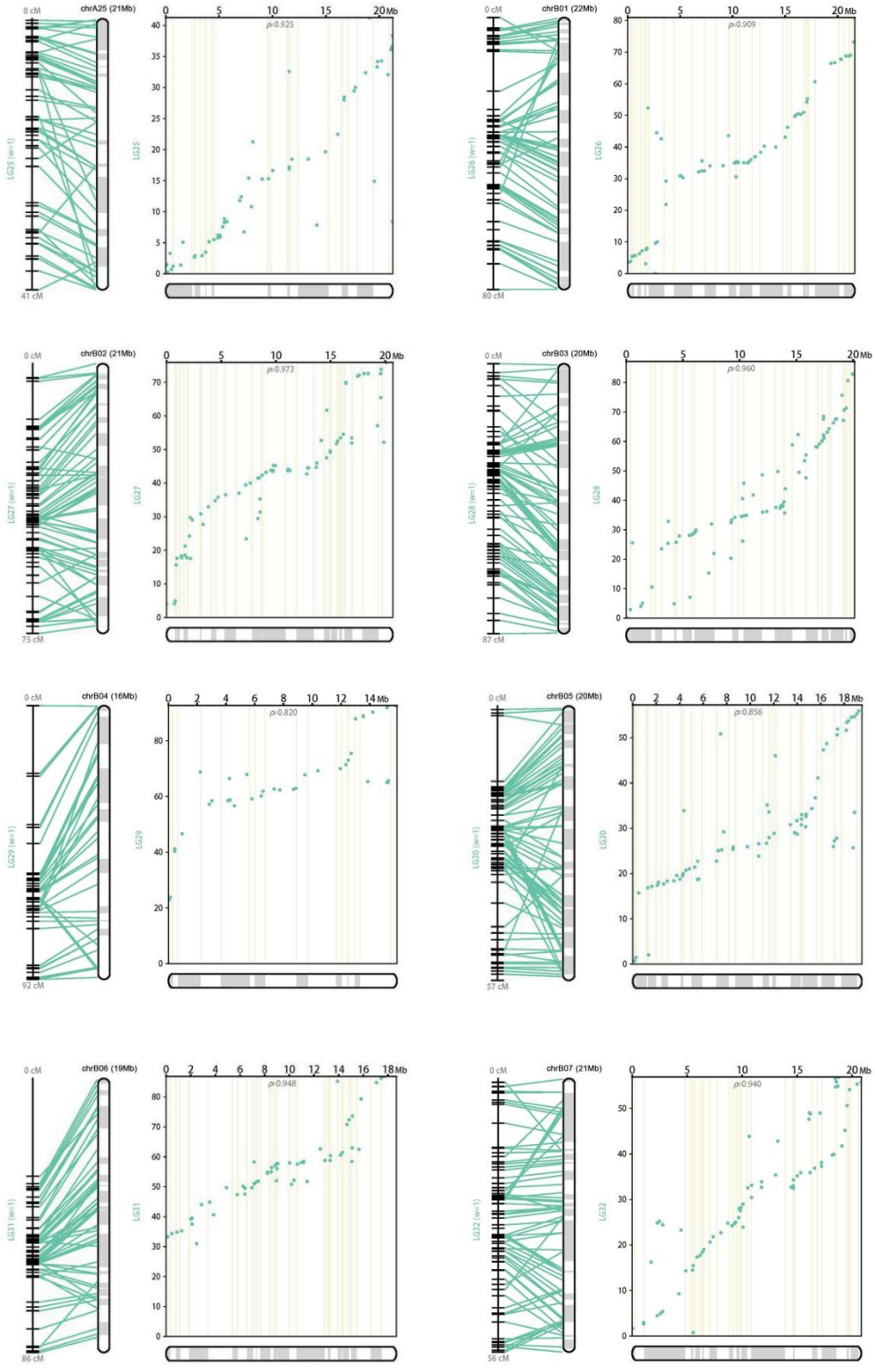


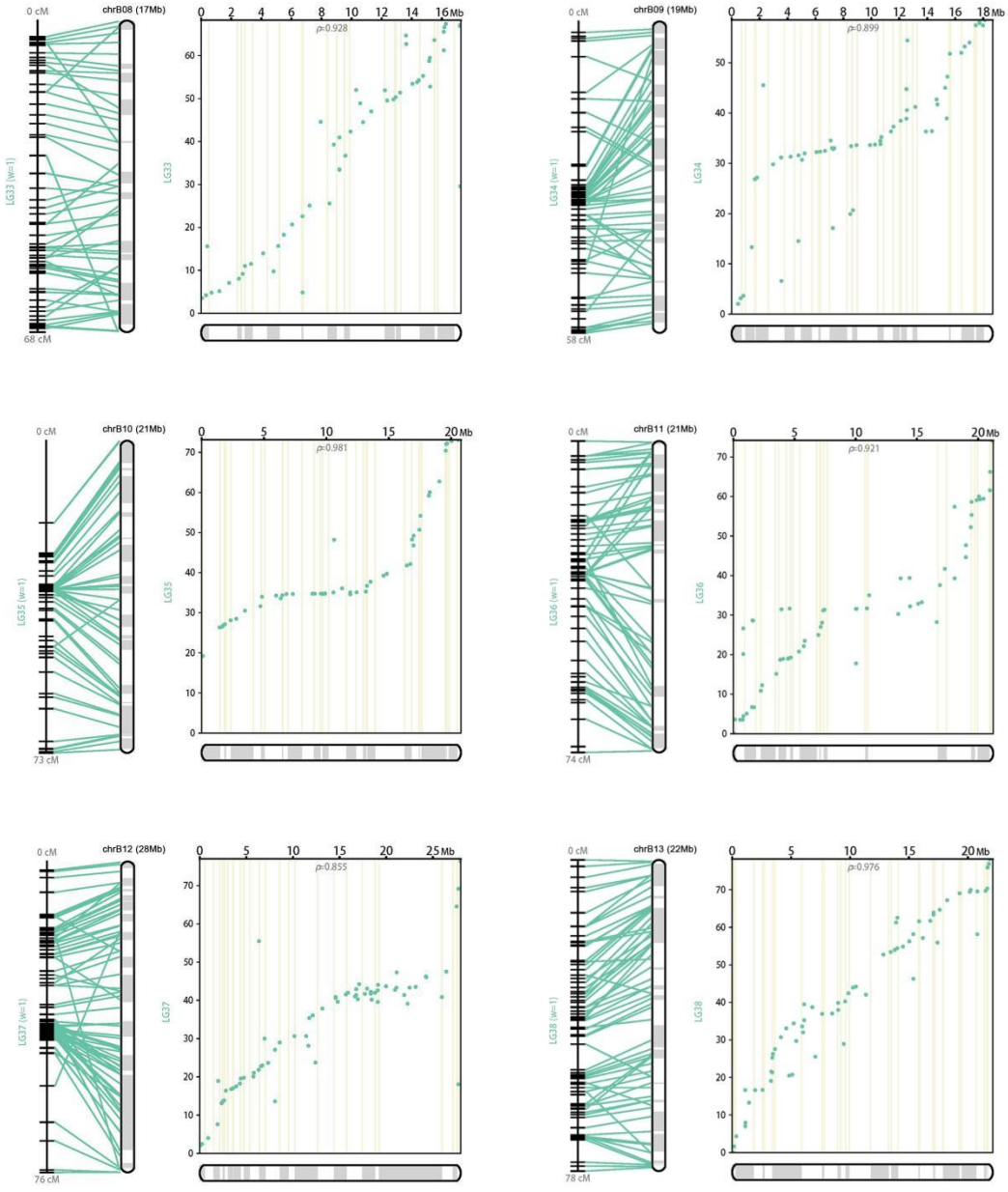
Fig. S2. Hic contact map from the goldfish. Genome-wide analysis of chromatin interactions in *C. auratus*.


chrA01 (26Mb)
LG01 (w=1)
0 cM
83 cM
LG01
ρ=0.934
chrA02 (20Mb)
LG02 (w=1)
82 cM
LG02
ρ=0.813
chrA03 (16Mb)
LG03 (w=1)
55 cM
LG03
ρ=0.833
chrA04 (21Mb)
LG04 (w=1)
77 cM
LG04
ρ=0.949
chrA05 (25Mb)
LG05 (w=1)
77 cM
LG05
ρ=0.804
chrA06 (24Mb)
LG06 (w=1)
68 cM
LG06
ρ=0.992
chrA07 (16Mb)
LG07 (w=1)
41 cM
LG07
ρ=0.948
chrA08 (23Mb)
LG08 (w=1)
55 cM
LG08
ρ=0.873


chrA09 (29Mb)
LG09
chrA10 (27Mb)
LG10
chrA11 (16Mb)
LG11
chrA12 (18Mb)
LG12
chrA13 (31Mb)
LG13
chrA14 (25Mb)
LG14
chrA15 (24Mb)
LG15
chrA16 (12Mb)
LG16


chrA17 (21Mb)
p=0.944
chrA18 (24Mb)
p=0.993
chrA19 (15Mb)
p=0.824
chrA20 (20Mb)
p=0.945
chrA21 (18Mb)
p=0.905
chrA22 (18Mb)
p=0.950
chrA23 (19Mb)
p=0.875
chrA24 (18Mb)
p=0.955


chrA25 (21Mb)
LG25 (w=1)
LG25
ρ=0.925
chrB01 (22Mb)
LG26 (w=1)
LG26
ρ=0.909
chrB02 (21Mb)
LG27 (w=1)
LG27
ρ=0.973
chrB03 (20Mb)
LG28 (w=1)
LG28
ρ=0.960
chrB04 (16Mb)
LG29 (w=1)
LG29
ρ=0.820
chrB05 (20Mb)
LG30 (w=1)
LG30
ρ=0.856
chrB06 (19Mb)
LG31 (w=1)
LG31
ρ=0.948
chrB07 (21Mb)
LG32 (w=1)
LG32
ρ=0.940


0 cM
chrB08 (17Mb)
LG33 (wr=1)
68 cM
0 2 4 6 8 10 12 14 16 Mb
p=0.928
LG33
0 cM
chrB09 (19Mb)
LG34 (wr=1)
58 cM
0 2 4 6 8 10 12 14 16 18 Mb
p=0.899
LG34
0 cM
chrB10 (21Mb)
LG35 (wr=1)
73 cM
0 5 10 15 20 Mb
p=0.981
LG35
0 cM
chrB11 (21Mb)
LG36 (wr=1)
74 cM
0 5 10 15 20 Mb
p=0.921
LG36
0 cM
chrB12 (28Mb)
LG37 (wr=1)
76 cM
0 5 10 15 20 25 Mb
p=0.855
LG37
0 cM
chrB13 (22Mb)
LG38 (wr=1)
78 cM
0 5 10 15 20 Mb
p=0.976
LG38

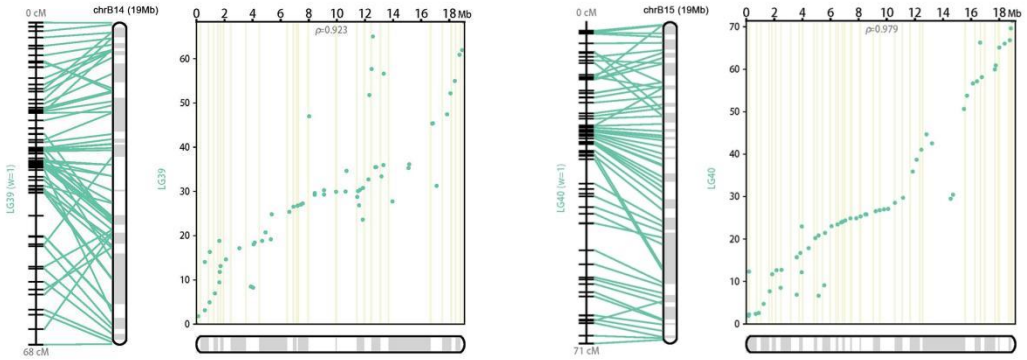


Fig. S3. Pseudo-chromosome level assembly collinearity to genetic maps with ALLMAPS. Pseudo-chromosomes were constructed by Hi-C data using ALLHiC (80, 81). A genetic map from crucian carp (82) was used to evaluate the accuracy of chromosome level assembly using ALLMAPS (83) with equal weights. Green lines connect the goldfish genome scaffolds to the linkage group of crucian carp.

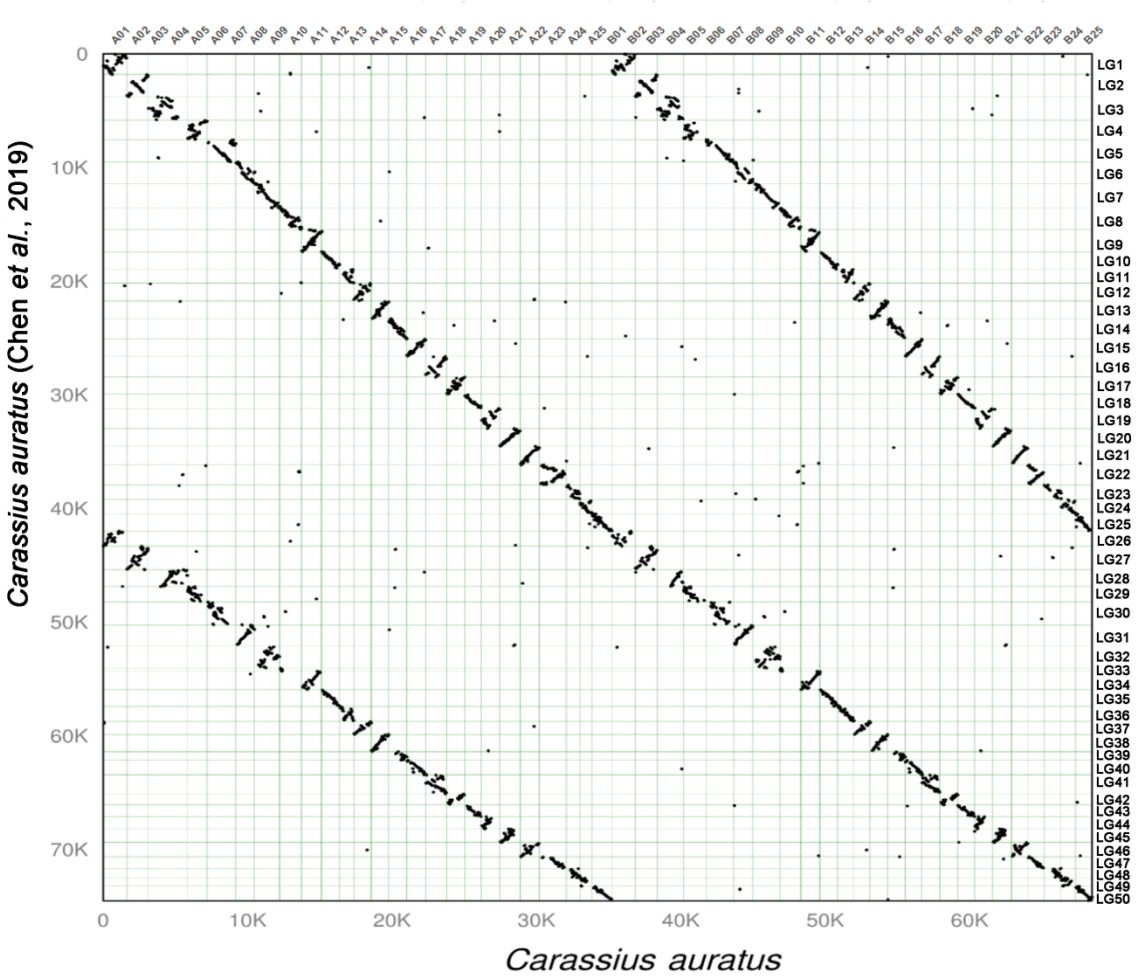


Fig. S4. Alignment of new and published *C. auratus* chromosome dot plots. Homologous gene pairs between the two goldfish species are represented by a dark dot. The genome sequence of *C. auratus* (x axis) was aligned and plotted against that of the published reference strain (y axis) by MUMmer. The diagonal line confirms the very high similarity of both genomes. The dots represent the positions of conserved DNA sequences in the genomes. The middle, main diagonal line indicates the collinearity between the homologous chromosome pairs of each ChrA01-A25 and ChrB01-B25 between the two goldfish assemblies (i.e. A-A' and B-B'), whereas the bottom left and upper right diagonal lines show the collinearity of the homeologous chromosome pairs between subgenome A in one assembly and subgenome B in the other one, revealing the signals from the WGD ohnologs (i.e. A-B' and B-A').

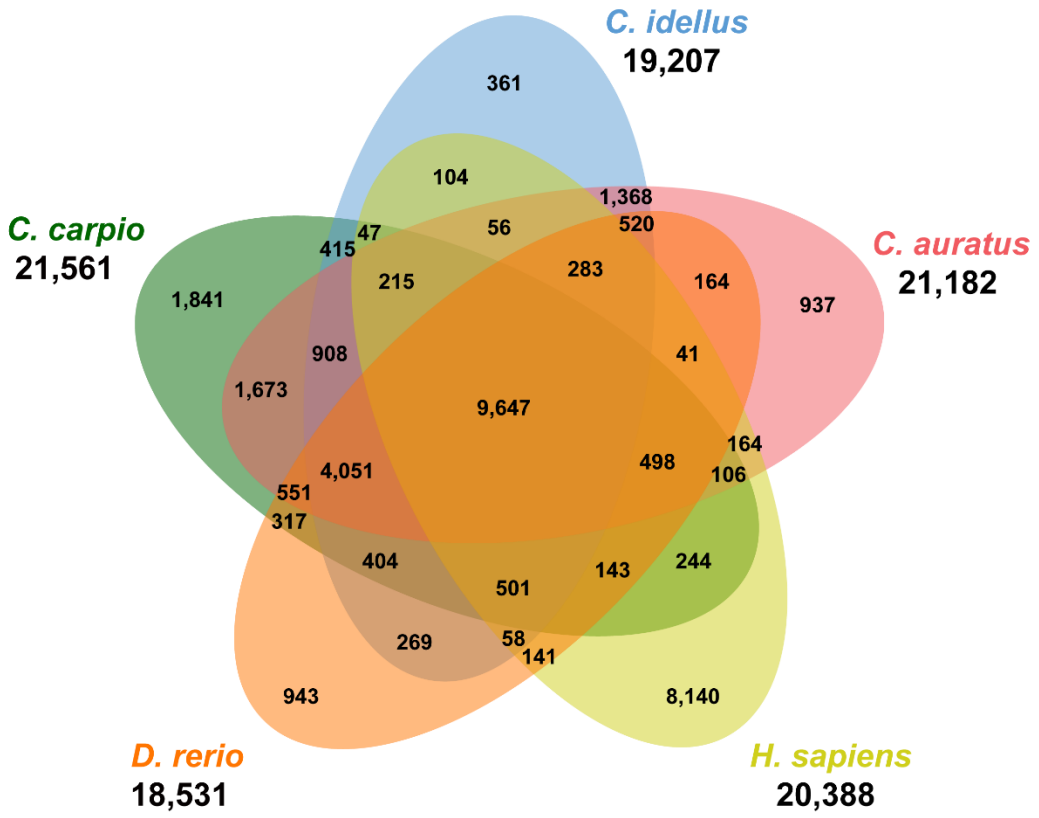


Fig. S5. Venn diagram of *Carassius auratus, Cyprinus carpio, Ctenopharyngodon idellus, Danio rerio,* and human. The orthologous gene clusters were shown in the diagram.

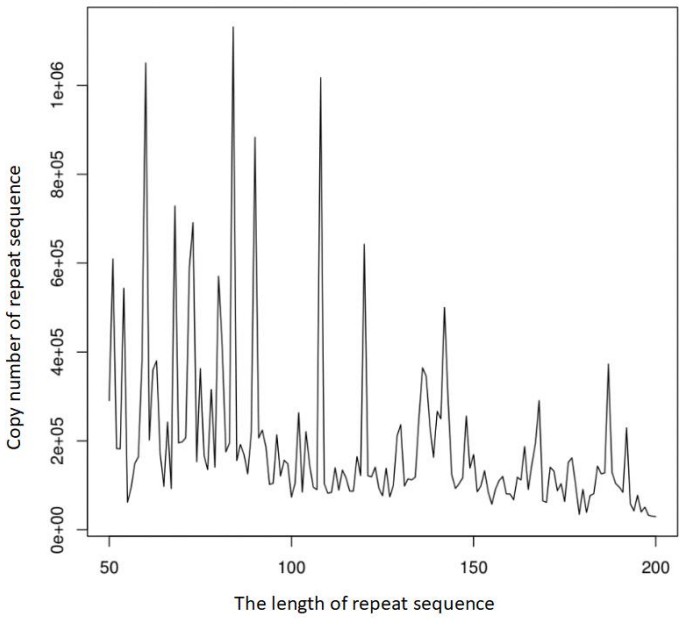


Fig. S6. Prediction of centromere model. The centromere type of goldfish was monocentric. A base monomer length was 84 bp.

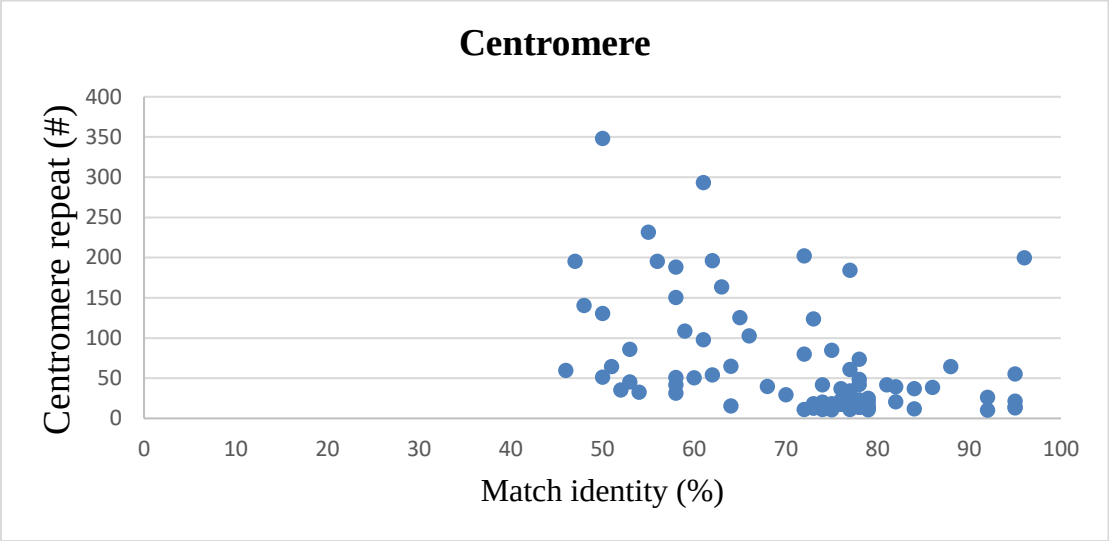


Fig. S7. Correlation between copy number of repeat and the match identity. The match identity between monomers decreases as the copy number of a repeat increases. The scatter diagram was drawn by a data matrix of centromeres predicted in genome BMerged2A. As shown, the low single-copy numbers of the centromere (⩽100) have a high matching rate (⩾70%), in contrast with the high copy numbers. This result provides insights into the centromere characteristics of the goldfish genome, revealing an inverse relationship between copy number and match identity between the centromeric repeats.

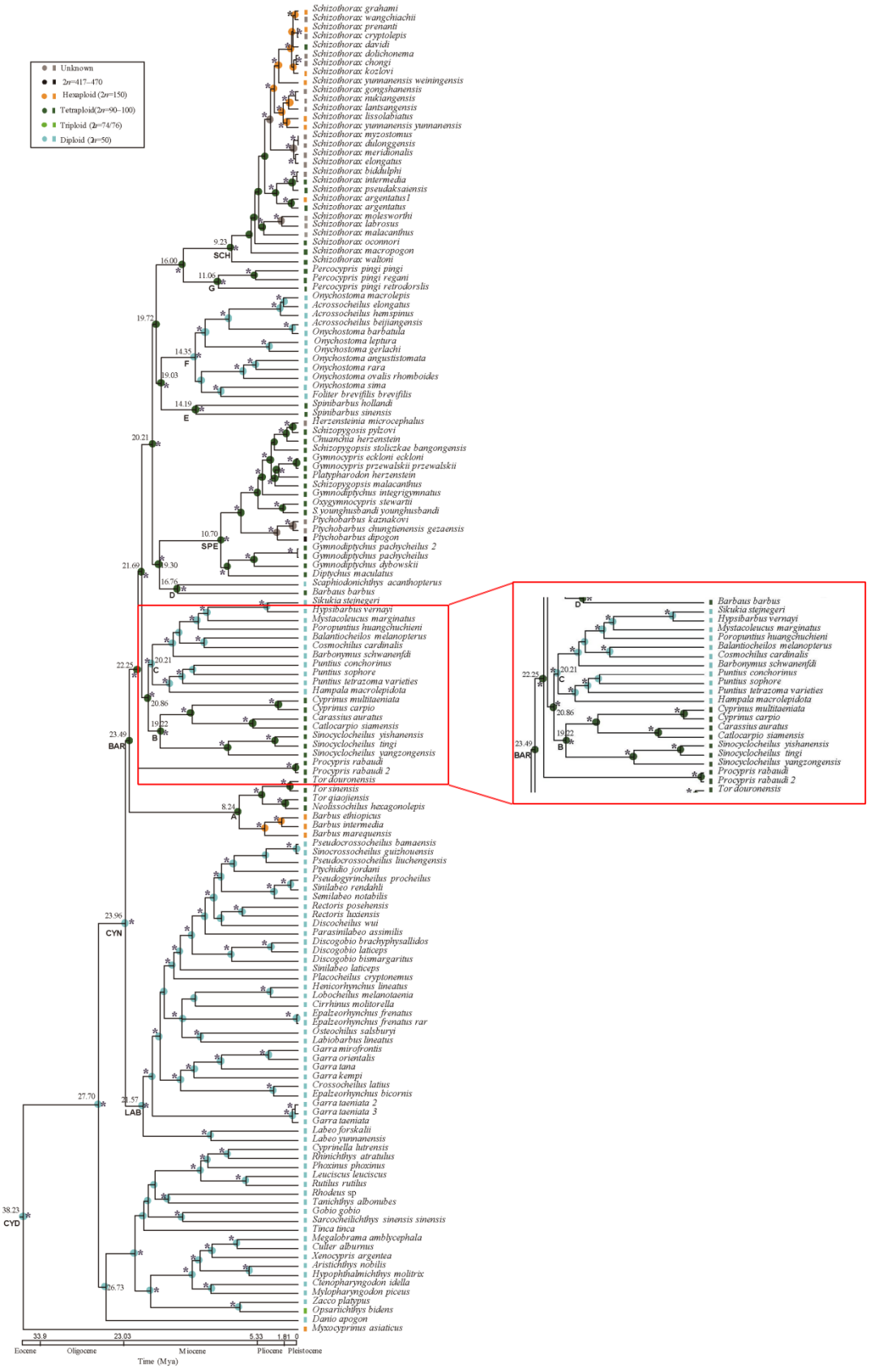
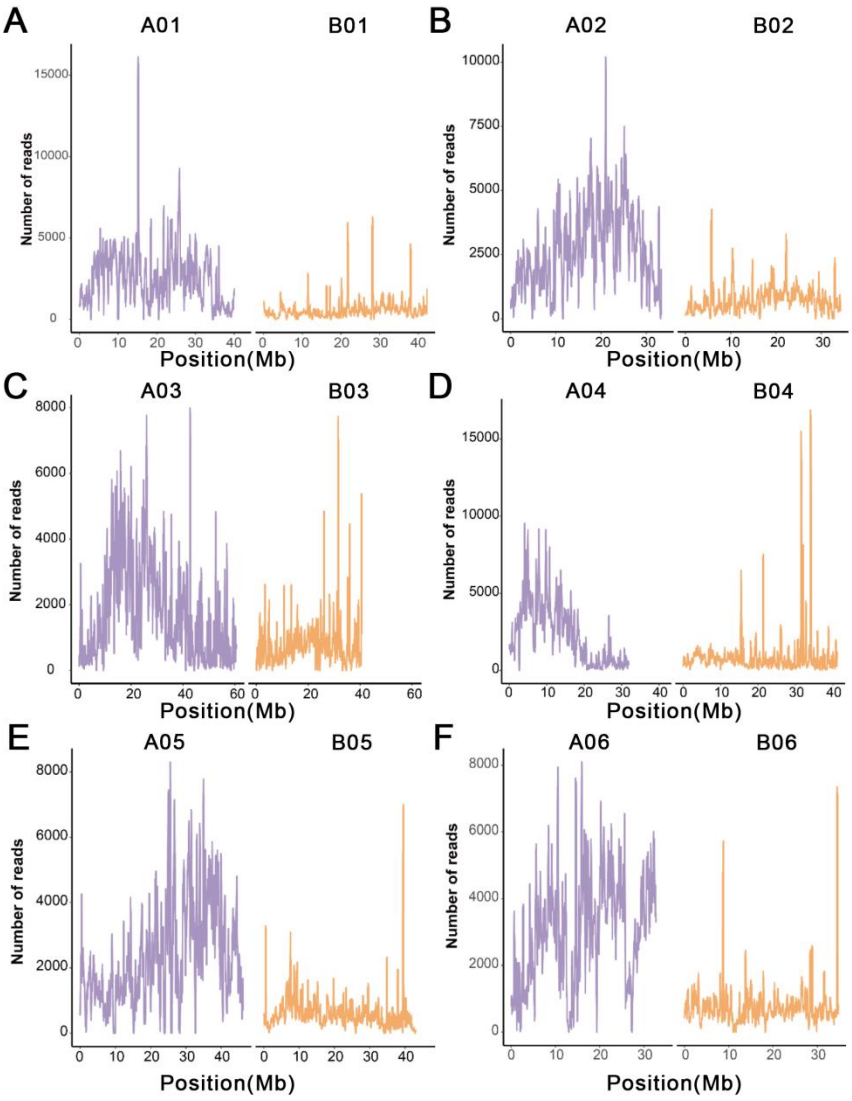
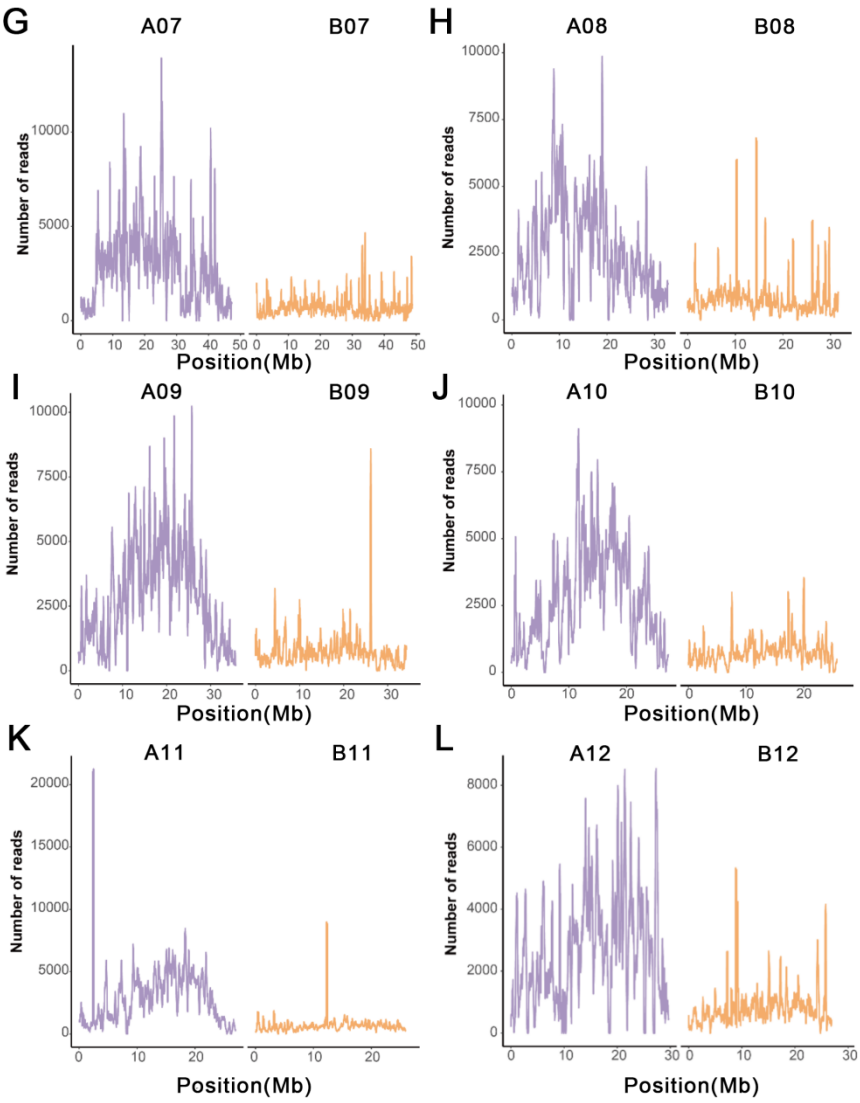
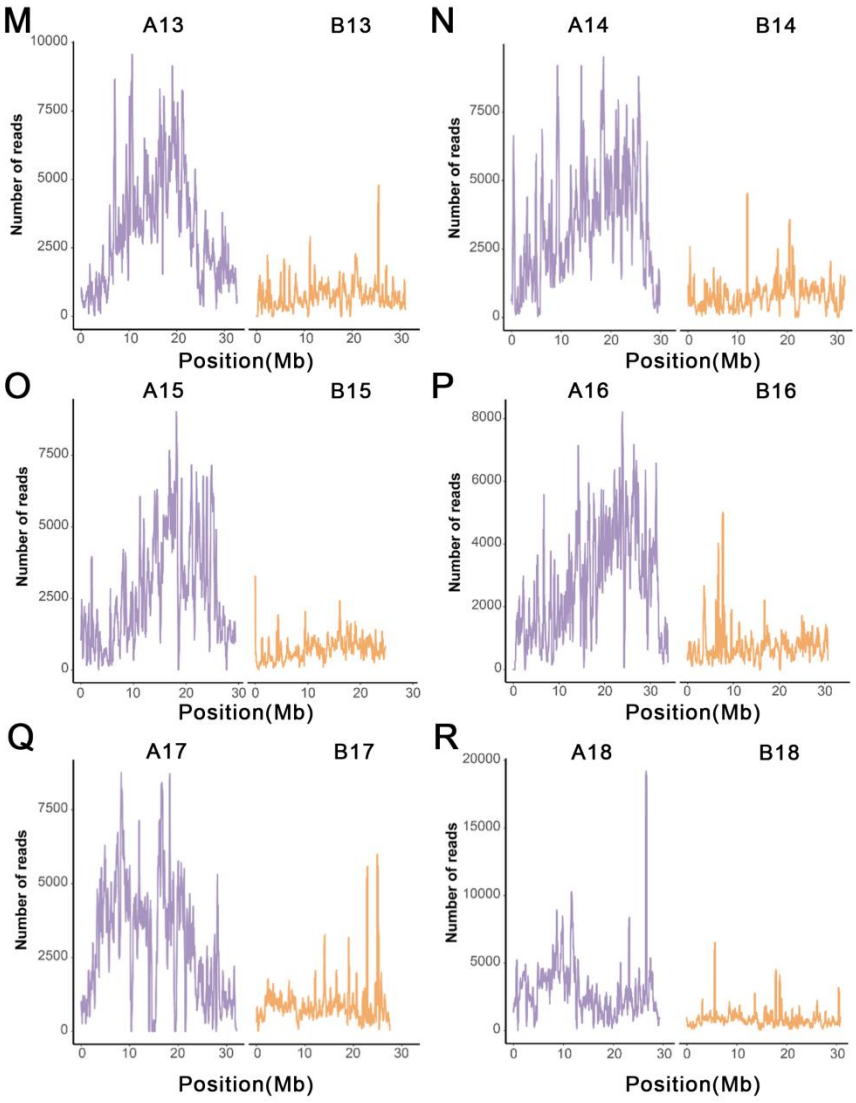
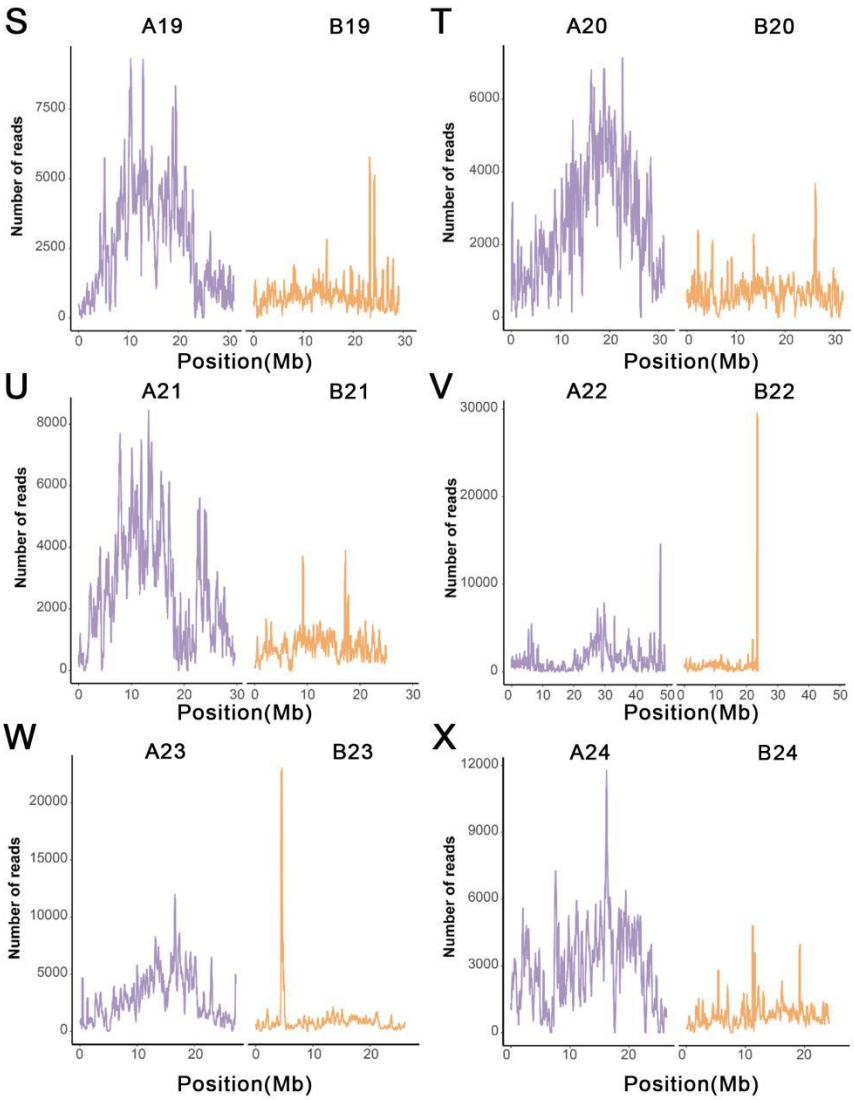

Unknown
2n=417–470
Hexaploid (2n=150)
Tetraploid(2n=90–100)
Triploid (2n=74/76)
Diploid (2n=50)
Schizothorax grahami
Schizothorax wangchiachii
Schizothorax prenanti
Schizothorax cryptolepis
Schizothorax davidi
Schizothorax dolichonema
Schizothorax chongi
Schizothorax kozlovi
Schizothorax yunnanensis weiningensis
Schizothorax gongshanensis
Schizothorax nukiangensis
Schizothorax lantsangensis
Schizothorax lissolabiatus
Schizothorax yunnanensis yunnanensis
Schizothorax myzostomus
Schizothorax dulongensis
Schizothorax meridionalis
Schizothorax elongatus
Schizothorax biddulphi
Schizothorax intermedia
Schizothorax pseudaksaiensis
Schizothorax argentatus1
Schizothorax argentatus
Schizothorax molesworthi
Schizothorax labrosus
Schizothorax malacanthus
Schizothorax oconnori
Schizothorax macropogon
Schizothorax waltoni
Percocypris pingi pingi
Percocypris pingi regani
Percocypris pingi retrodorslis
Onychostoma macrolepis
Acrossocheilus elongatus
Acrossocheilus hamspinus
Acrossocheilus beijiangensis
Onychostoma barbatula
Onychostoma leptura
Onychostoma gerlachi
Onychostoma angustistomata
Onychostoma rara
Onychostoma ovalis rhomboides
Onychostoma sima
Folifer brevifilis brevifilis
Spinibarbus hollandi
Spinibarbus sinensis
Herzensteinia microcephalus
Schizopygosis pylzovi
Chuanchia herzenstein
Schizopygopsis stoliczkae bangongensis
Gymnocypris eckloni eckloni
Gymnocypris przewalskii przewalskii
Platypharodon herzenstein
Schizopygopsis malacanthus
Gymnodiptychus integrigymnatus
Oxygymnocypris stewartii
S younghusbandi younghusbandi
Ptychobarbus kaznakovi
Ptychobarbus chungtienensis gezaensis
Ptychobarbus dipogon
Gymnodiptychus pachycheilus 2
Gymnodiptychus pachycheilus
Gymnodiptychus dybowskii
Diptychus maculatus
Scaphiodonichthys acanthopterus
Barbaus barbus
Sikukia stejnegeri
Hypsibarbus vernayi
Mystacoleucus marginatus
Poropuntius huangchuchieni
Balantiocheilos melanopterus
Cosmochilus cardinalis
Barbonymus schwanenfdi
Puntius conchorinus
Puntius sophore
Puntius tetrazoma varieties
Hampala macrolepidota
Cyprinus multitaeniata
Cyprinus carpio
Carassius auratus
Catlocarpio siamensis
Sinocyclocheilus yishanensis
Sinocyclocheilus tingi
Sinocyclocheilus yangzongensis
Procypris rabaudi
Procypris rabaudi 2
Tor douronensis
Tor sinensis
Tor qiaojiensis
Neolissochilus hexagonolepis
Barbus ethiopicus
Barbus intermedia
Barbus marequensis
Pseudocrossocheilus bamaensis
Sinocrossocheilus guizhouensis
Pseudocrossocheilus liuchengensis
Ptychidio jordani
Pseudogyrincheilus procheilus
Sinilabeo rendahli
Semilabeo notabilis
Rectoris posehensis
Rectoris luxiensis
Discocheilus wui
Parasinilabeo assimilis
Discogobio brachyphysallidos
Discogobio laticeps
Discogobio bismargaritus
Similabeo laticeps
Placocheilus cryptonemus
Henicorhynchus lineatus
Lobocheilus melanotaenia
Cirrhinus molitorella
Epalzeorhynchus frenatus
Epalzeorhynchus frenatus rar
Osteochilus salsburyi
Labiobarbus lineatus
Garra mirofrontis
Garra orientalis
Garra tana
Garra kempi
Crossocheilus latius
Epalzeorhynchus bicornis
Garra taeniata 2
Garra taeniata 3
Garra taeniata
Labeo forskalii
Labeo yunnanensis
Cyprinella lutrensis
Rhinichthys atratulus
Phoxinus phoxinus
Leuciscus leuciscus
Rutilus rutilus
Rhodeus sp
Tanichthys albonubes
Gobio gobio
Sarcocheilichthys sinensis sinensis
Tinca tinca
Megalobrama amblycephala
Culter alburnus
Xenocypris argentea
Aristichthys nobilis
Hypophthalmichthys molitrix
Ctenopharyngodon idella
Mylopharyngodon piceus
Zacco platypus
Opsariichthys bidens
Danio apogon
Myxocyprinus asiaticus
9.23
SCH
16.00
11.06
G
19.72
14.35
F
19.03
14.19
E
20.21
10.70
SPE
21.69
19.30
16.76
D
22.25
20.21
C
20.86
19.22
B
23.49
BAR
8.24
A
23.96
CYN
27.70
20.57
LAB
38.23
CYD
26.73
33.9
23.03
5.33
1.81
0
Eocene
Oligocene
Miocene
Pliocene
Pleistocene
Time (Mya)

Fig. S8. Cyprininae phylogeny [modified from (85)]. The six representative diploid Barbinae species (2n = 50), including *Puntius semifasciolatus, Hypsibarbus vernayi, Mystacoleucus marginatus, Balantiochelos melanopterus, Barbonymus schwanenfdi,* and *Hampala macrolepidota* are underlined.


A
A01
B01
B
A02
B02
C
A03
B03
D
A04
B04
E
A05
B05
F
A06
B06
Number of reads
Position(Mb)


G
A07
B07
H
A08
B08
I
A09
B09
J
A10
B10
K
A11
B11
L
A12
B12
Number of reads
Position(Mb)


M
A13
B13
N
A14
B14
O
A15
B15
P
A16
B16
Q
A17
B17
R
A18
B18
Number of reads
Position(Mb)


S
A19
B19
T
A20
B20
U
A21
B21
V
A22
B22
W
A23
B23
X
A24
B24
Number of reads
Position(Mb)

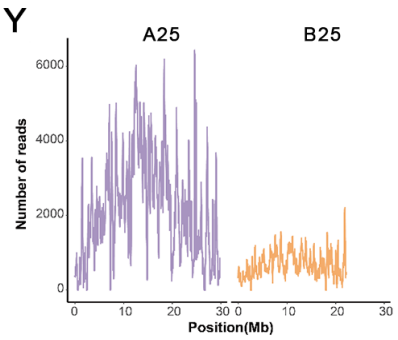


Fig. S9. Distribution of *Puntius semifasciolatus* reads mapped against the 25 goldfish homeologous chromosome pairs. The A (ChrA01~A25) and B (ChrB01-B25) subgenomes are marked with purple and orange lines, respectively.

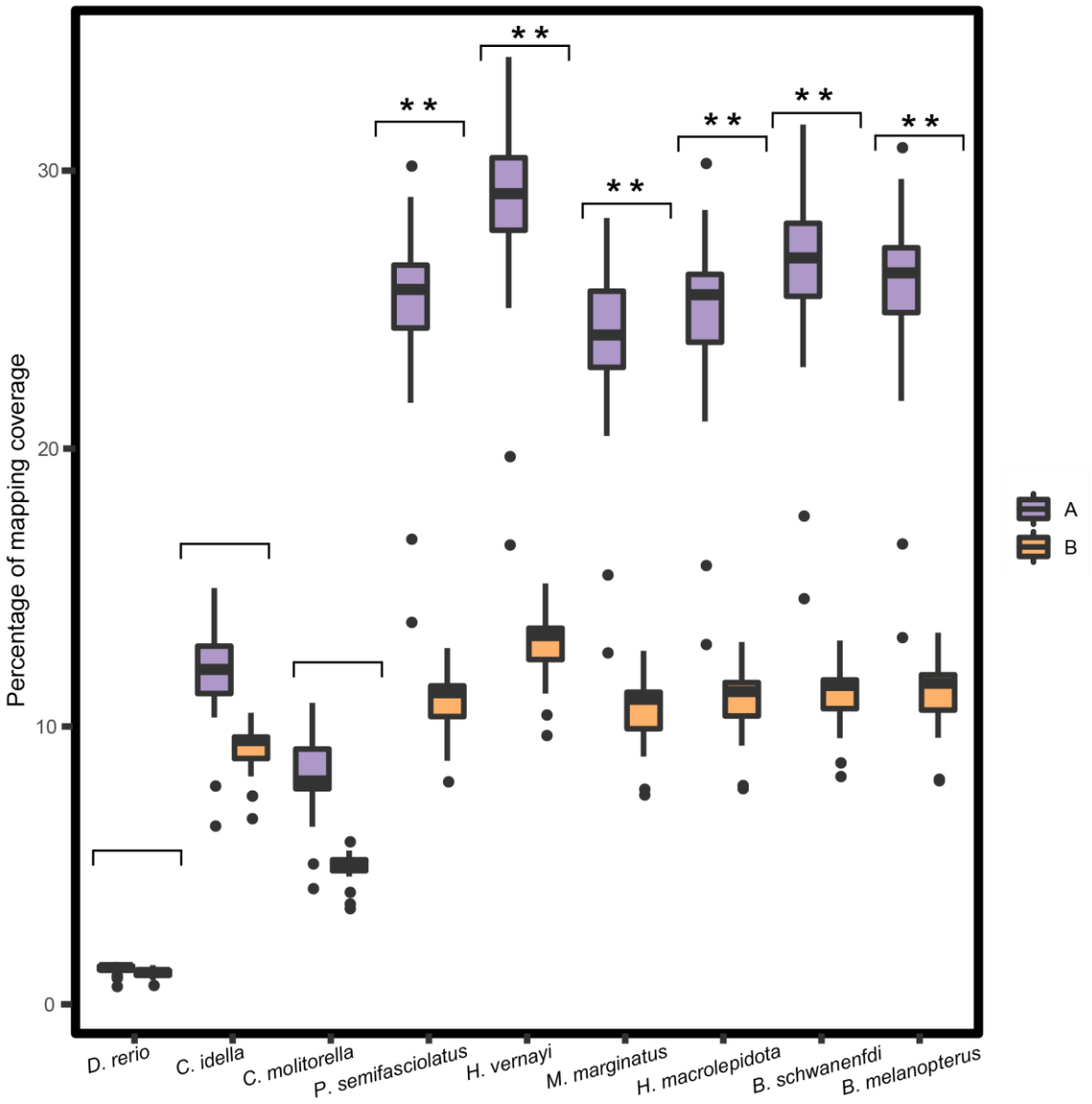


Fig. S10. Mapping coverage of *D. rerio*, *C. idella* and representative *Barbinae* species against the 25 goldfish homeologous chromosome pairs. Boxplots from left to right represent *D. rerio*, *C. idella*, *C. molitorella*, *P. semifasciolatus*, *H. vernayi*, *M. arginatus*, *H. macrolepidota*, *B. schwanenfdi,* and *B. melanopterus*. The X-axis and Y-axis indicate chromosome position (Mb) and number of reads mapped on the goldfish chromosomes, respectively. The A (ChrA01~A25) and B (ChrB01-B25) subgenomes are marked with purple and orange boxes, respectively.

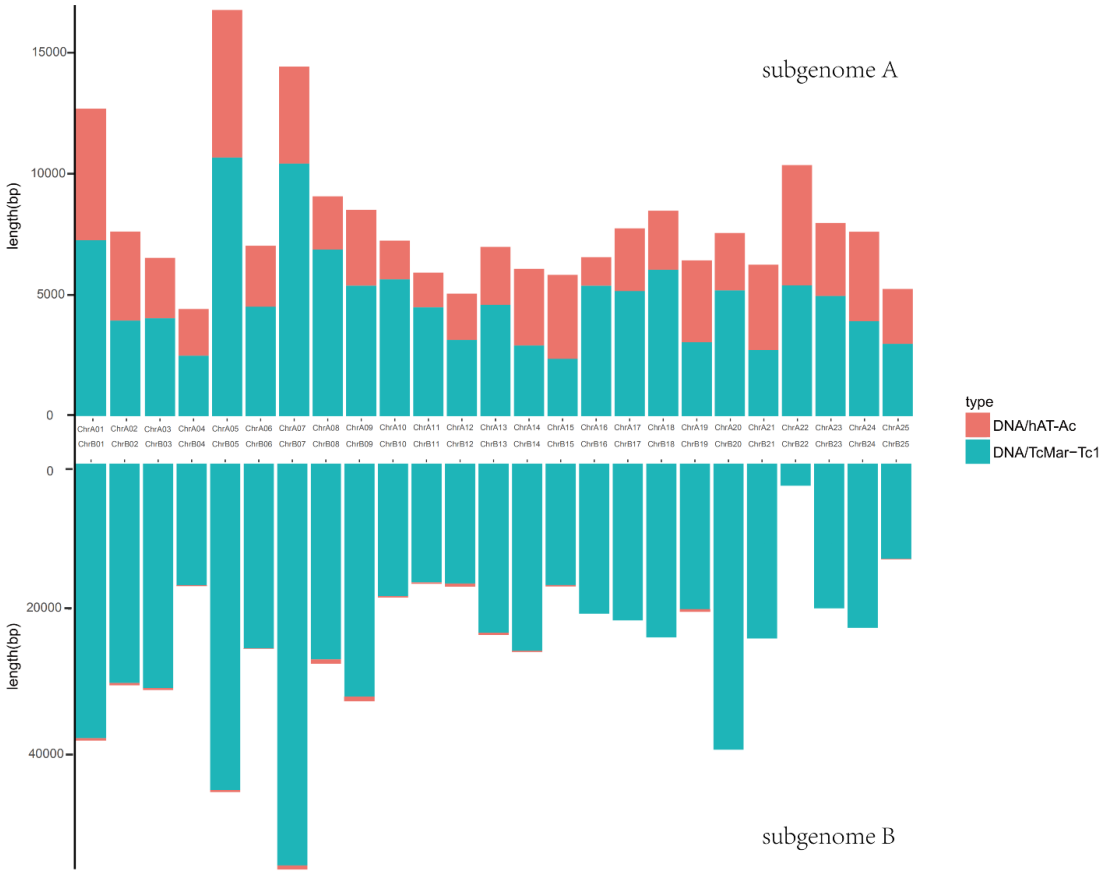


Fig.11 Distribution of the subgenome-specific transposons on each goldfish chromosome.

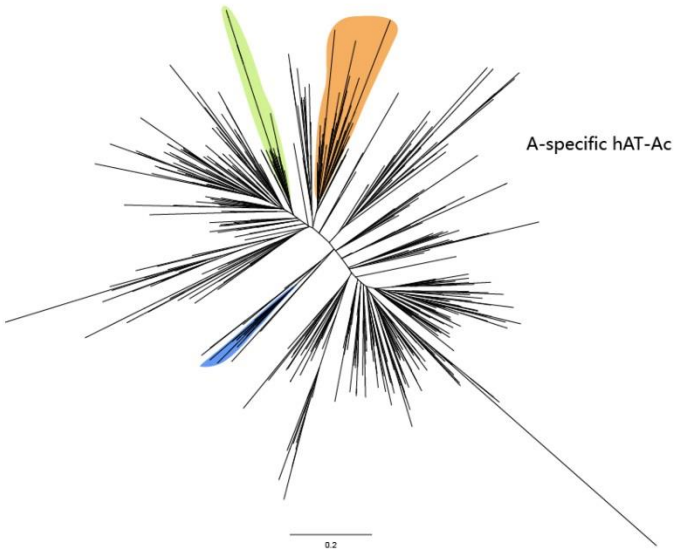


Fig.12 Phylogenetic tree of the hAT-Ac element expansion in the goldfish genome, built by means of Jukes-Cantor corrected distances. Subclusters with enough members to determine accurate timings are highlighted.

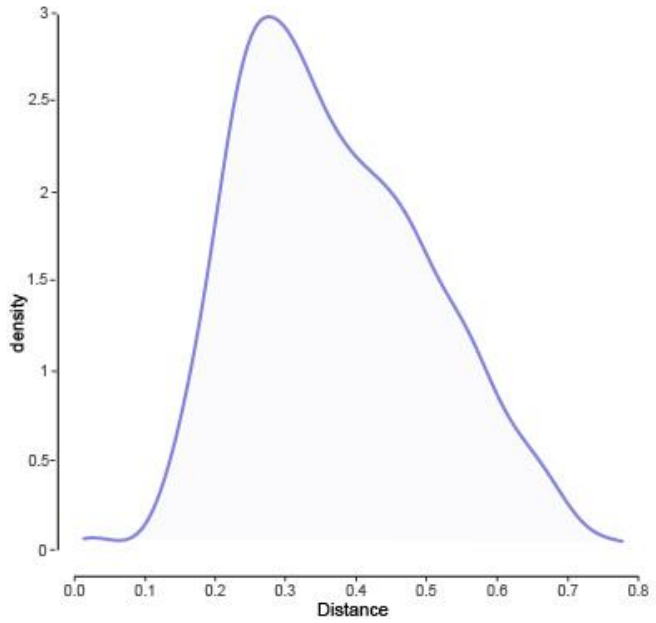


Fig.13 Jukes-Cantor distance of the selected subfamily sequences of transposons against their conservative sequences.

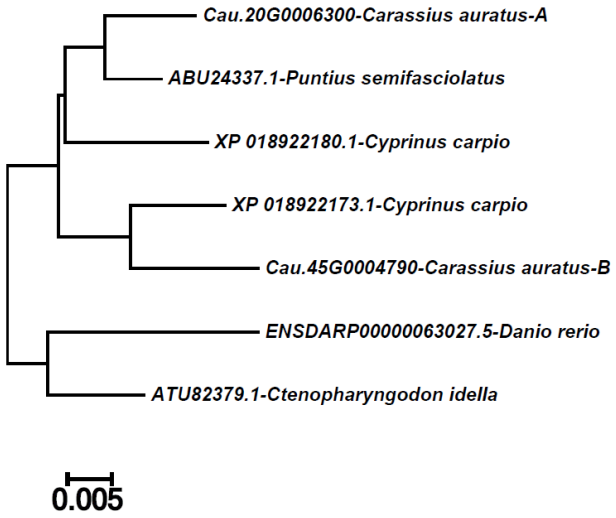


Fig. S14 Phylogenetic tree of connective tissue growth factor-like from Puntius semifasciolatus, goldfish fish, common carp, glass carp and zebrafish.

Fig. S15. Dot-plot of alignment of the genomes of *C. auratus* and *D. rerio*.

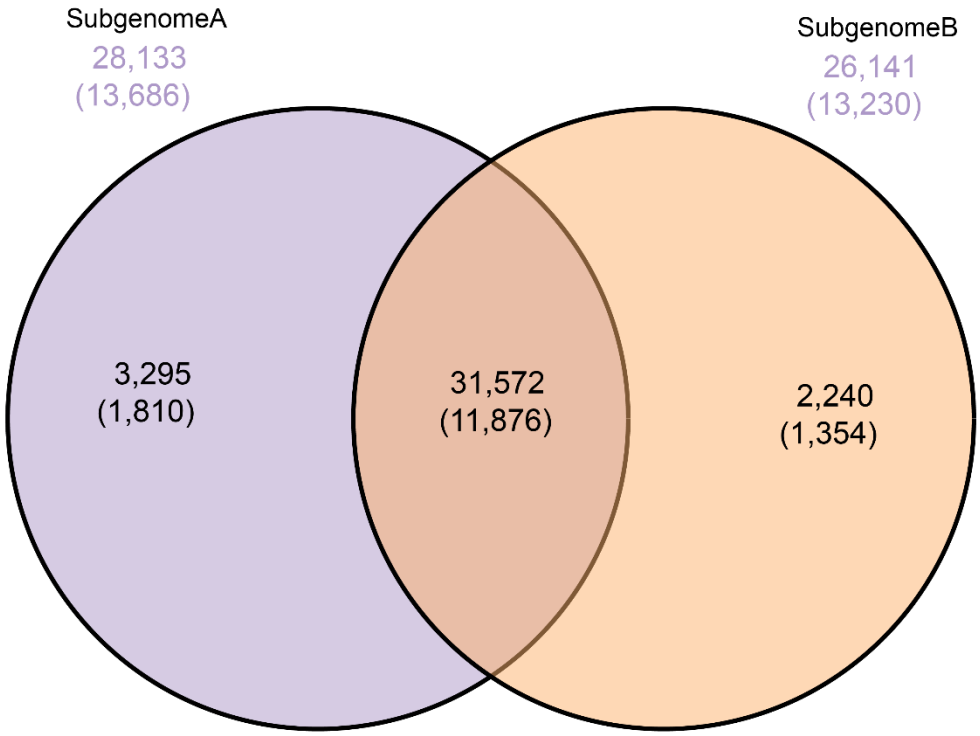


Fig. S16. Venn diagram statistics for genes and gene clusters of subgenomes A and B in *C. auratus*. Subgenomes A and B contain contain 28,133 genes (13,686 gene clusters) and 26,141 genes (13,230 clusters), respectively. A total of 31,572 genes (11,876 gene clusters) were shared between subgenomes A and B, of which 3,295 genes (1,810 gene clusters) were specific to subgenome A and 2,240 genes (1,354 gene clusters) to subgenome B.

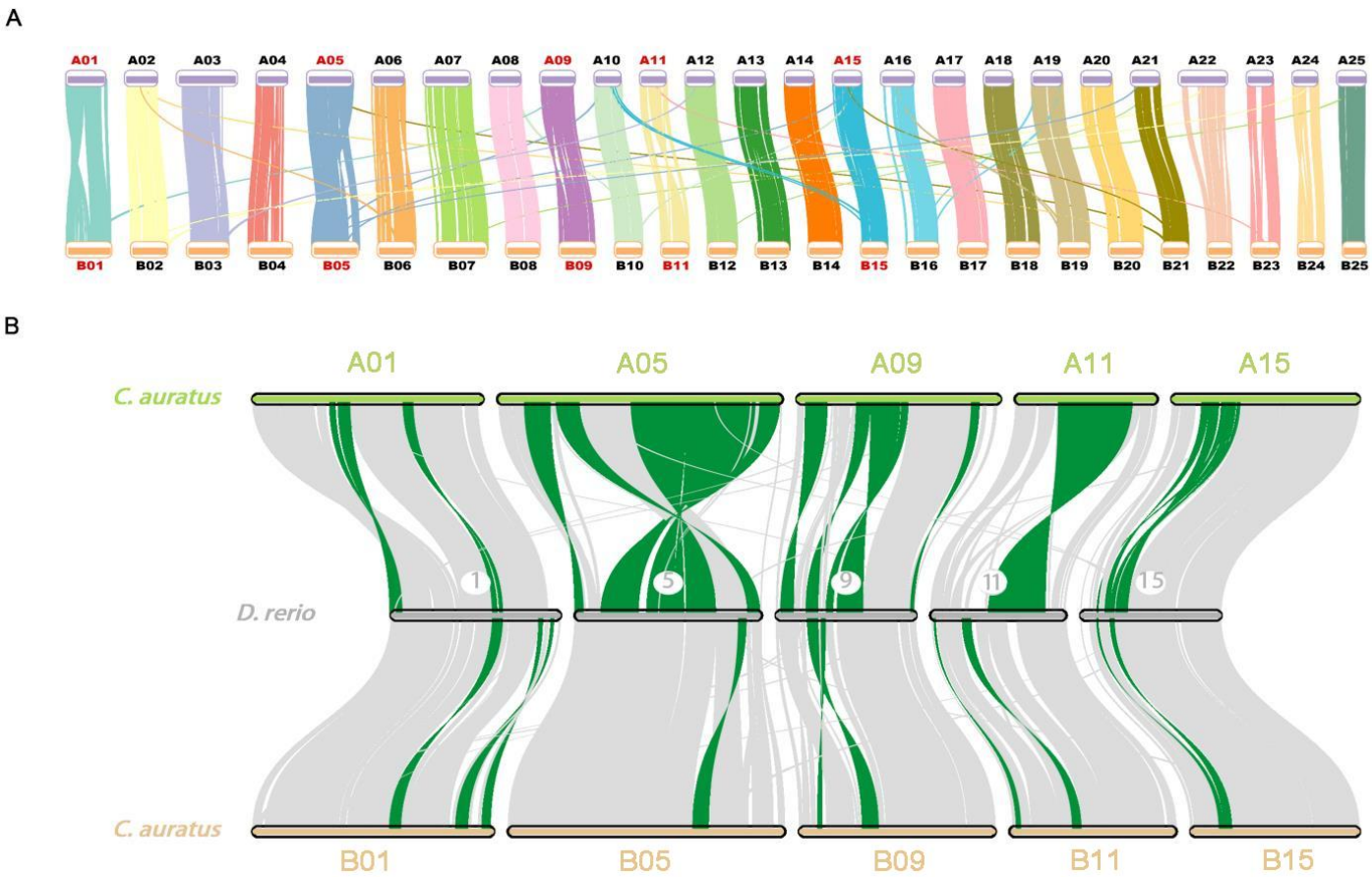


Fig. S17. Synteny analysis of the genomes. (A) Synteny analysis of subgenomes A and B in *C. auratus*. Schematic representation of synteny between subgenomes A and B. Each line connects a pair of orthologous genes between subgenomes A and B. (B) Genomic alignments between ChrA01, ChrA05, ChrA09, ChrA11, ChrA15, ChrB01, ChrB05, ChrB09, ChrB11, and ChrB15 in *Carassius auratus* and corresponding chromosomes in *Ctenopharyngodon idellus* and *Danio rerio*. The inverted regions are shown with green lines.

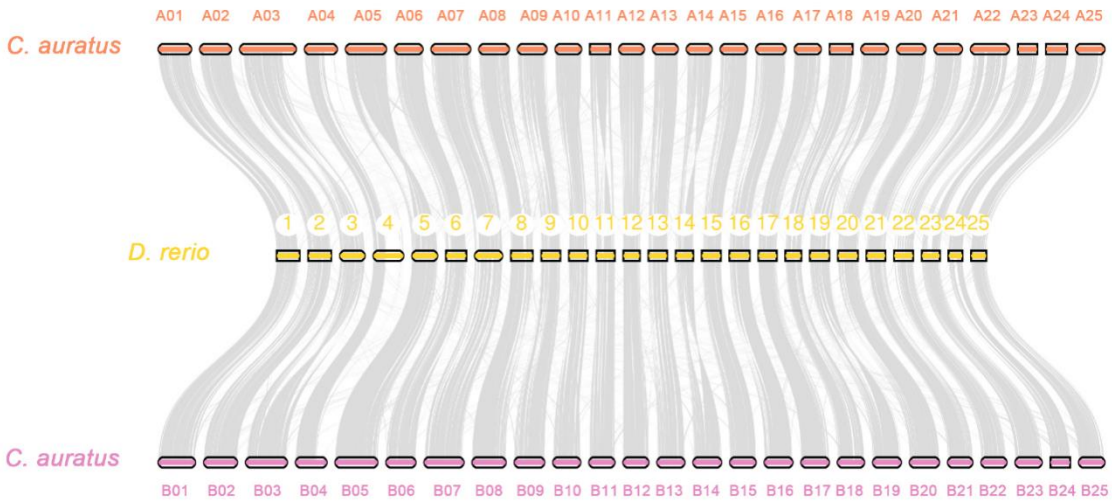


Fig. S18. Schematic representation of the synteny analysis between *Carassius auratus* and *Danio rerio*. Each line connects a pair of orthologous genes.

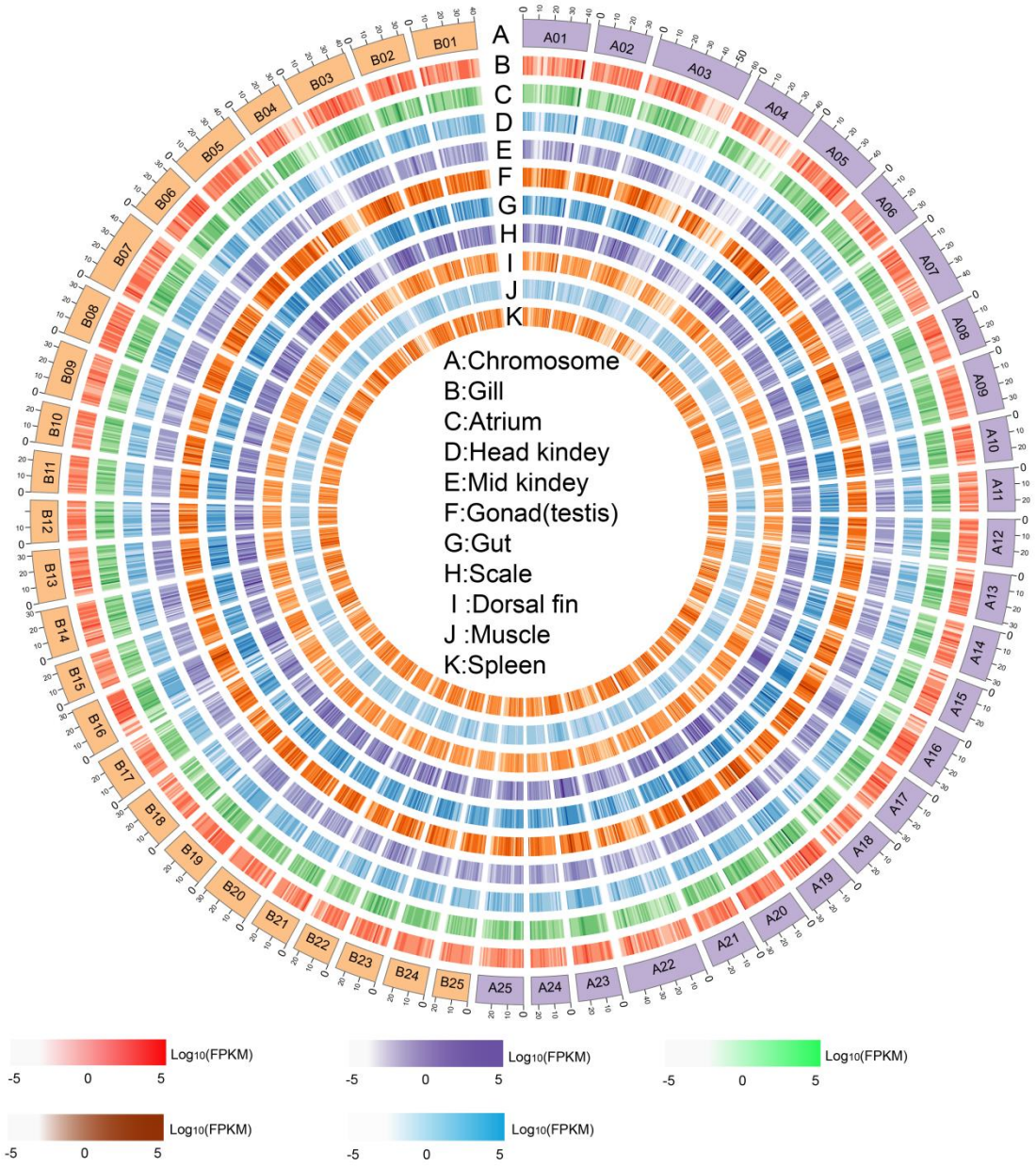


Fig. S19. Genomic expression pattern for ten tissues of *Carassius auratus*. Circos plot showing the expression abundance across the 50 chromosomes for ten tissues. From outer ring to the center, A, chromosome; B, gill; C, atrium; D, head kindey; E, mid kindey; F, gonad (testis); G, gut; H, scale; I, dorsal fin; J, muscle; and K, spleen.

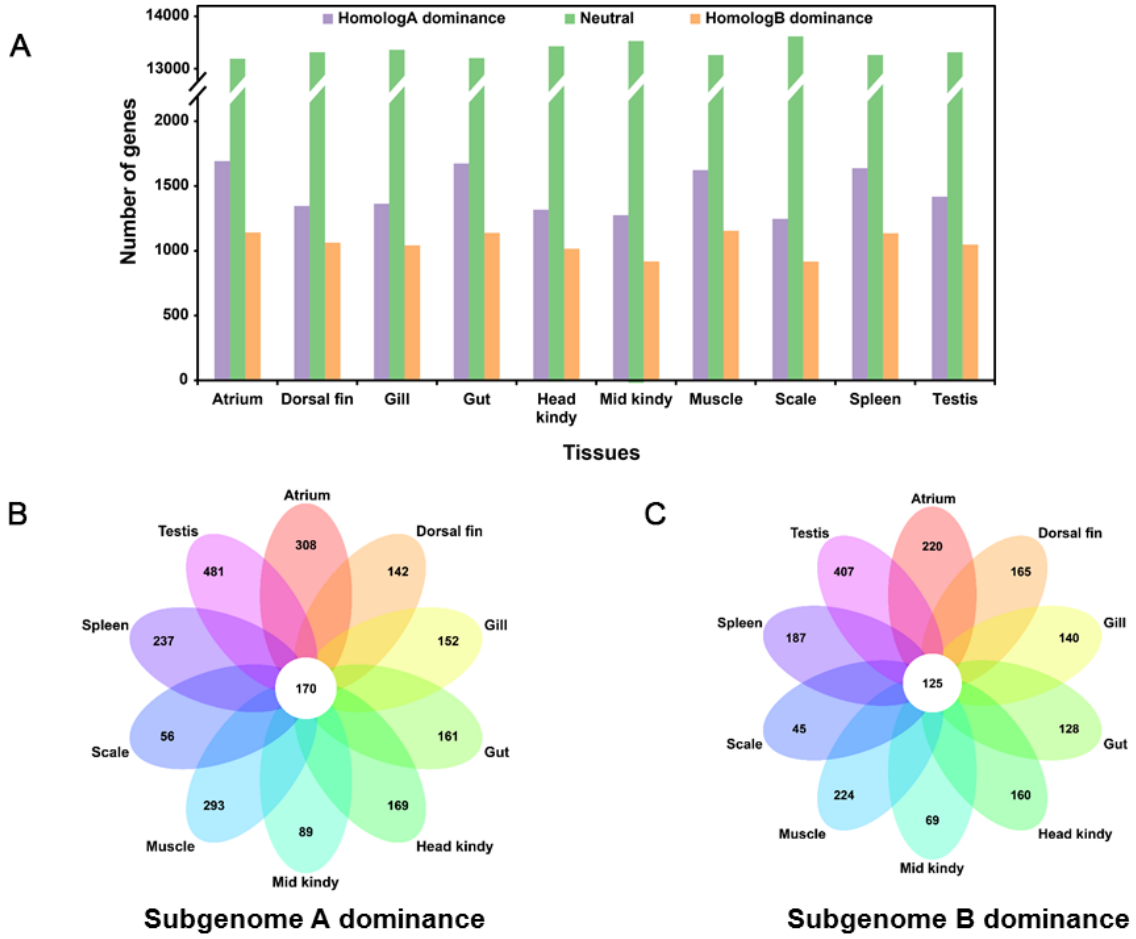


Fig. S20. Asymmetrical homoeologous gene expression in *C. auratus*. (A) Number of genes preferentially expressed [> $\log_2$(FC)] with homolog A, homology B, or neutral dominance in ten tissues (atrium, dorsal fin, gill, gut, head kidney, mid kidney, muscle, scale, spleen, and testis) in *C. auratus*. (b and c) Venn diagrams indicating genes specifically preferentially expressed in each of the tissues with homolog A (B) and homolog B (C) dominance. The central white circos represent the preferentially expressed genes shared in the ten tissues with subgenomes A or B dominance.

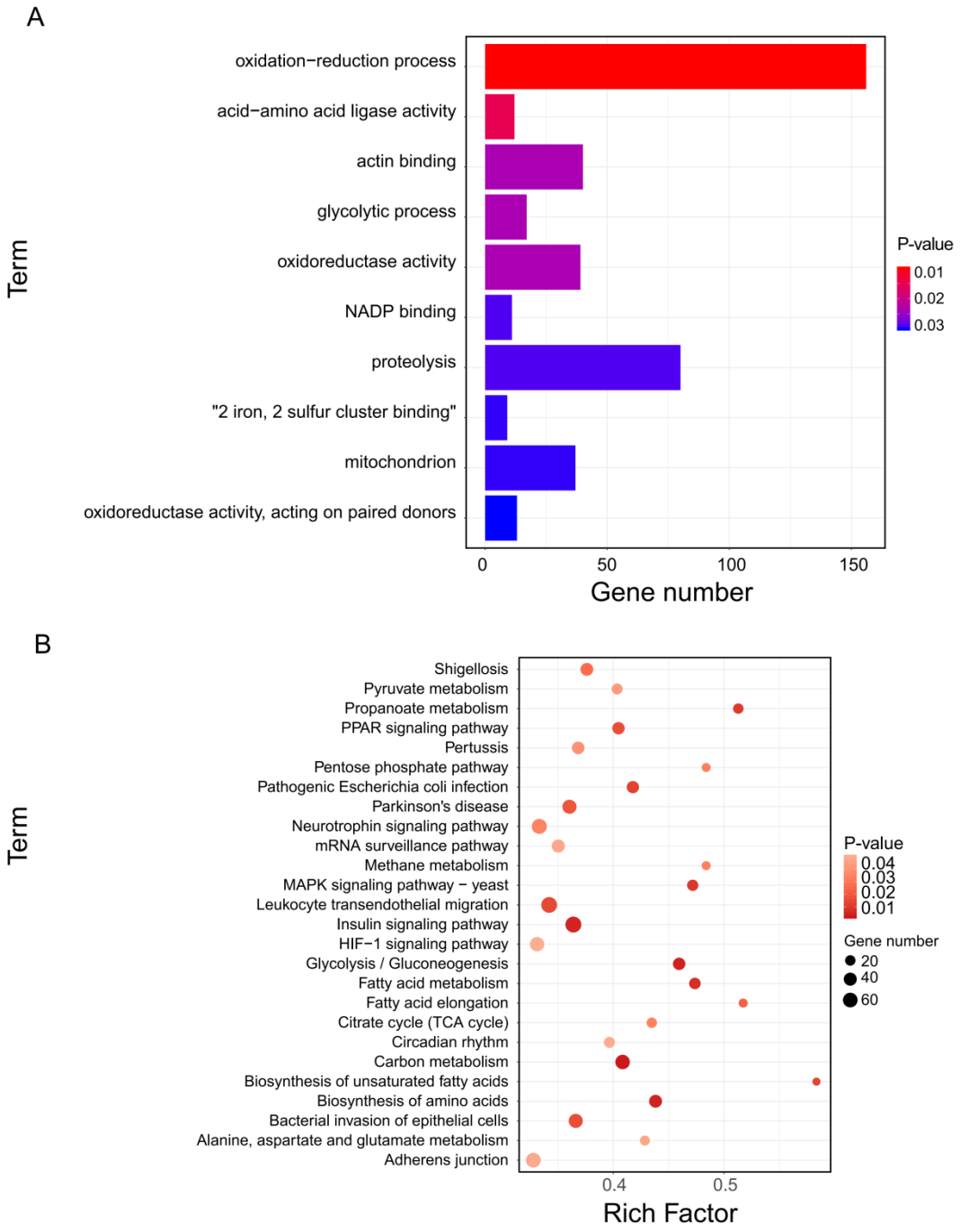


Fig. S21. GO term and KEGG pathway enrichment for 4,123 dominant expressed genes of subgenome A in *C. auratus*. (A) GO term enrichment showing that the dominantly expressed homoeologs of subgenome A were primarily enriched in the oxidation-reduction process, acid-amino ligase activity, actin binding, glycolytic process, and oxidoreductase activity, etc. (B) KEGG pathway analysis indicating that the predominantly expressed homoeologs of subgenome A mainly participate in the insulin signaling pathway, leukocyte transendothelial migration, neurotrophin signaling pathway, and HIF-1 signaling pathway. The size and color of the circles are used to distinguish significance.

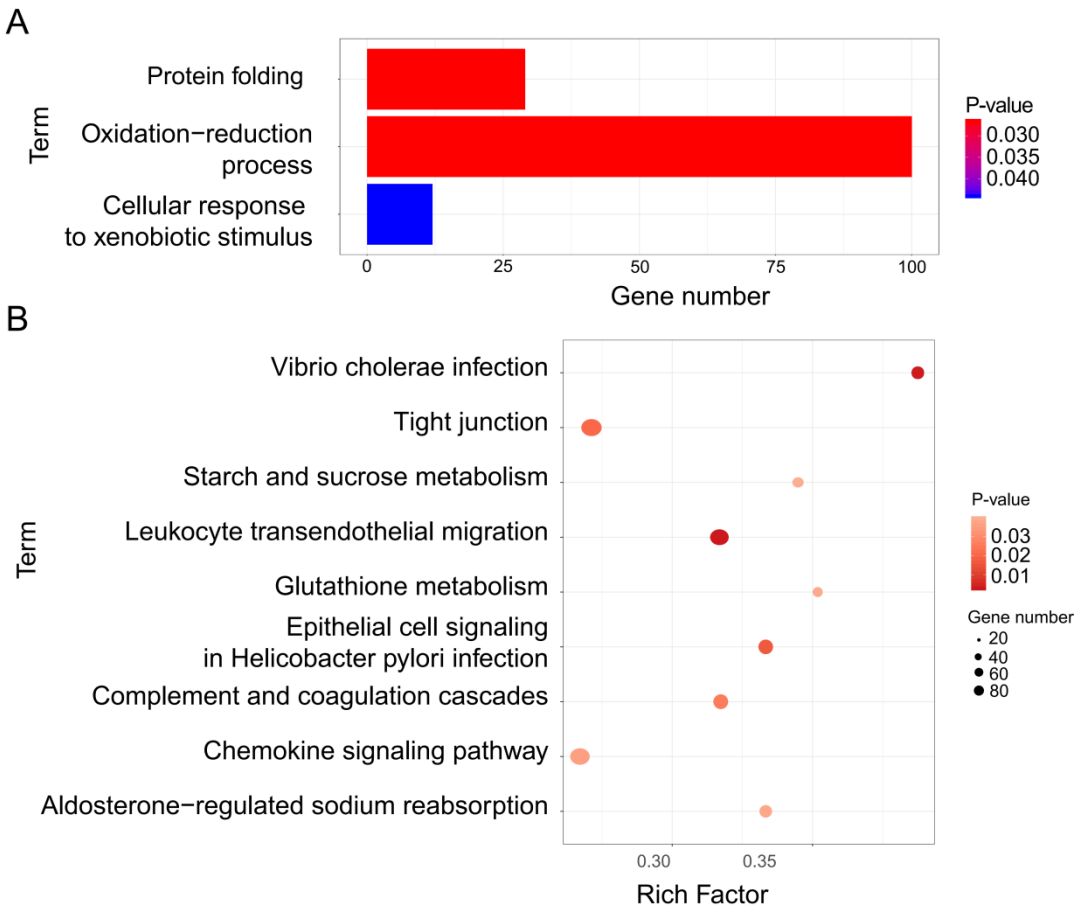


Fig. S22. GO term and KEGG pathway enrichment for the 2,941 dominantly expressed genes of subgenome B in *C. auratus*. (a) GO term enrichment showing that the dominantly expressed homoeologs of subgenome B were primarily enriched in protein folding, oxidation-reduction process, and stimulus response. (b) KEGG pathway analysis indicating that the predominantly expressed homoeologs of subgenome B mainly participate in processes, such as tight junctions, chemokine signaling pathways, and coagulation cascades. The size and color of the circles are used to distinguish significance.

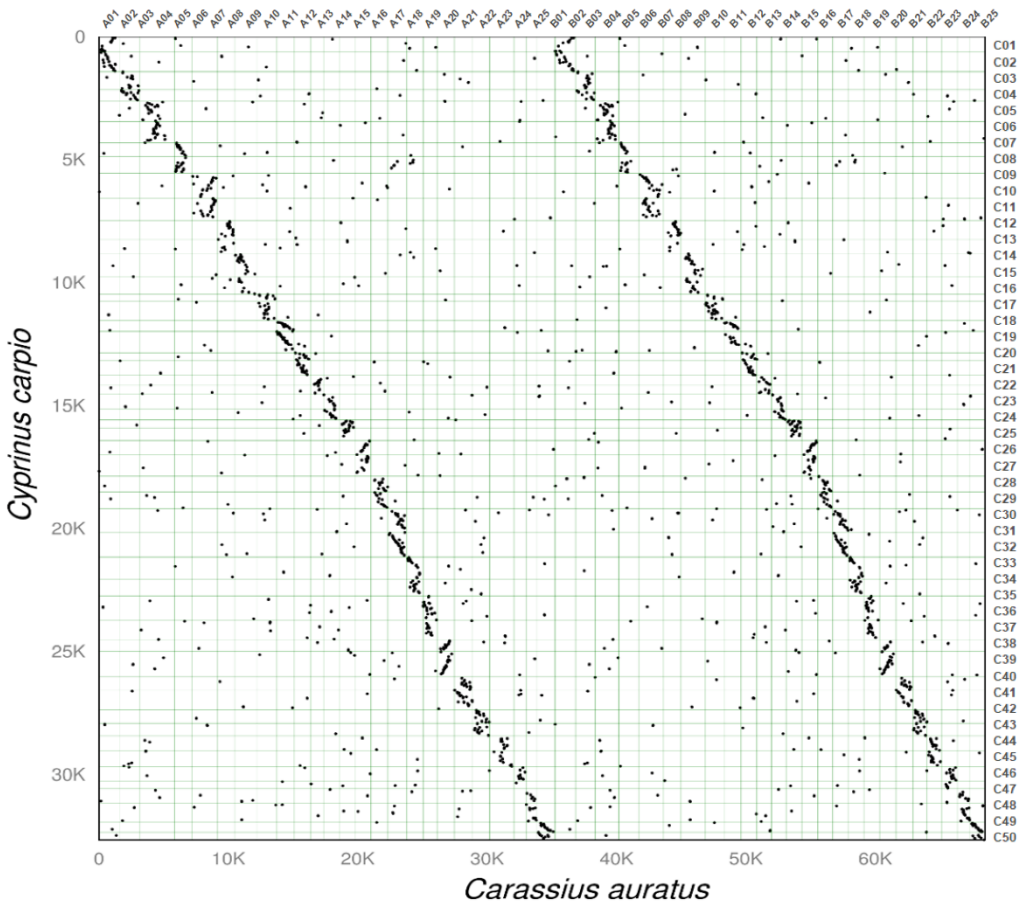


Fig. S23. Dot-plot of alignment of *Carassius auratus* genome with *Cyprinus carpio* genome.

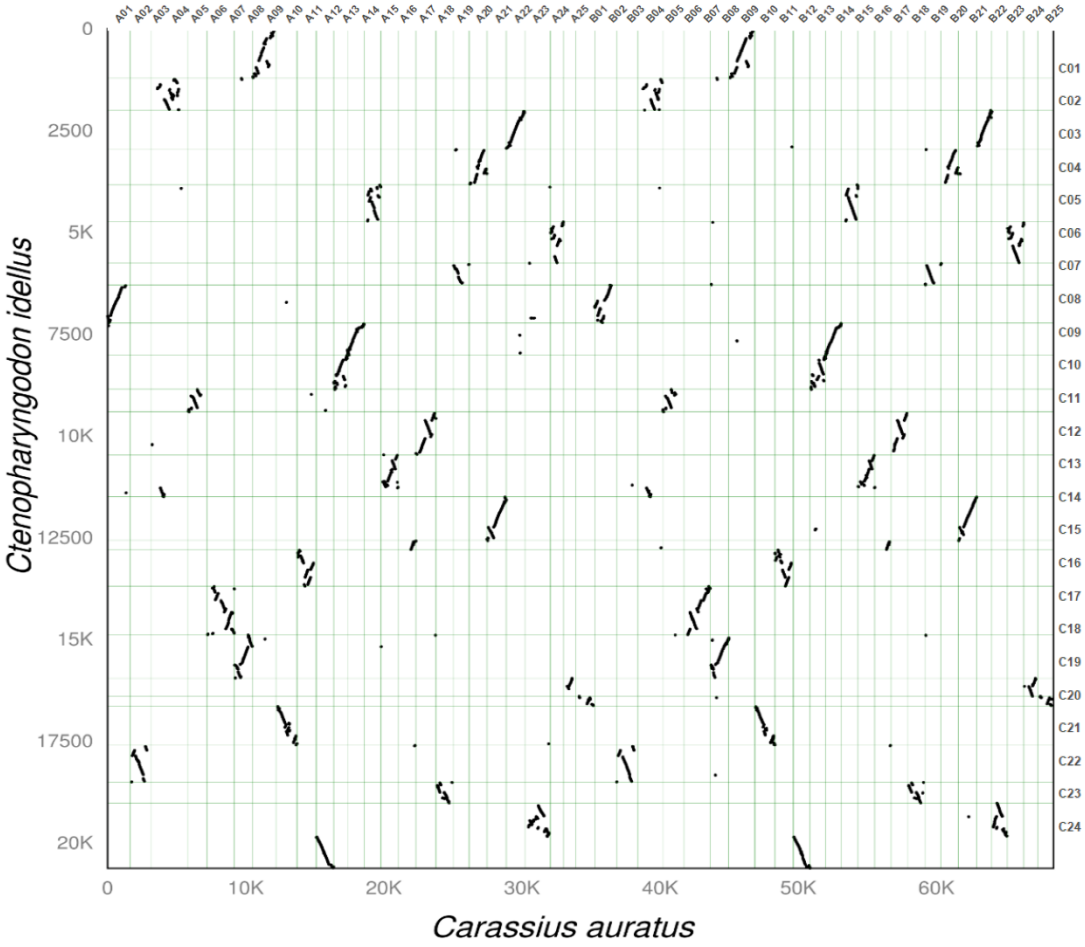


Fig. S24. Alignment of the genomes of *Carassius auratus* with that of *Ctenopharyngodon idellus*.

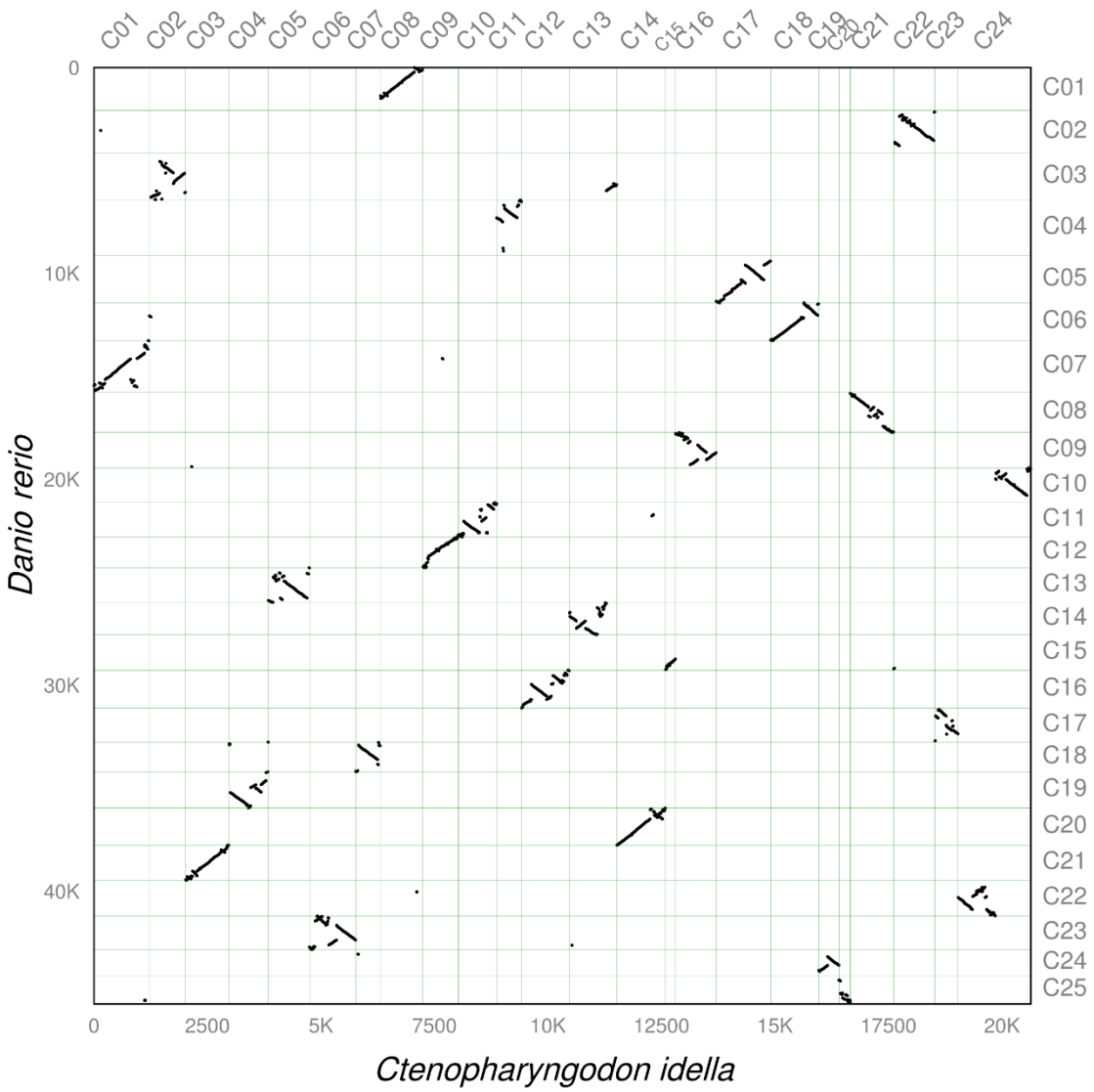
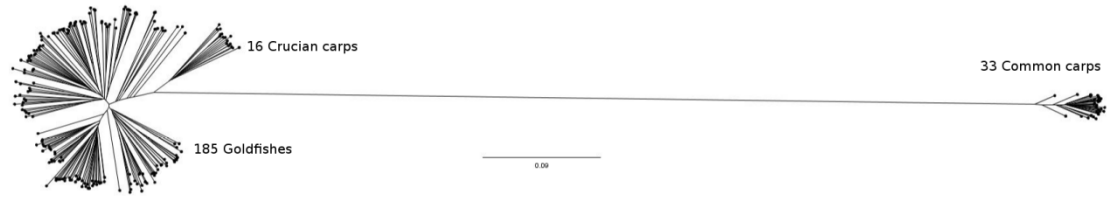


Fig. S25. Alignment of the genome of *Danio rerio* with that of *Ctenopharyngodon idellus*.

Fig. S26. Map of the resequenced samples collected in China. A total of 16 wild crucian carps were collected from the enclosed waters of the potential area of the goldfish origin, including the Chinese provinces of Zhejiang (3), Jiangxi (2), Henan (2), Jiangsu (1), Shangdong (1), Heilongjiang (1), Fujian (4), Yunnan (1), and Anhui (1).

A



B

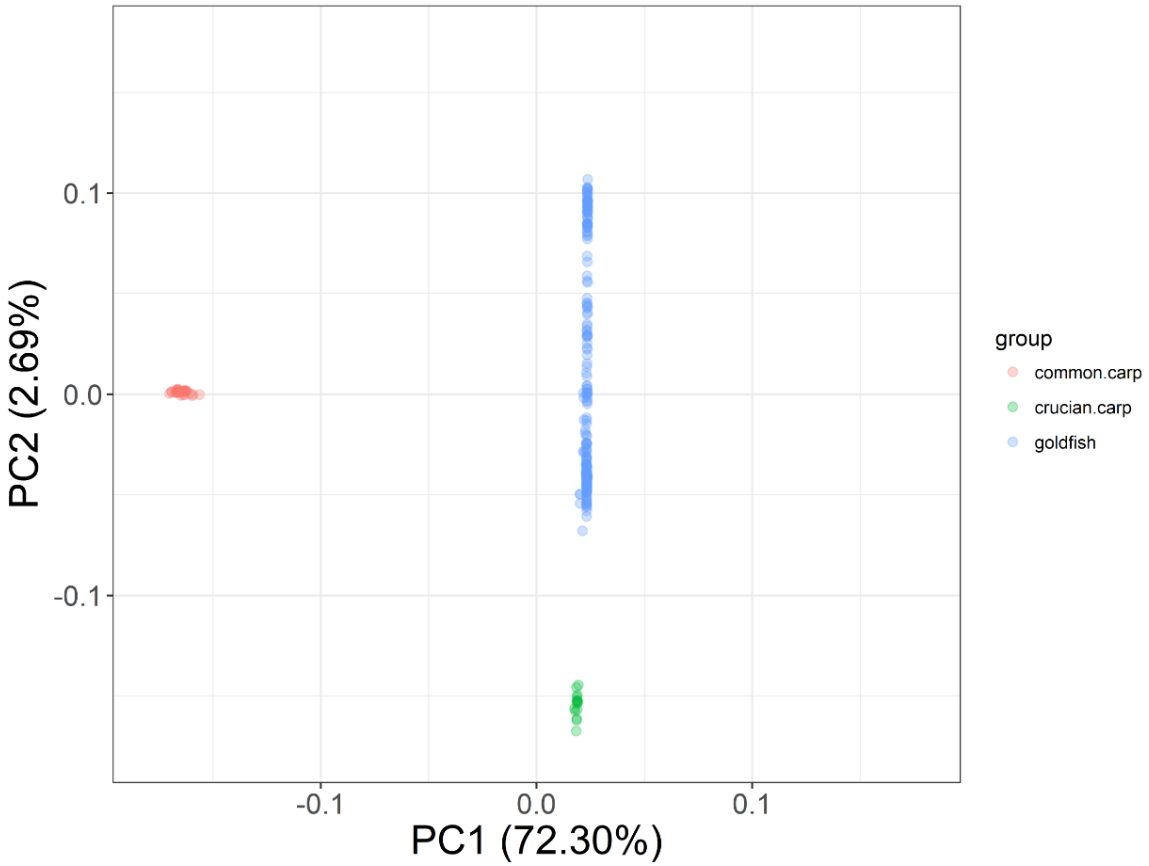


Fig. S27. 1-IBS genetic distance-based phylogeny tree constructed by fastME and visualized by FigsTree. (a) Identity-By-State (IBS) value calculated by plink --ibs for all pairs of 234 samples (33 *Cyprinus carpio*, 16 crucian carps, and 185 goldfishes). (b) PCA dotplot with the same three groups (33 *Cyprinus carpio*, 16 crucian carps, and 185 goldfishes).

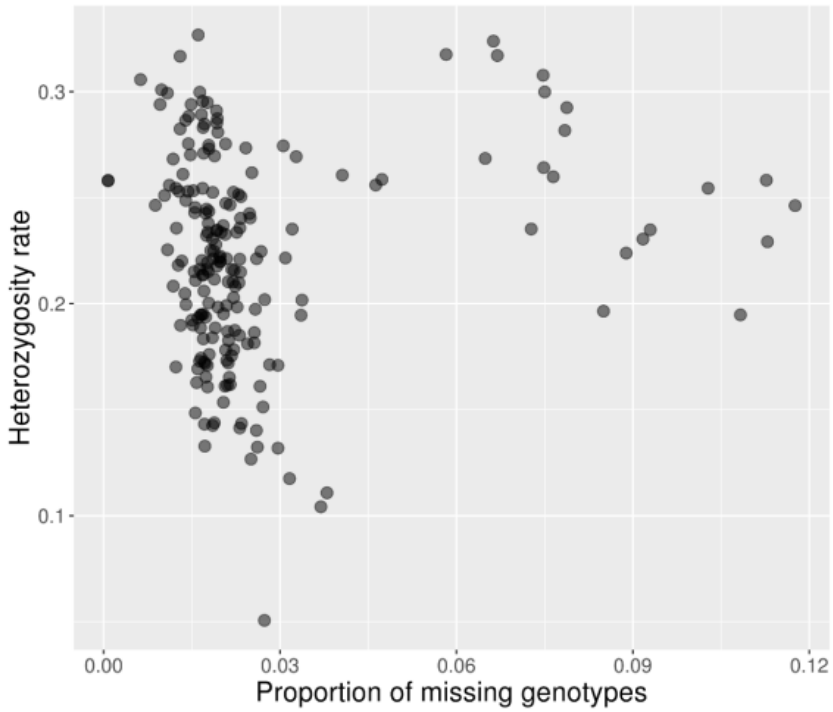


Fig. S28. Relationship between the heterozygosity rate and the proportion of missing genotypes in 201 goldfish populations.

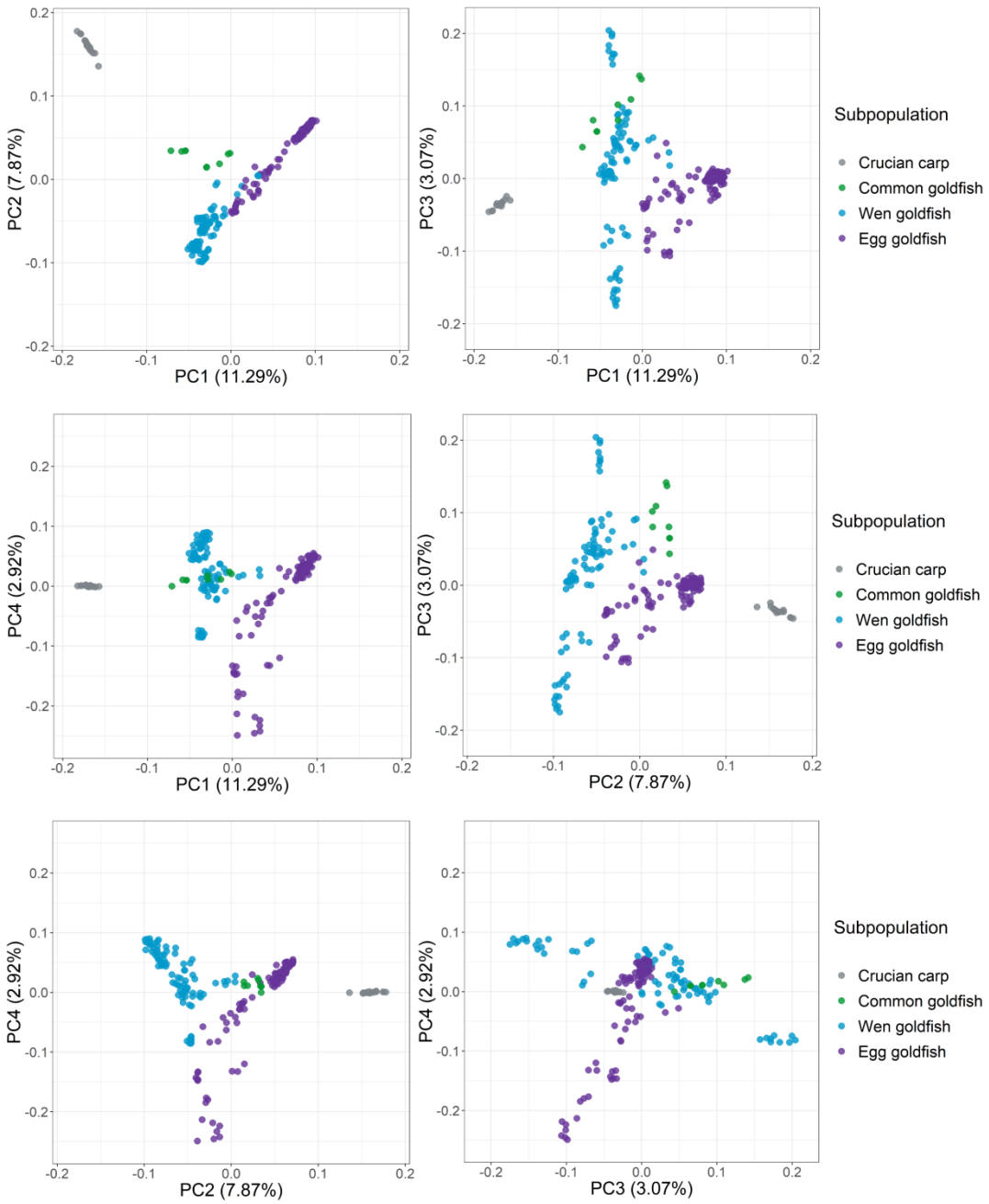


Fig. S29. Principal component SNP variation analysis for four subpopulations within 185 goldfish and 16 crucian carp samples. Samples from the subpopulations of crucian carp (CC), common goldfish (CG), Wen goldfish (WG), and Egg goldfish (EG) are shown. The plots show the first two principal components.

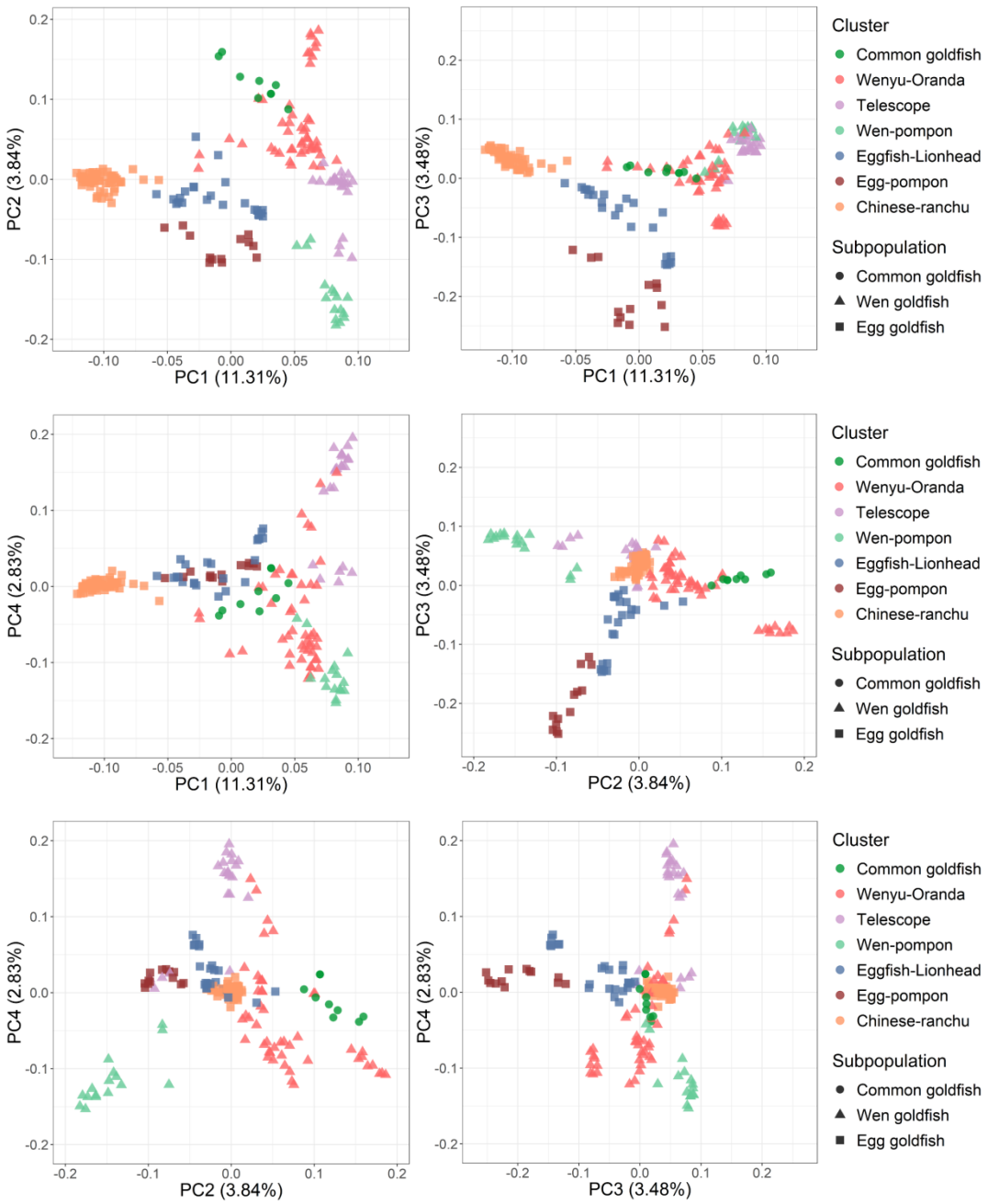
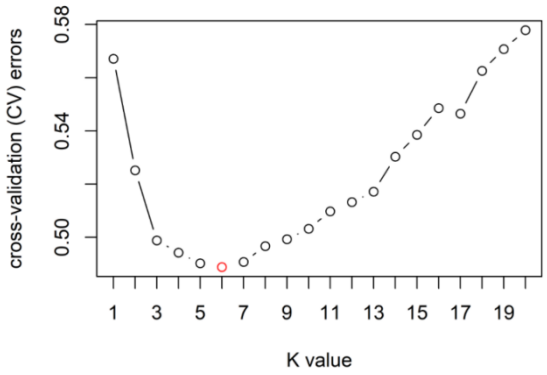


Fig. S30. Principal component SNP variation analysis for seven clusters within the 185 goldfish samples. Samples from the cluster of common goldfish, Wenyu-Oranda, telescope, Wen-pompon, Eggfish-Lionhead, Egg-pompon, and Chinese-ranchu are shown. The plots show the first two principal components.

A



B

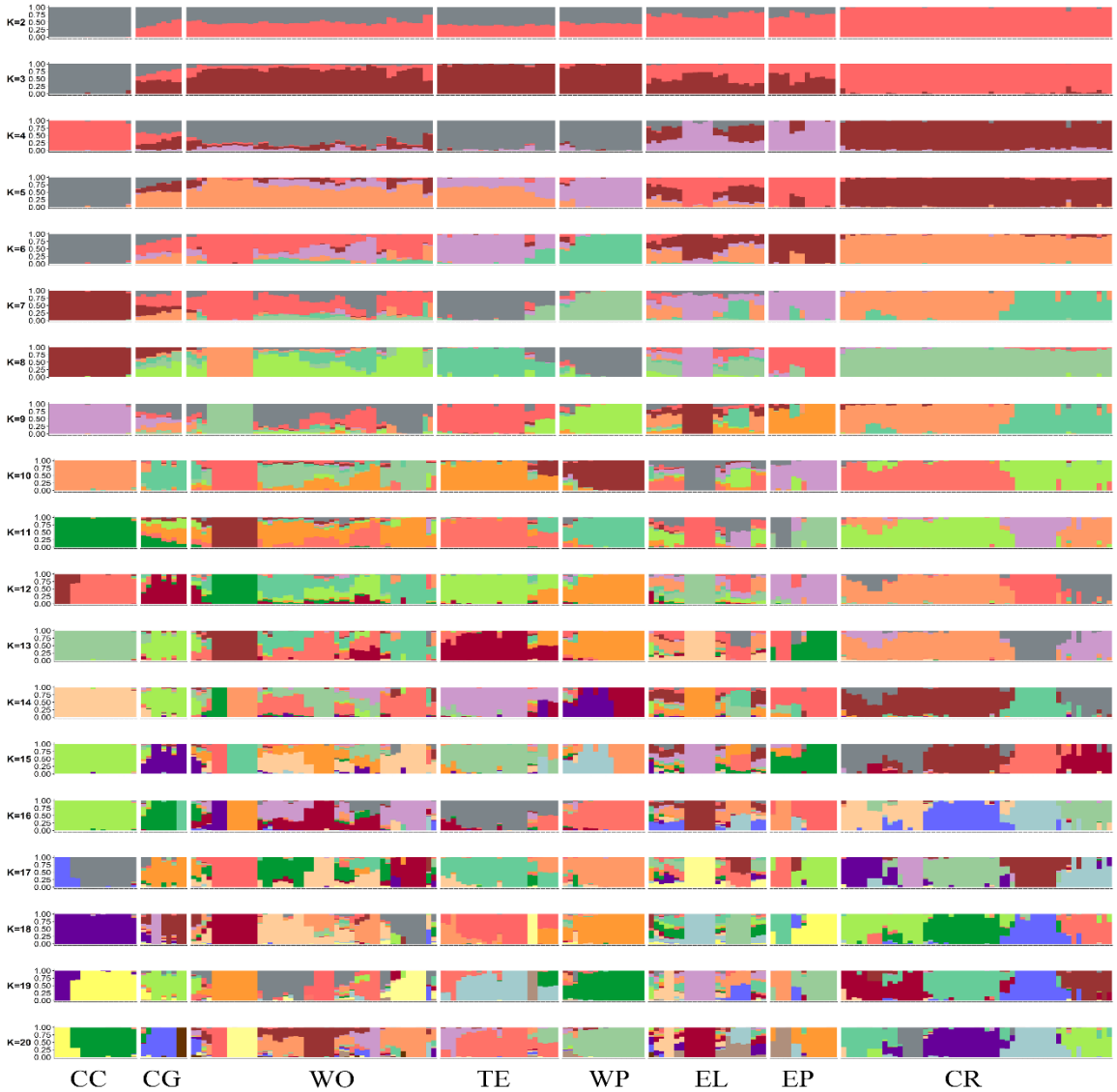


Fig. S31. Admixture results from K=2 to K=20. (a) Cross-validation errors showing K=6 is the optimal population cluster grouping. (b) Population from K=2 to K=20 showing the population structure. The number of groups (K value) is represented by different color codes, for instance K=6 can be divided into six groups by different colors as follows: group 1 of CC,

group 2 of CG and WO, group 3 of TE, group 4 of WP, group 5 of EL and EP, and group 6 of CR. The consanguinity of the WO species is closest to that of the common goldfish, indicating that the genes of the WO species are shared with the CG. CC, crucian carp; CG, common goldfish; WO, Wenyu-Oranda; TE, Telescope; WP, Wen-pompon; EL, Eggfish-lionhead; EP, Egg-pompon; CR, Chinese-ranchu.

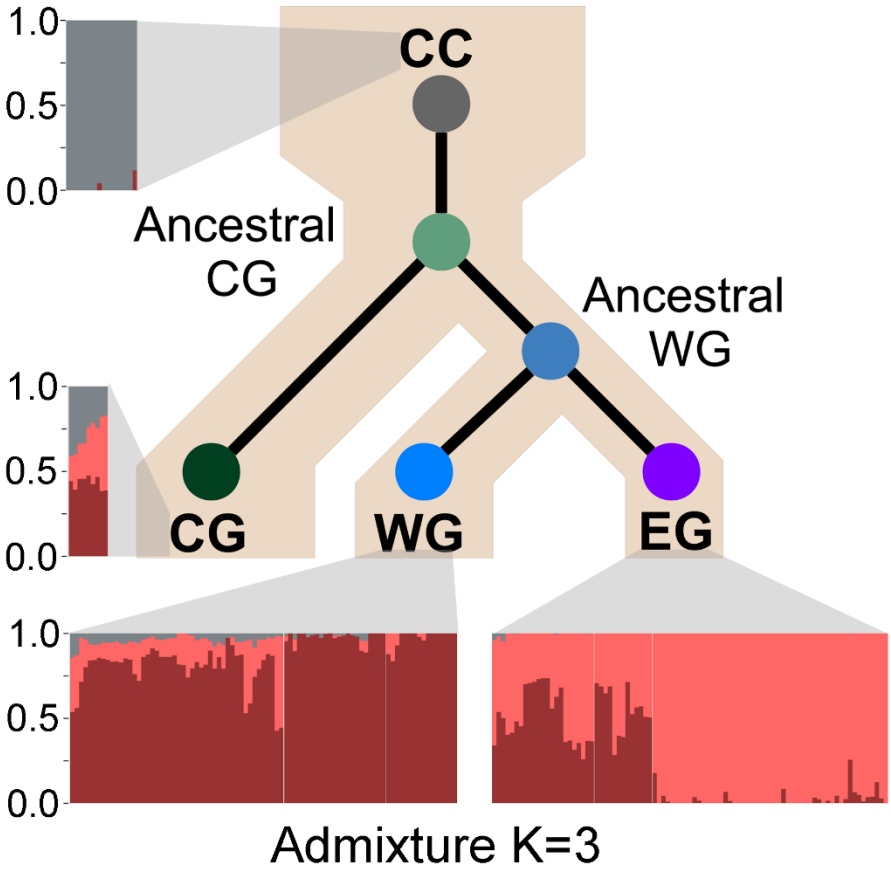


Fig. S32. Schematic diagram of the variation in the goldfish domestication history. Population from K=3 to show the population structure, CC, crucian carp; CG, common goldfish; WG, wen-goldfish; and EG, egg-goldfish. The number of groups (K value) is represented by different color codes; for instance, K=3 can be divided into three groups by different colors as follows: group 1 of CG, group 2 of WG, and group 3 of EG. The consanguinity of the WG species is the closest to that of the common goldfish, whereas EG is derived from WG by domestication and artificial selection, indicating that the genes of the WG species are shared with the EG species.

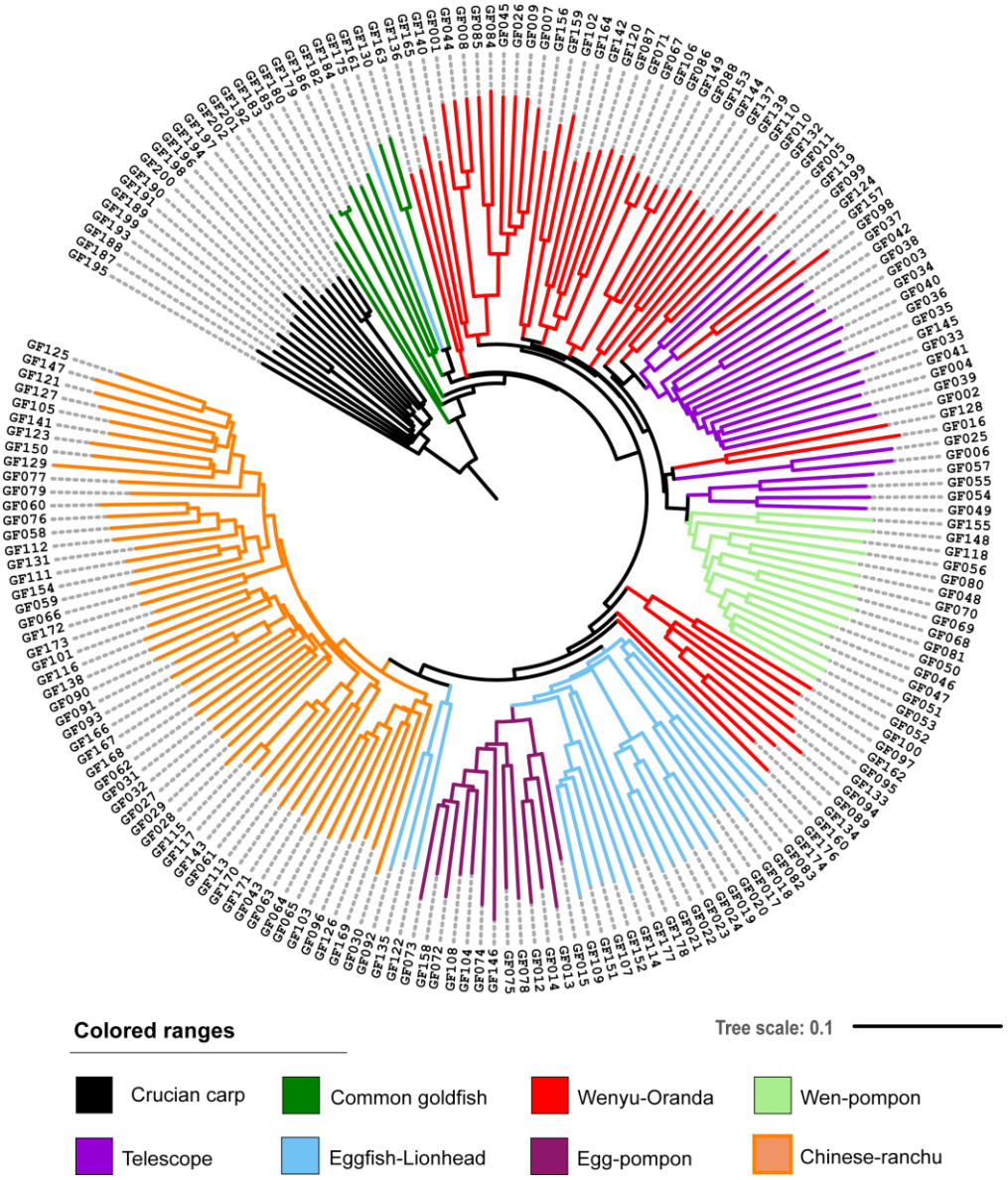


Fig. S33. Population phylogenetic relationship among 201 goldfish accessions. Binary search tree of 201 goldfish accessions (including 16 crucian carps) based on genetic distance. Color bars indicate accessions with different variation characteristics. The scale bar shows substitutions per site. The phylogenetic tree showed that crucian carp and goldfish were divided into two groups, Egg-goldfish and Wen-goldfish.

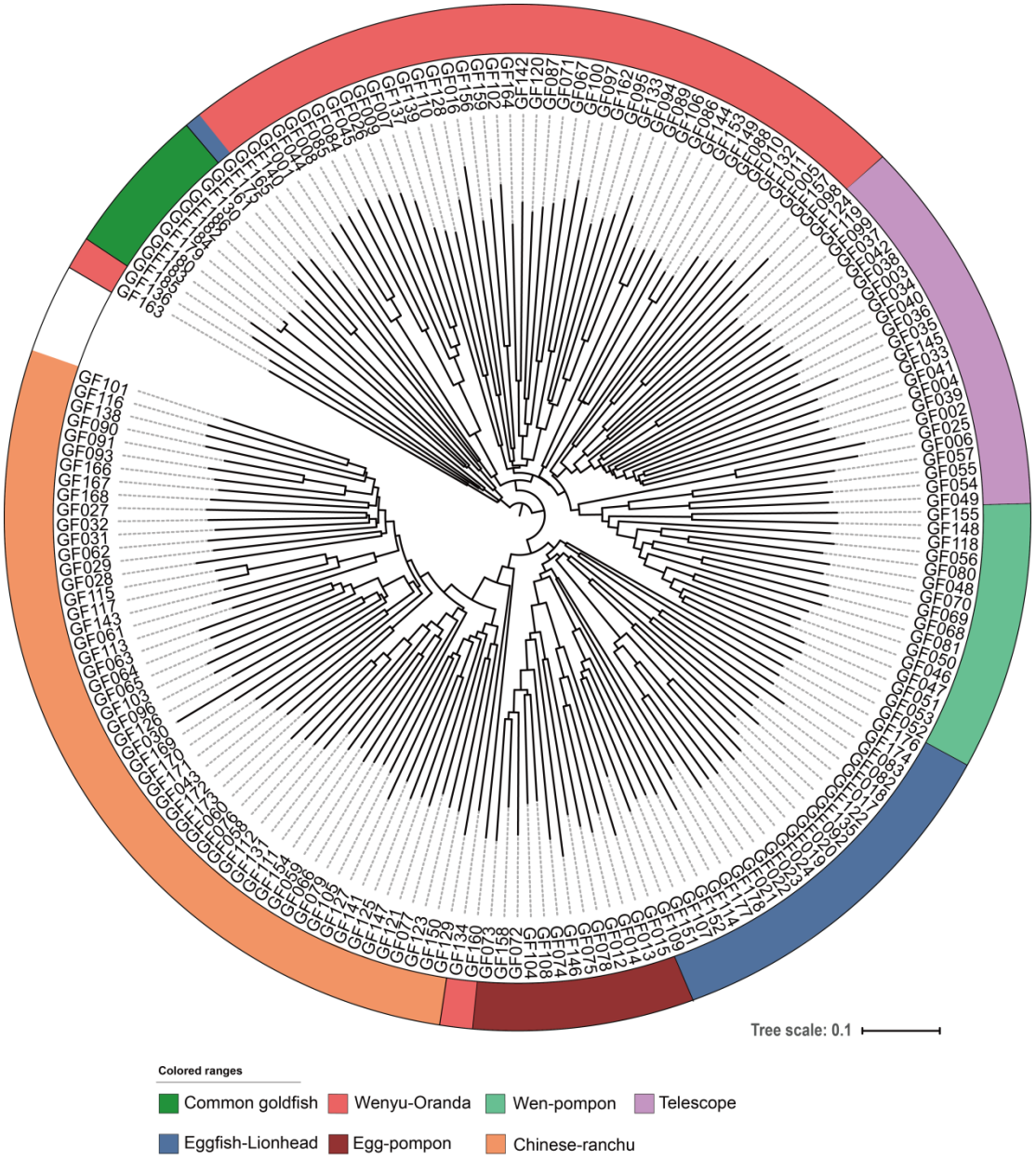


Fig. S34. Population phylogenetic relationship among 185 goldfish accessions. Binary search tree of 185 goldfish accessions (except for 16 crucian carps) based on genetic distance. Color bars indicate accessions with different variation characteristics. The scale bar shows substitutions per site. Among the 185 goldfish, common goldfish and Chinese-ranchu were significantly independent, whereas Egg-goldfish and Wen-goldfish were divided into two groups. The results indicate that the genetic relationship between the varieties telescope, Wen-pompon, and Egg-pompon was relatively close.

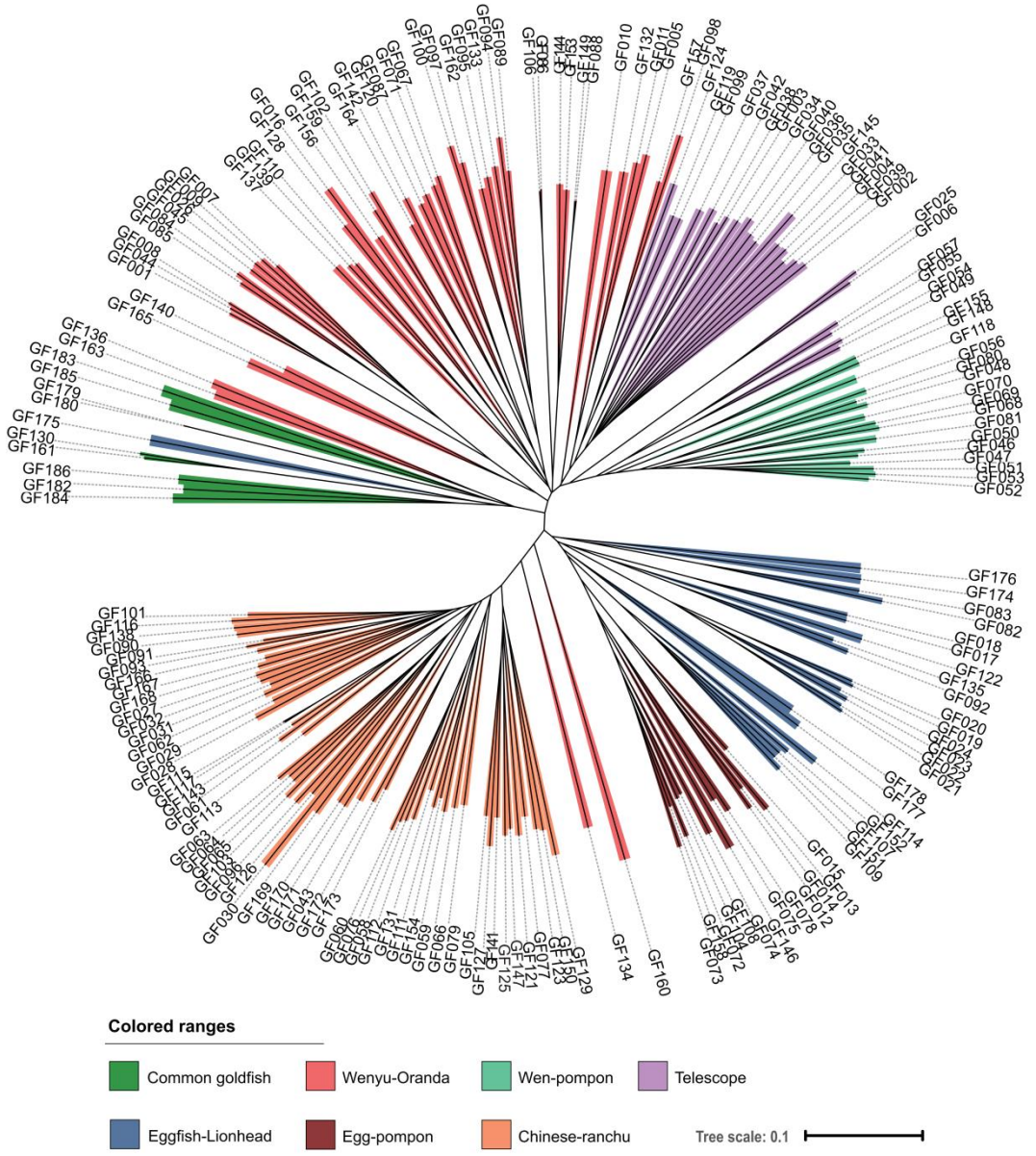


Fig. S35. Population phylogenetic relationship among 185 goldfish accessions. Neighbor-joining clustering of common goldfish, Wenyu-Oranda, telescope, Wen-pompon, Eggfish-Lionhead, Egg-pompon, and Chinese-ranchu based on genetic distance. Branch color indicates membership in one of eight classified goldfish populations. The scale bar shows substitutions per site. Among the 185 goldfish accessions, common goldfish and Chines-ranchu were significantly independent, whereas Egg-goldfish and Wen-goldfish were divided into two groups. The result indicated that the genetic relationship between the varieties teletescope, Wen-pompon, and Egg-pompon was relatively close.

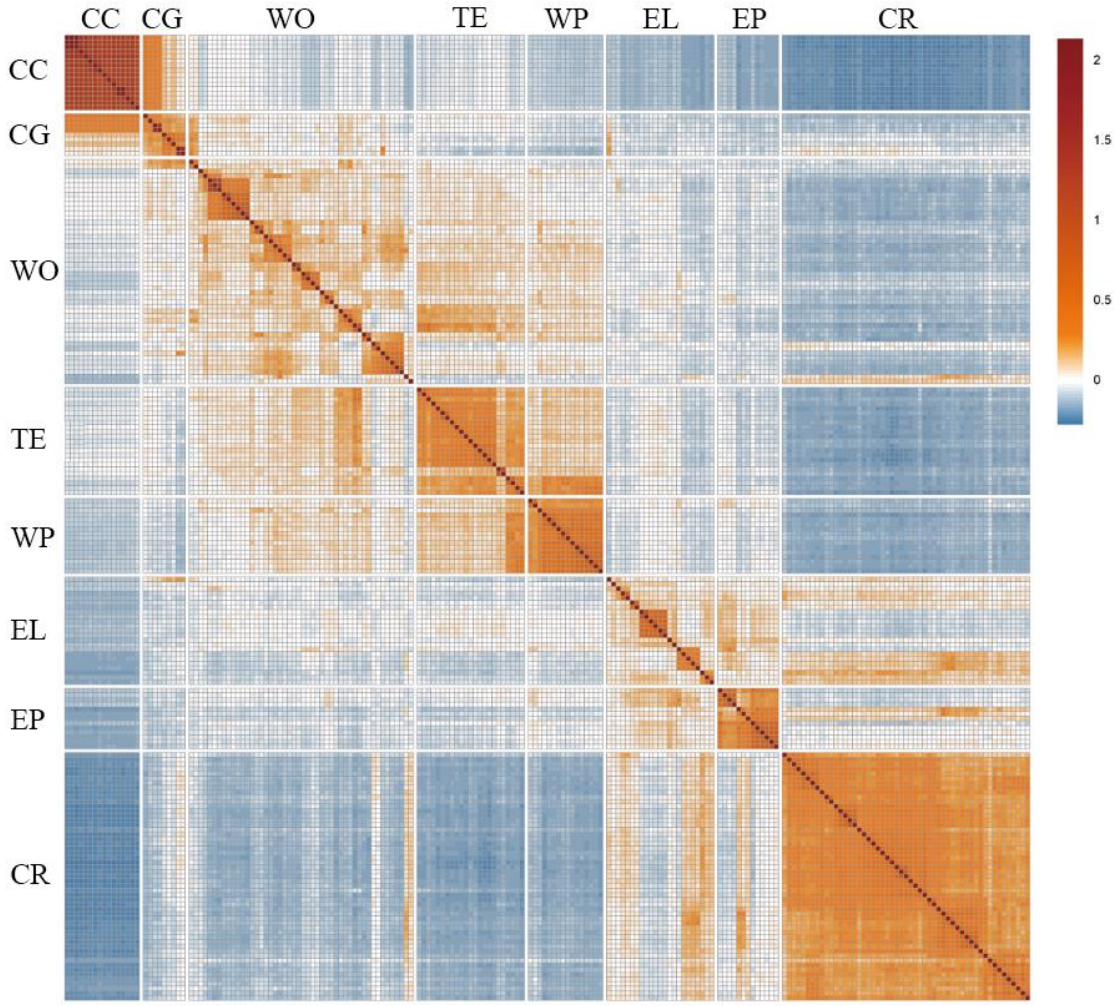


Fig. S36. Heatmap of pairwise genetic relatedness score matrix of genotyped individuals of curcian carps and goldfishes. The scale unit is the genetic relatedness score. CC, crucian carp; CG, common goldfish; WO, Wenyu-Oranda; TE, telescope; WP, Wen-pompon; EL, Eggfish-Lionhead; EP, Egg-pompon; and CR, Chinese-ranchu.

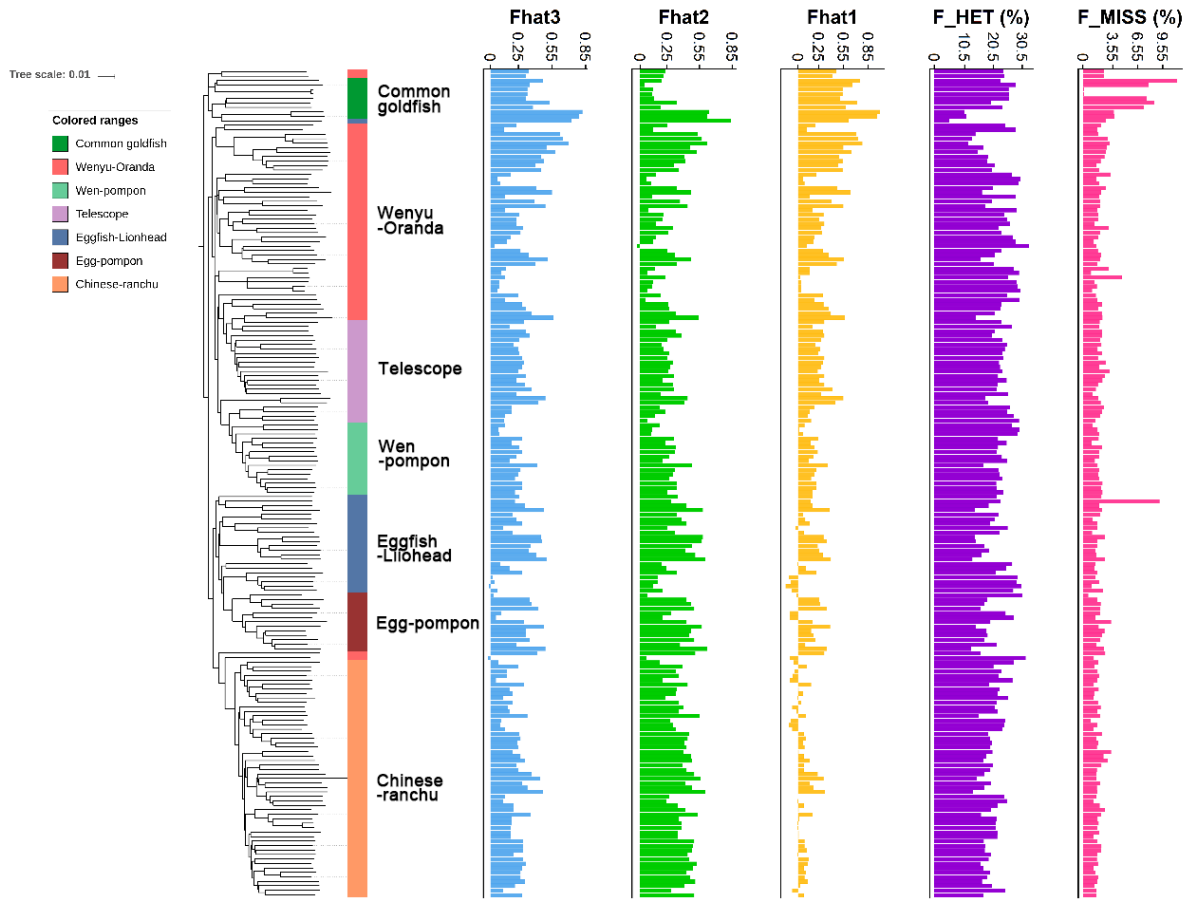


Fig. S37. Five genetic features of individuals across the phylogenetic tree of 185 goldfish. Fhat1~Fhat3, three inbreeding coefficients calculated by plink; F_HET (%), proportion of heterozygous genotypes; and F_MISS (%), proportion of missing genotypes. As shown in the figure, the inbreeding number of each sample was analyzed. There was no error in the heterozygous and gene deletion ratio of the 185 goldfishes.

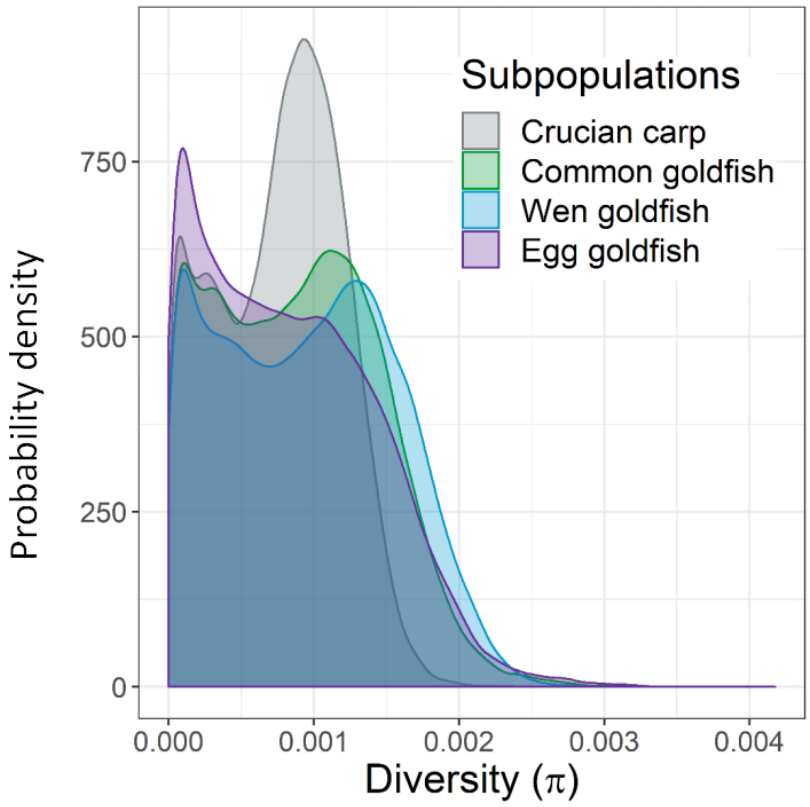


Fig. S38. Level of genetic diversity (π) in each subpopulation of goldfish and crucian carp. The Y-axis is the probability density of the distribution of the average *P* value in a nonoverlapping window of 100 kb.

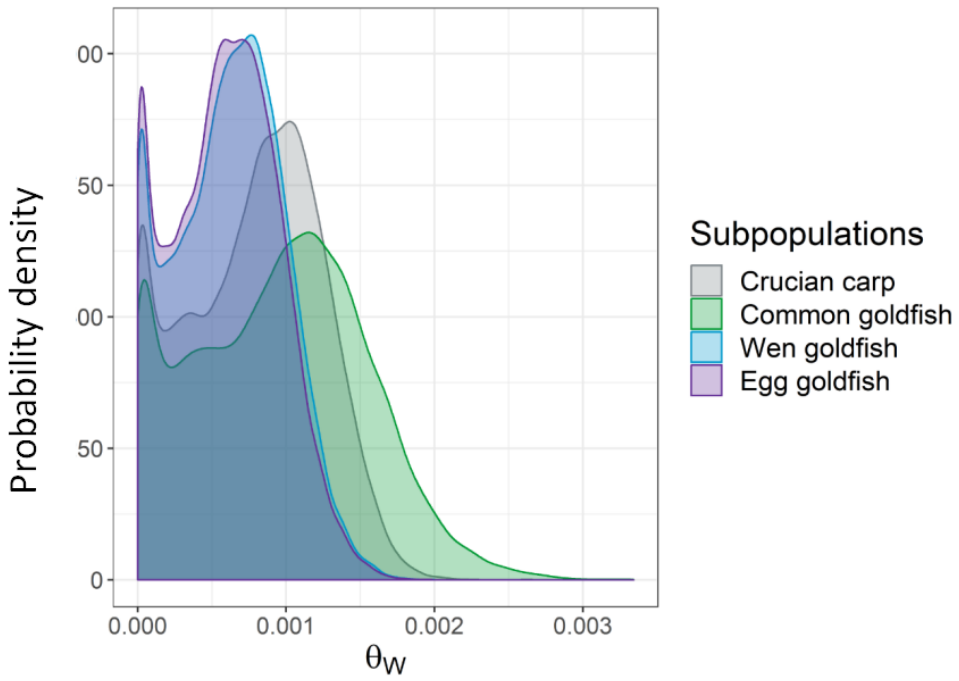


Fig. S39. Level of Watterson's theta(θw) in each subpopulation of goldfish and crucian carp. The Y-axis is the probability density of the distribution of the average *P* value in a nonoverlapping window of 100 kb.

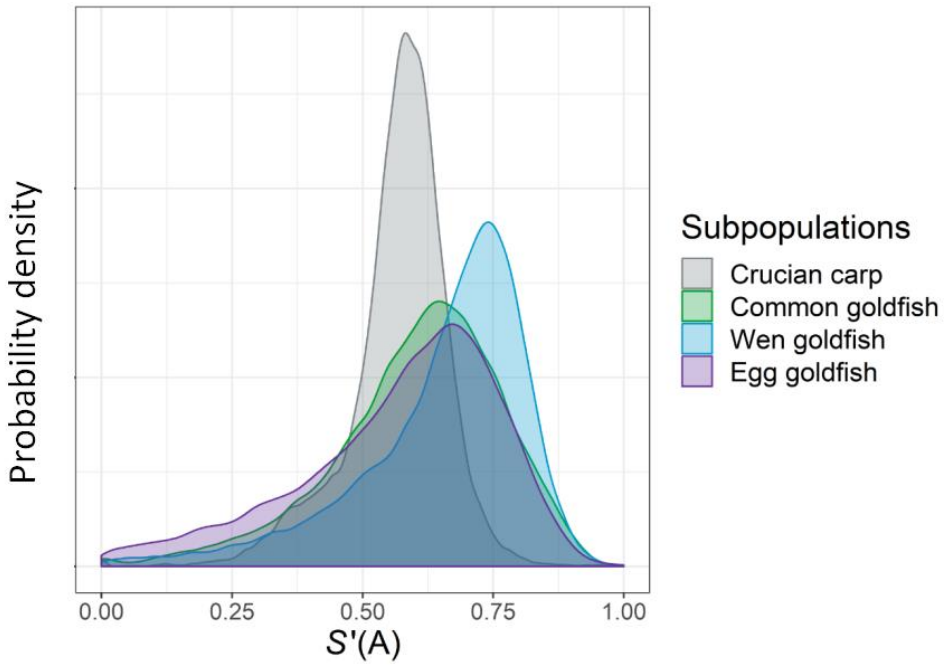


Fig. S40. Distributions of allele frequency-based relative Shannon entropy denoted by S'(A), calculated in the crucian carp subpopulation, common goldfish subpopulation, Wen-goldfish subpopulation, and Egg-goldfish subpopulation. The Y-axis is the probability density of the distribution of the average *P* value in a nonoverlapping window of 100 kb.

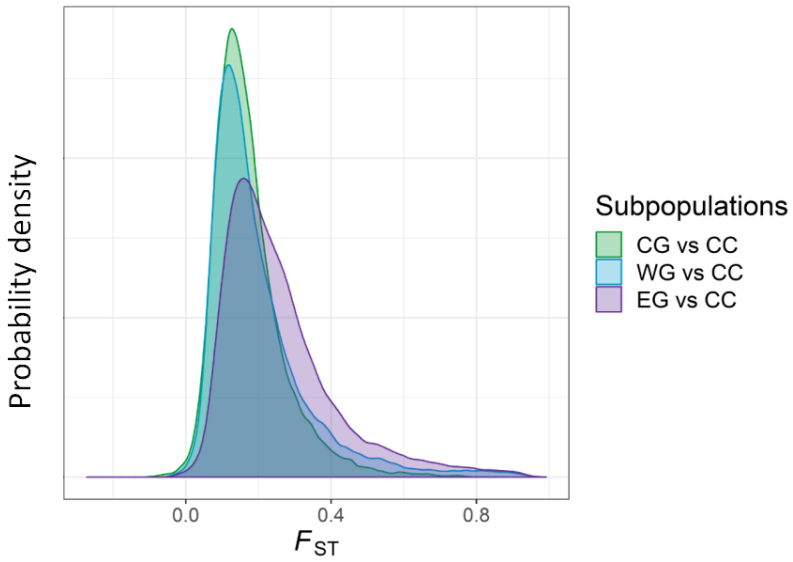


Fig. S41. Fst value distribution of goldfish subpopulations. The Y-axis is the probability density of the distribution of the average *P* value in a nonoverlapping window of 100 kb. CG, common goldfish; CC, crucian carp; WG, Wen-goldfish; and EG, Egg-goldfish.

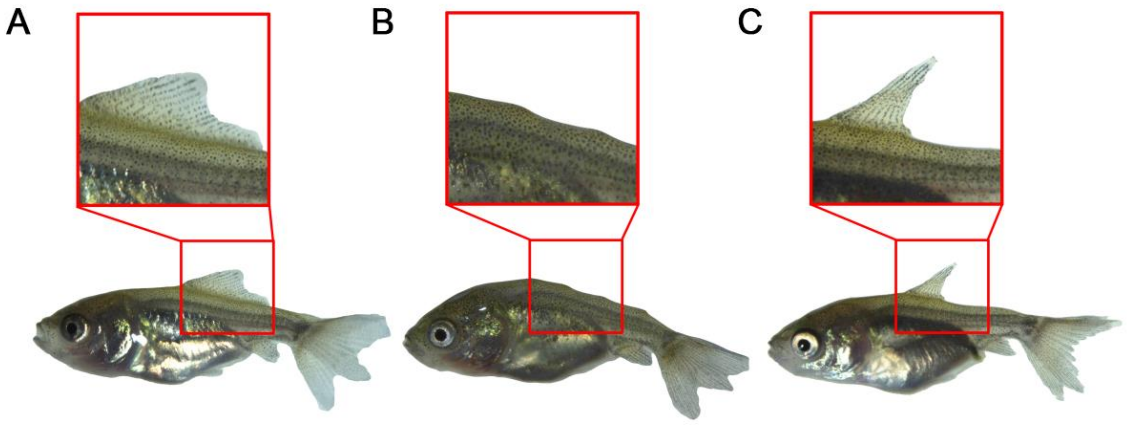


Fig. S42. Phenotypic variants of the dorsal fin in the F1 population. (A) Normal dorsal fin. (B) Lack of dorsal fin. (C) Abnormal dorsal fin. The sharp abnormal and normal dorsal fins are an acute and obtuse triangle, respectively.

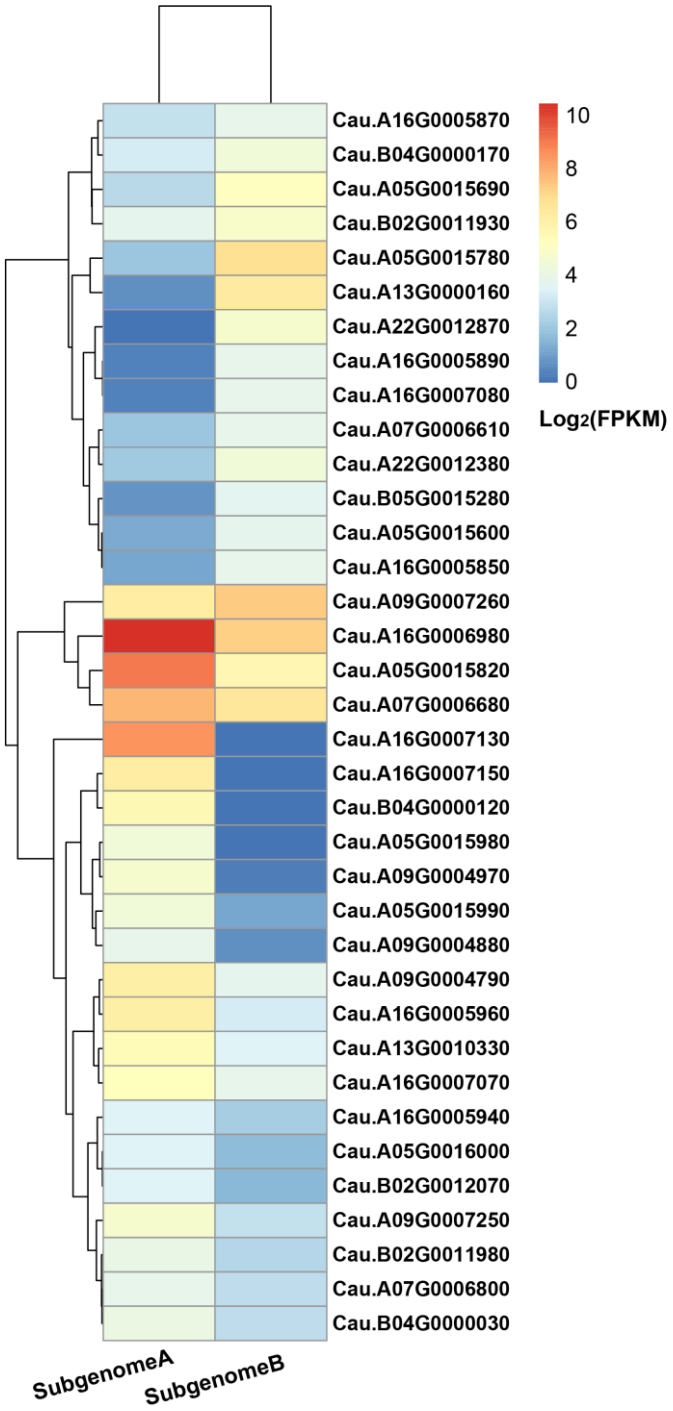


Fig. S43. Primary of dorsal fin GWAS associated genes dominant expressed in subgenome A. Of the 378 dominant expressed genes in subgenomes A or B, 36 genes were clustered by base expression.

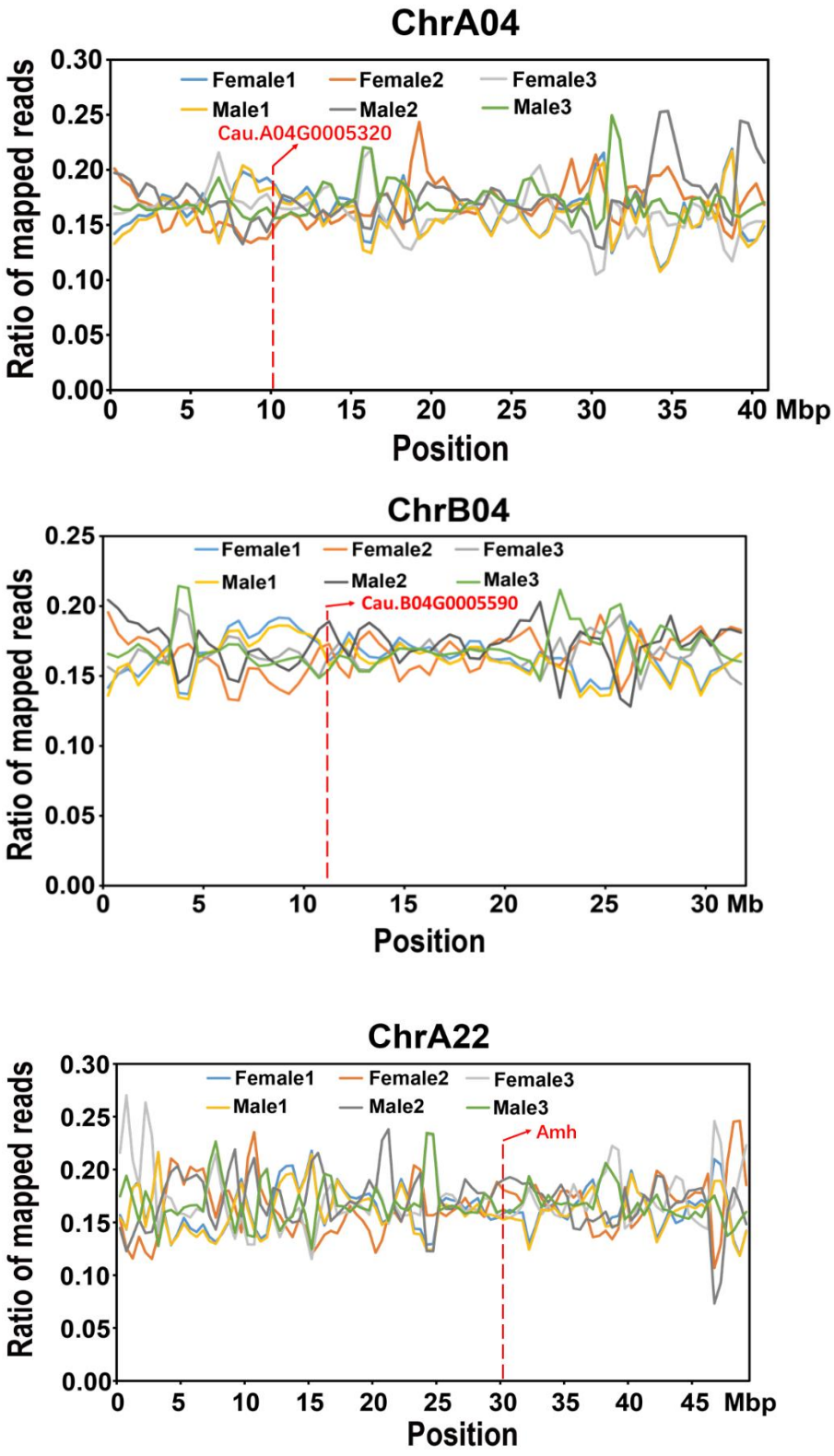


Fig. S44 Distribution of reads from three females and three males mapped against the goldfish ChrA04, ChrB04 and ChrA22. The X-axis and Y-axis indicate chromosome position (Mb) and ratio of mapped reads on the goldfish chromosomes, respectively.

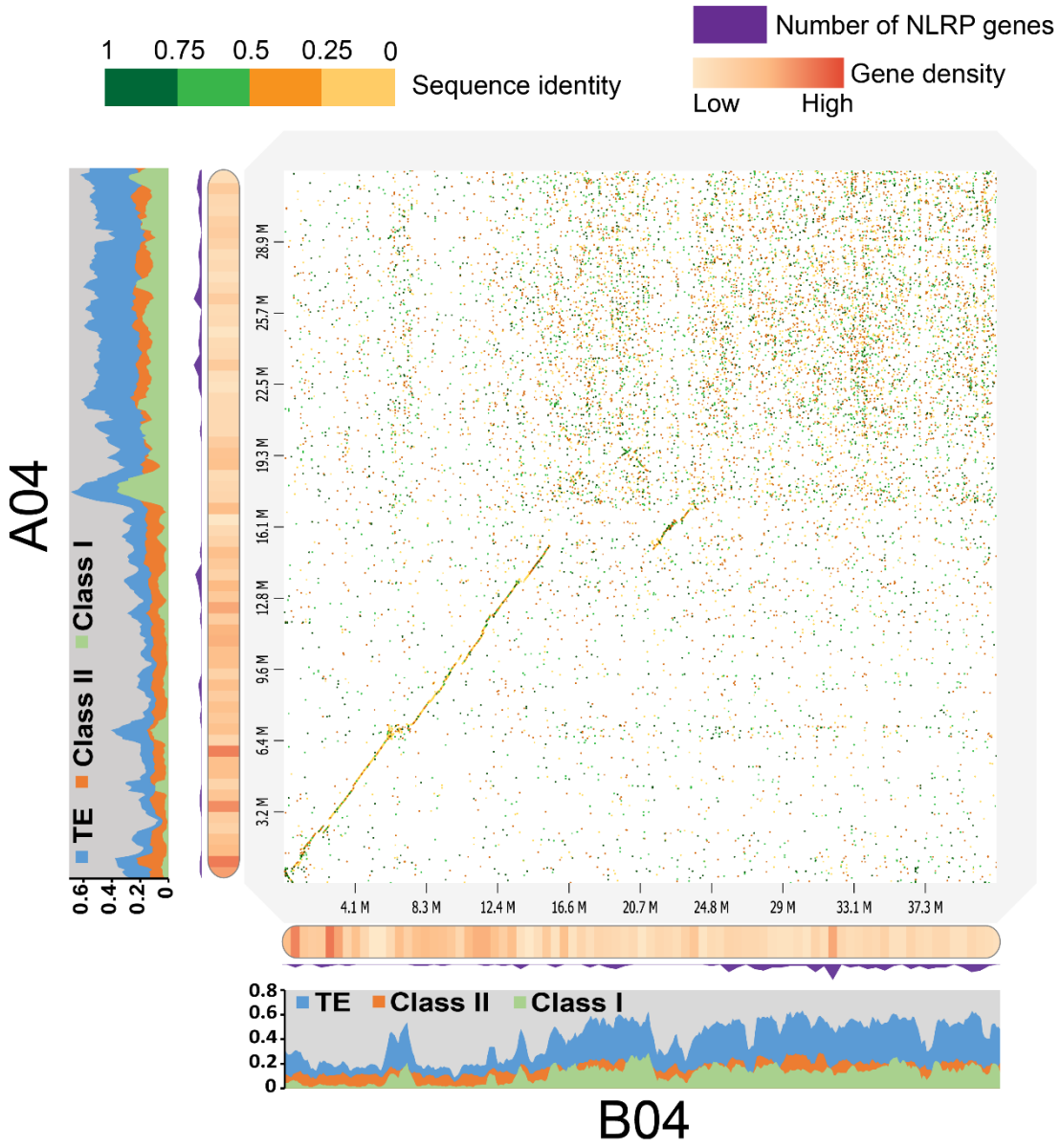


Fig. S45. Structural overview and comparison of chromosomal pseudomolecules ChrA04 and ChrB04**.** The distributions of transposable elements (TEs) are represented as stacked areas, and the density of innate immunity-related genes is highlighted with purple peaks within the corresponding regions. In the dotplot comparison, note the genomic character of the nonsynteny region, indicating that the highly divergent region displayed pooled collinearity and abundant TE and innate immunity-related gene (NLRP) content. The gene content is represented with a heatmap.

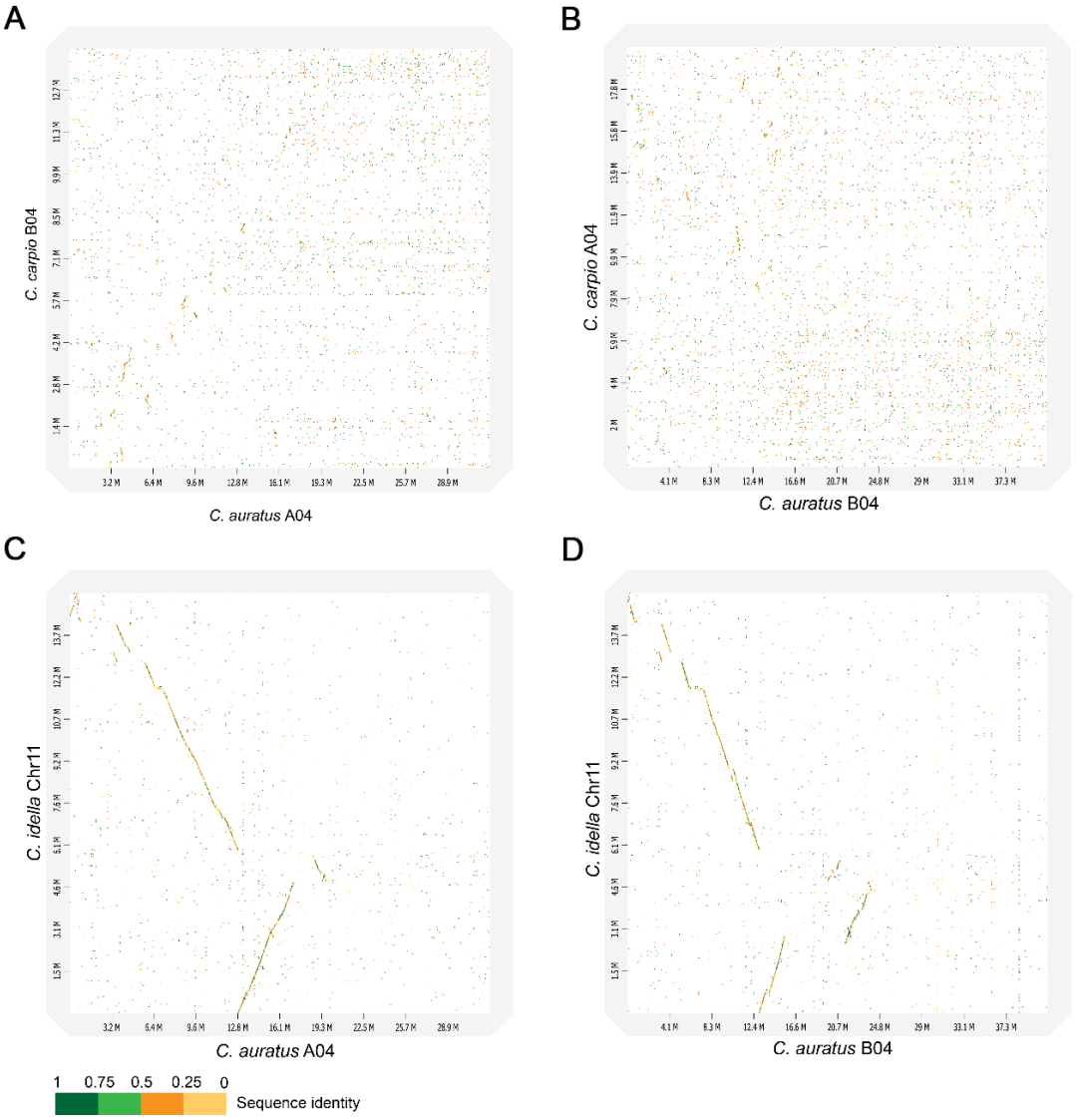


Fig. S46. Structural overview and comparison of chromosomal pseudomolecules ChrA04 and ChrB04 of *C. auratus* with the homologous chromosomes of *C. carpio* and *C. idella*. (A and C) Collinearity between ChrA04 of *C. auratus* with its homolog chromosome Chr7 in *C. carpio* and Chr11 in *C. idella*. (B and D) Collinearity between ChrB04 of *C. auratus* with its homolog chromosome Chr8 in *C. carpio* and Chr11 in *C. idella*.

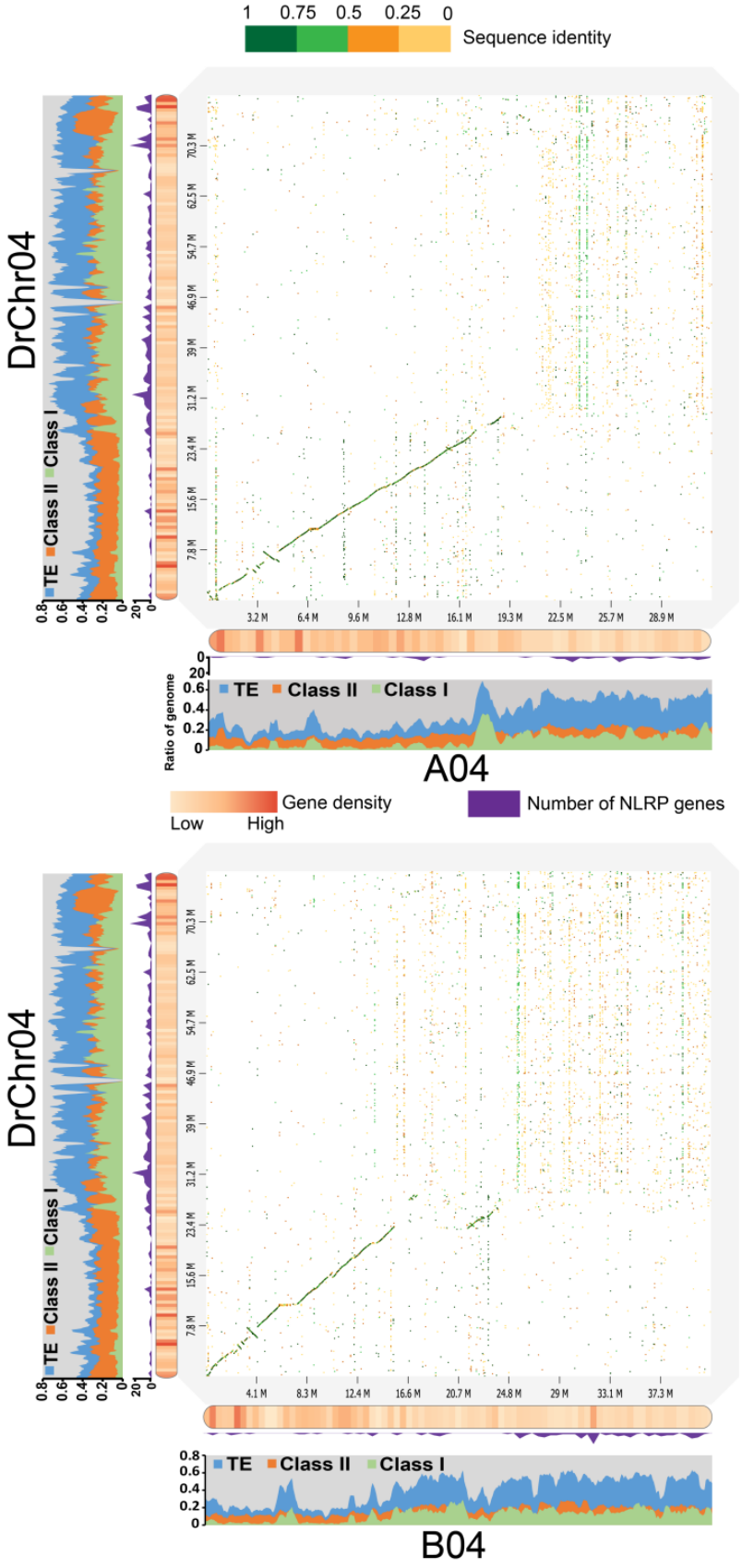


Fig. S47. Structural overview and comparison of chromosomal pseudomolecules ChrA04 and ChrB04 of *C. auratus* with the homologous chromosome Chr04 of *D.rerio.*

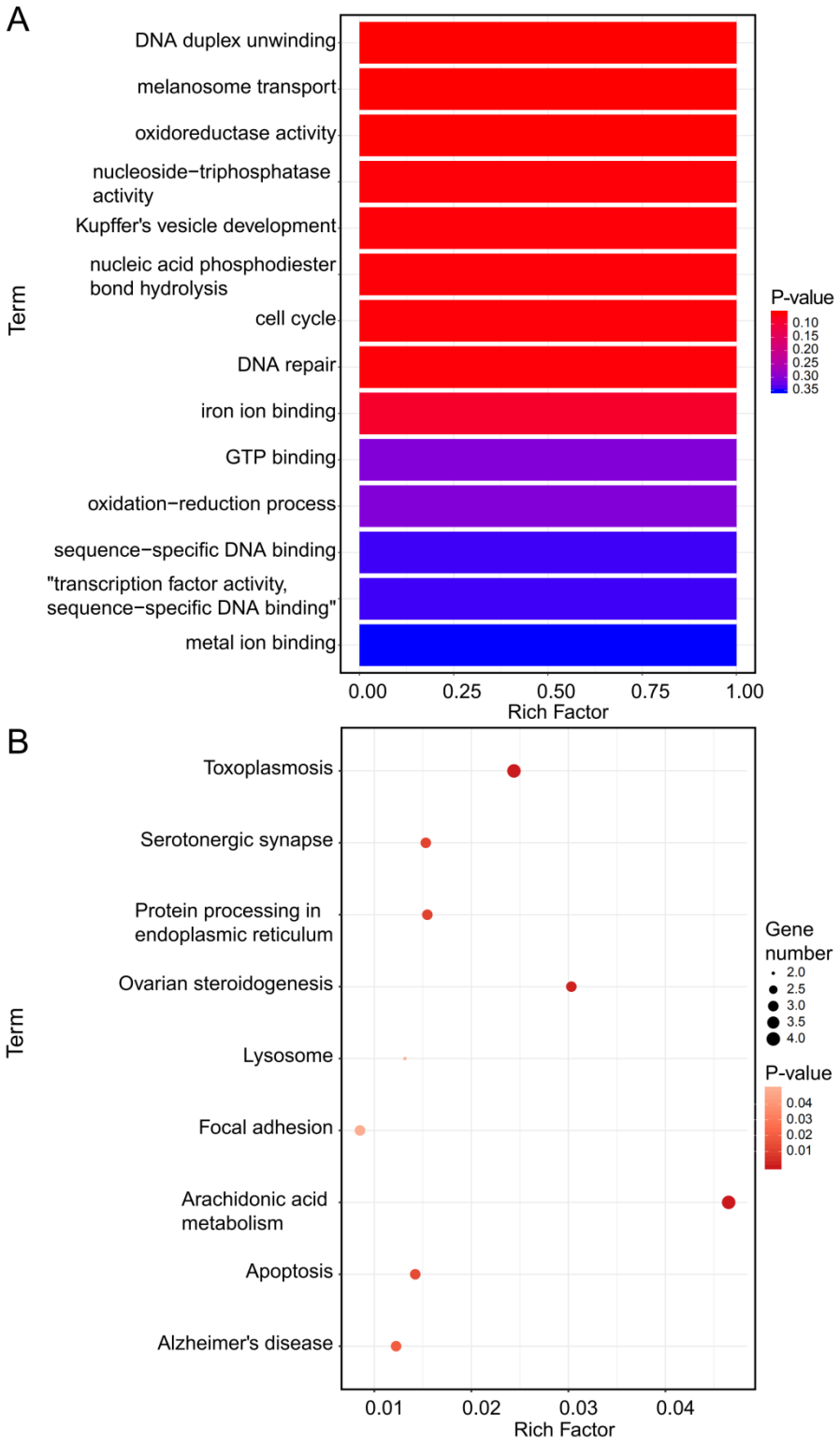


Fig. S48. GO term and KEGG pathway enrichment for the genes in the noncollinear region of ChrA04 in *C. auratus*. (A) GO term enrichment showing that the dominantly expressed homoeologs of Chr4 were primarily enriched in DNA duplex unwinding, oxidoreductase activity, etc. (B) KEGG pathway analysis indicating that dominantly expressed homoeologs of ChrA04 mainly participate in processes, such as protein processing, endoplasmic reticulum, and focal adhesion.

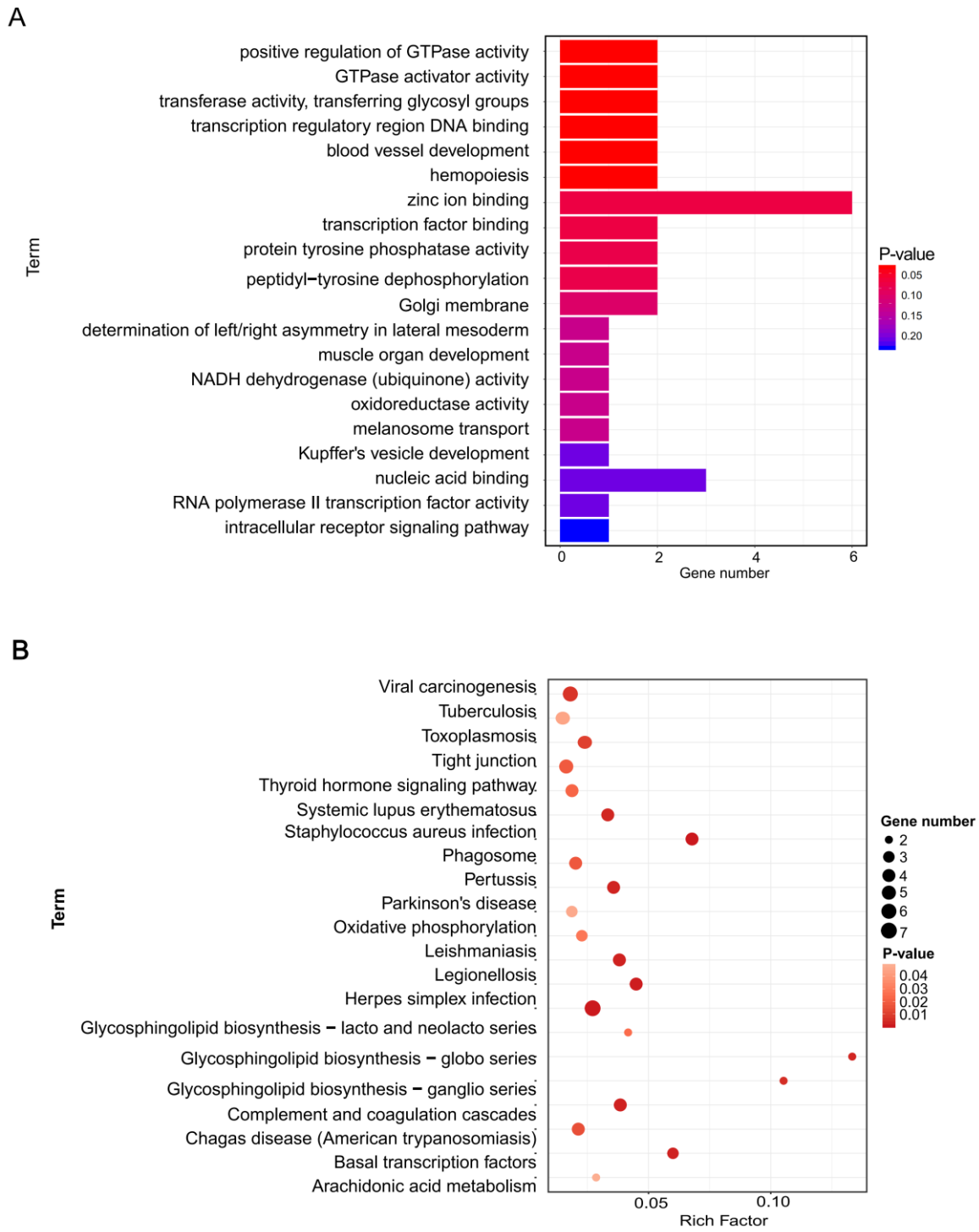


Fig. S49. GO term and KEGG pathway enrichment for the genes in the noncollinear region of ChrB04 in *C. auratus*. (A) GO term enrichment showing that the dominantly expressed homoeologs of ChrB04 were primarily enriched in factors, such as zinc ion binding, transcription factor binding, and GTPase activator activity. (B) KEGG pathway analysis indicating that the dominantly expressed homoeologs of ChrB04 mainly participate in processes such as tight junctions and oxidative phosphorylation.

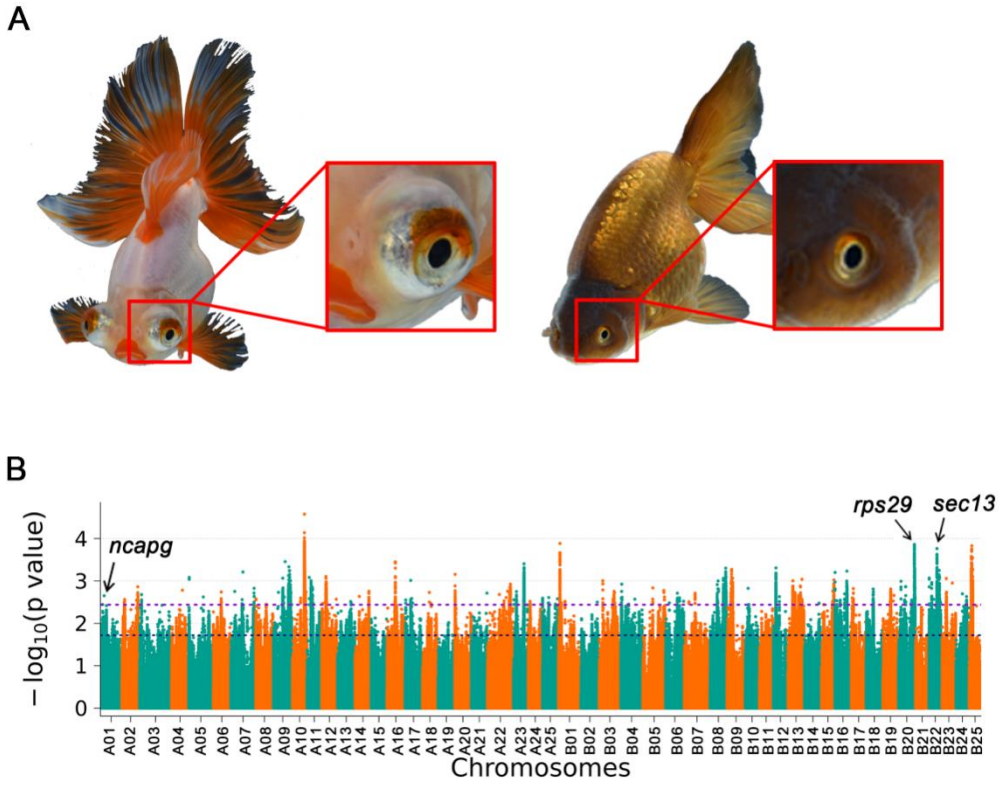


Fig. S50. Phenotypic variants of telescope eye in the goldfish population for GWAS. (Aa) Left, telescope eye: right, normal eye; (B) Manhattan plot for telescope eye GWAS in the goldfish population. Genes surrounding or covering peaks are indicated. The genome-wide threshold of 2.5 was defined to confirm selective sweeps by the top 1% of -$\log_{10}$ (*P* value) values.

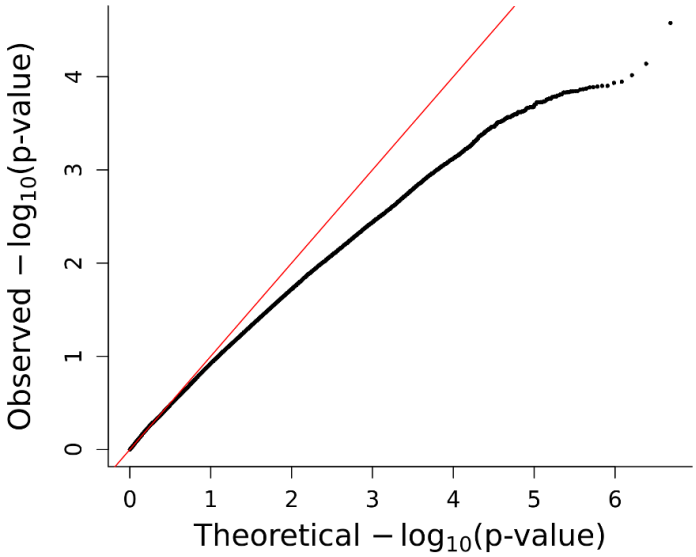


Fig. S51. Quantile-quantile plot for telescope eye GWAS under a general linear model.

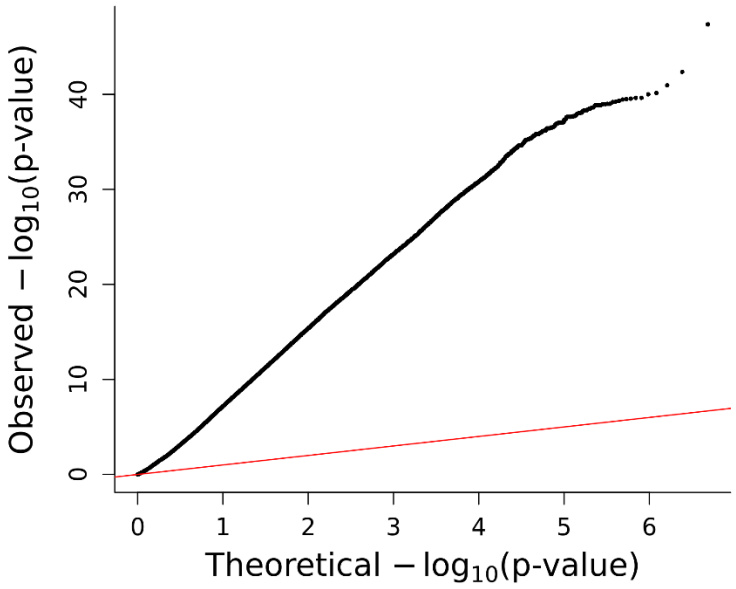


Fig. S52. Quantile-quantile plot for telescope eye GWAS under a general linear model.

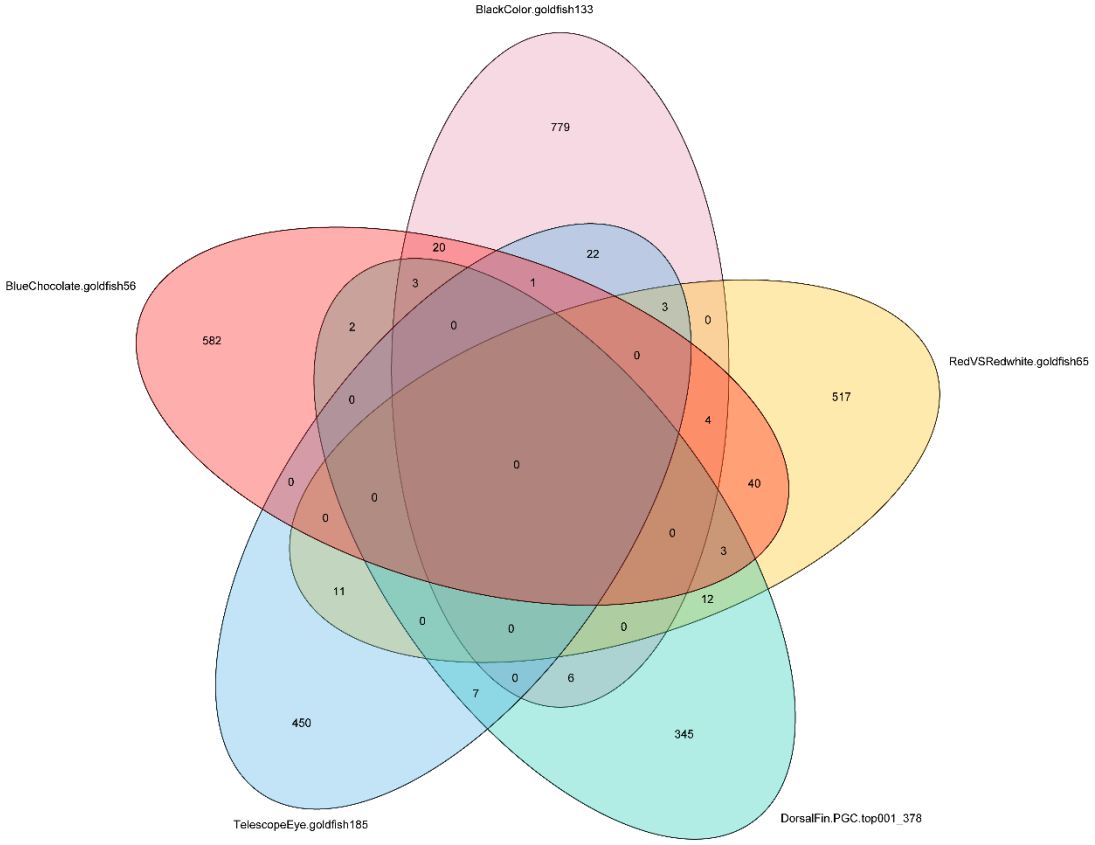


Fig. S53. Venn diagram of candidate genes from different phonotype GWAS.

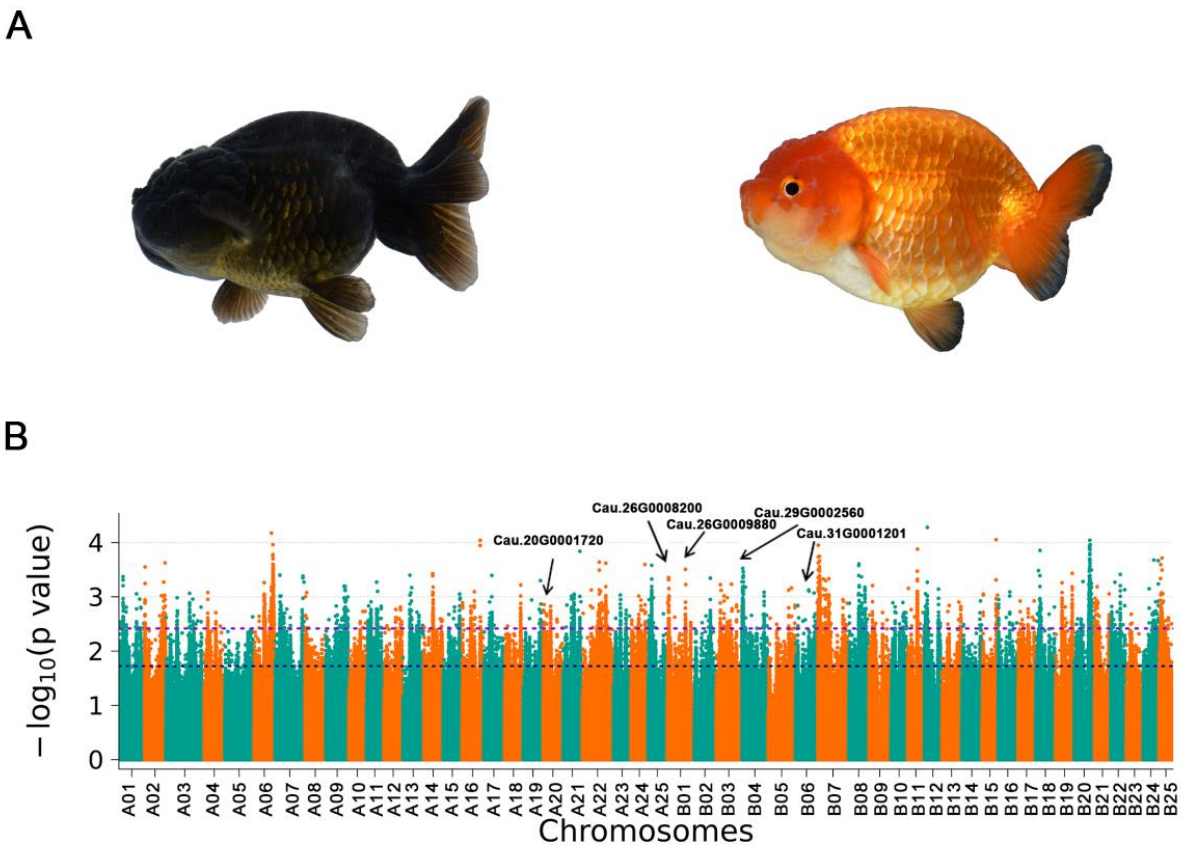


Fig. S54. Phenotypic variants of black-red color in goldfish population for GWAS. (A) Left, black morph: right, red morph; (B) Manhattan plot for black-red color GWAS in goldfish population. Genes surrounding or covering peaks are indicated. The genome-wide threshold of 2.5 was defined to confirm selective sweeps by the top 1% of -$\log_{10}$ (*P* value) values.

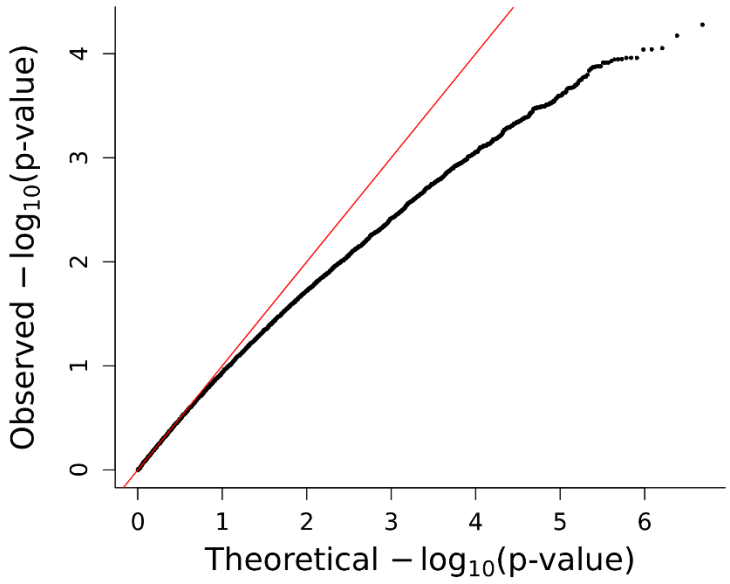


Fig. S55. Quantile-quantile plot for black-red color GWAS under a general linear model.

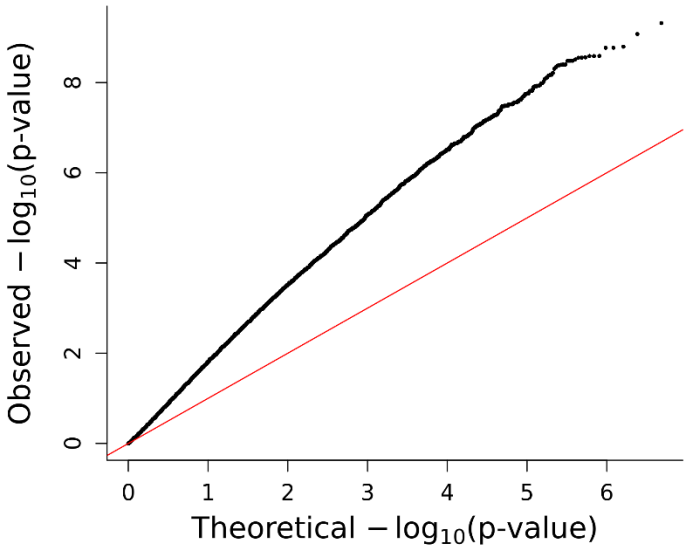


Fig. S56. Quantile-quantile plot for black-red color GWAS under a general linear model.

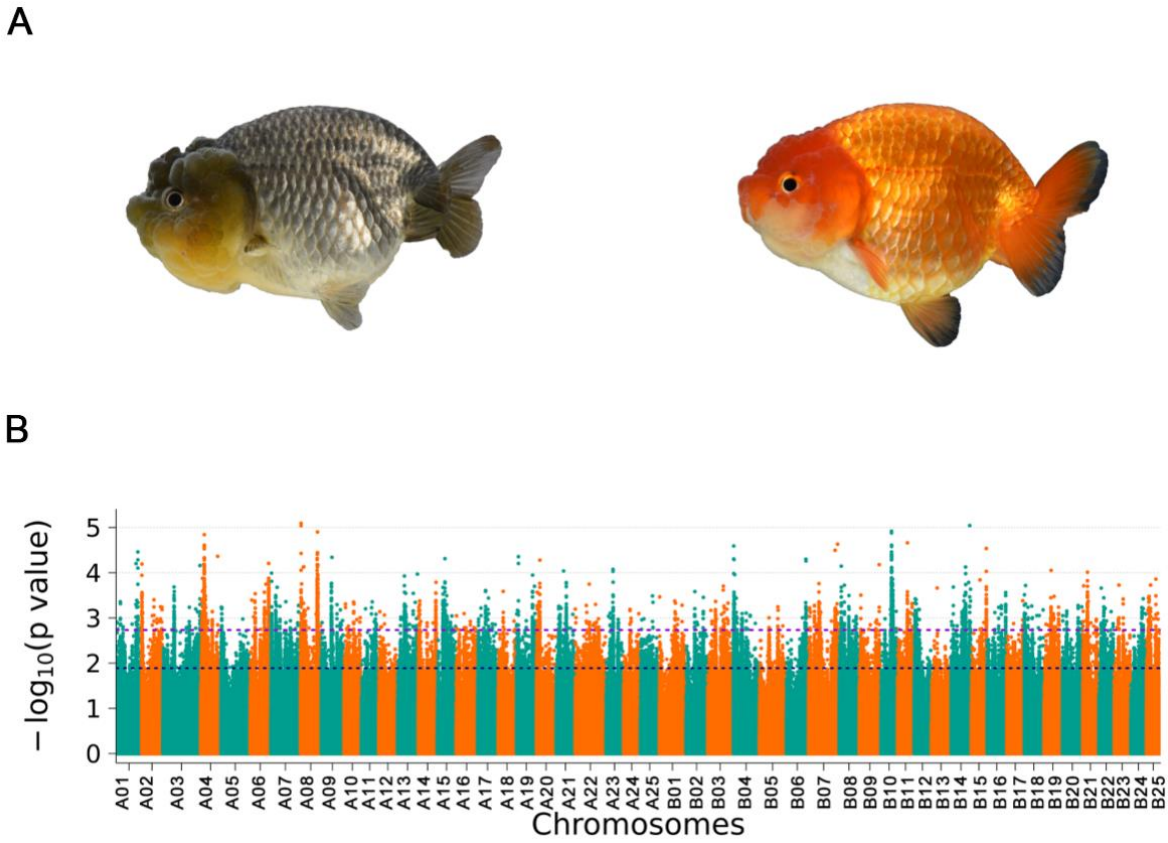


Fig. S57. Phenotypic variants of blue-red color in goldfish population. for GWAS. (A) Left, blue morph: right, red morph.; (B) Manhattan plot for blue-red color GWAS in goldfish population. Genes surrounding or covering peaks are indicated. The genome-wide threshold of 2.5 was defined to confirm selective sweeps by the top 1% of -$\log_{10}$ (*P* value) values.

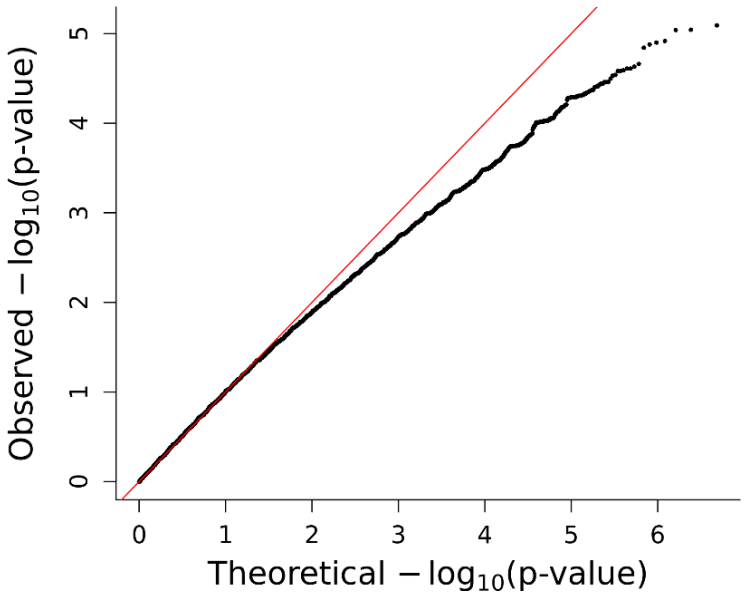


Fig. S58. Quantile-quantile plot for Blue-red color GWAS under a general linear model.

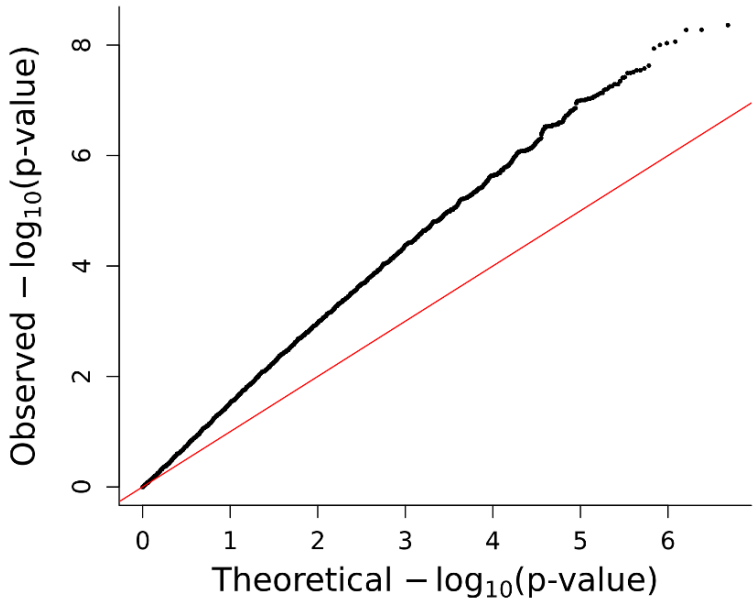


Fig. S59. Quantile-quantile plot for Blue-red color GWAS under a general linear model.

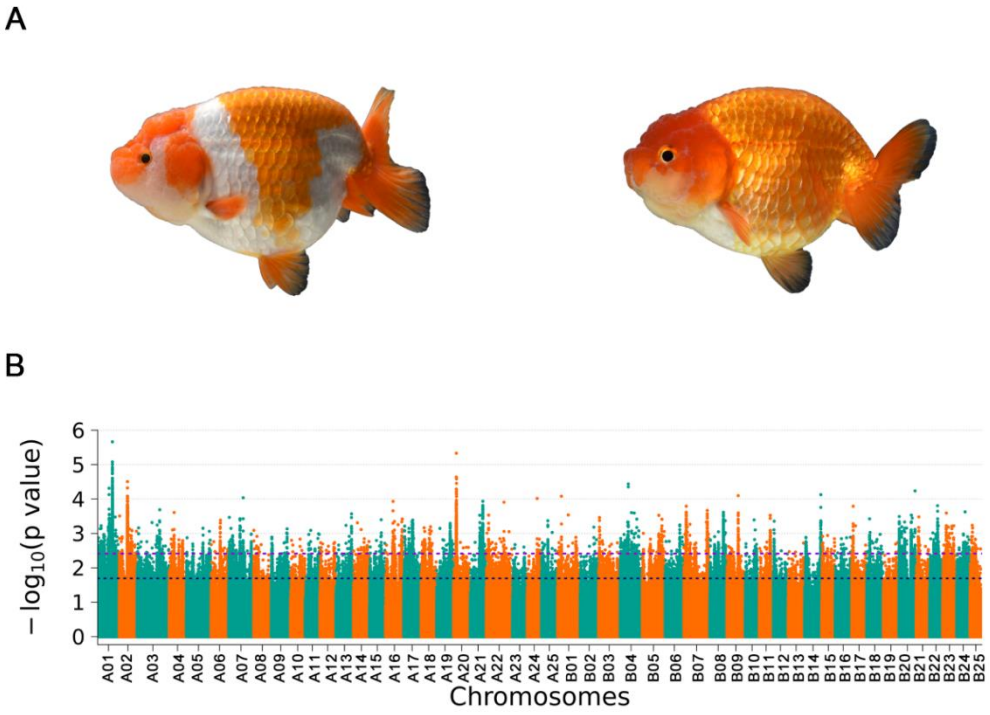


Fig. S60. Phenotypic variants of white-red color in goldfish population for GWAS. (A) Left, white morph.; right, red morph; (B) Manhattan plot for white-red color GWAS in goldfish population. Genes surrounding or covering peaks are indicated. The genome-wide threshold of 2.5 was defined to confirm selective sweeps by the top 1% of -$\log_{10}$ (*P* value) values.

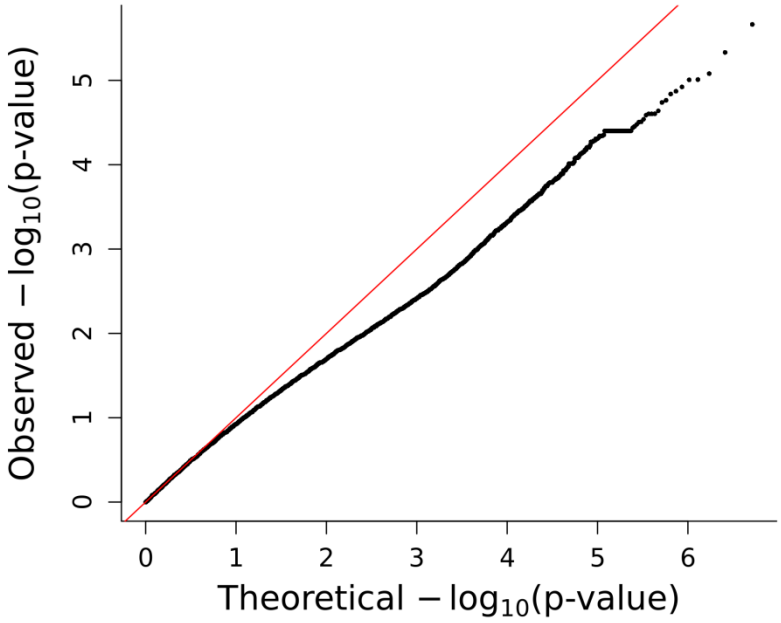


Fig. S61. Quantile-quantile plot for white-red color GWAS under a general linear model.

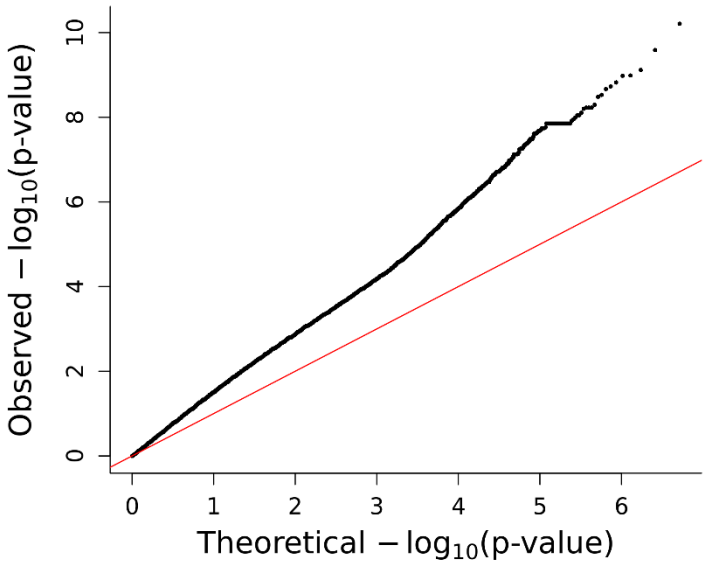


Fig. S62. Quantile-quantile plot for white-red color GWAS under a general linear model.

Table S1. Statistics of sequencing reads.

Items	Pacbio RSII	Pacbio corrected	Illumina X10
Total number of reads	17,755,597	10,185,636	428,890,719
Total number of sequenced bases (bp)	140,278,567,820	65,147,393,760	128,667,215,700
Mean reads length (bp)	7,900	6,396	150
N50 (bp)	12,525	9,262	150
Coverage (X)	82.5	38	75

Table S2. Statistics of genome assembly.

Strategy	Pacbio	Pacbio + bionano	Pacbio + bionano+ Hi-C
Assembly size (Gbp)	1.73	1.73	1.73
No. of contigs/Scaffolds	9,448	5,888	1,770
Maximum length (bp)	5,895,017	7,117,495	60,771,278
N90 (bp)	86,544	137,868	24,975,486
N80 (bp)	156,812	234,903	27,182,364
N70 (bp)	241,882	340,368	29,634,213
N60 (bp)	343,934	456,489	30,993,554
N50 (bp)	474,404	606,731	31,841,898
Average length (bp)	186,286	295,387	982,856

Table S3. Summary of BioNano data collection and assembly statistics.

	No.Molecule/genome maps	Total length (Mb)	Coverage	Molecule/map N50 (Kb)	Average molecule/ Scaffold (Kb)
Single molecules (>150 kb)	1,050,985	277,298	128	303	286
Map assembly	2,798	2,165	-	1,197	774

Characteristics	**Single molecules (>150 kb)**	**Map assembly**
Molecules/genome maps	1,050,985	2,798
Total length (Mb)	277,298	2,165
Coverage	128	-
Molecule/map N50 (Kb)	303	1,197
Average molecule/Scaffold (Kb)	286	774

Table S4. Statistics of Hi-C sequencing and mapping.

Statistics of mapping	
Clean Paired-end Reads	525,339,654
Mapped Reads	443,432,105
Mapped Ratio (%)	84.41
Unique Mapped Paired-end Reads	244,977,218
Unique Mapped Ratio (%)	46.63
Statistics of valid reads	
Unique Mapped Paired-end Reads	244,977,218
Valid Interaction Pairs	65,127,696
Valid Interaction Rate (%)	53.17
Dangling End Pairs	42,449,018
Dangling End Rate (%)	34.66
Religation Pairs	3,110,725
Religation Rate (%)	2.54
Self Circle Paired-end Reads	410,257
Self Circle Rate (%)	0.33
Dumped Paired-end Reads	11,390,913
Dumped Rate (%)	9.30

Table S5. Statistics of chromosome level scaffolding base Hi-C.

Subgenome A/B	No. of genes in A/B	No. of contigs in A/B	Total Length of A/B(bp)	The ratio of *Puntius semifasciolatus* reads mapping against on goldfish (%)
A01/B01	1,352/1,246	168/122	44,997,851/42,097,245	3.55/0.91
A02/B02	1,217/1,236	140/83	36,707,285/34,697,925	3.14/0.88
A03/B03	1,979/1,490	234/144	63,381,152/43,070,474	3.74/1.21
A04/B04	977/936	235/101	45,030,301/33,333,602	2.62/1.38
A05/B05	1,632/1,533	146/108	47,090,836/44,685,497	3.82/0.95
A06/B06	1,080/1,071	117/78	36,347,942/33,331,409	3.52/0.95
A07/B07	1,403/1,442	167/132	50,744,615/48,543,534	4.81/1.23
A08/B08	1,235/1,135	138/100	33,638,161/34,220,593	2.98/0.93
A09/B09	1,133/1,043	105/104	37,372,278/35,086,782	3.59/0.90
A10/B10	1,023/975	62/84	28,605,678/26,941,415	2.60/0.69
A11/B11	803/905	71/83	28,441,722/27,135,006	2.78/0.63
A12/B12	988/931	93/84	31,099,852/28,243,374	2.76/0.80
A13/B13	985/970	101/81	33,044,272/33,691,517	3.57/0.85
A14/B14	977/949	101/89	33,122,789/31,886,446	3.56/0.91
A15/B15	1,025/903	99/79	32,194,103/26,057,057	2.77/0.95
A16/B16	1,180/1,013	98/96	35,082,966/31,798,904	3.22/0.81
A17/B17	1,062/987	81/65	32,776,654/28,718,717	3.29/0.89
A18/B18	859/909	111/92	33,617,607/33,194,216	3.03/0.96
A19/B19	1,089/1,017	77/86	31,871,233/30,387,218	2.92/0.85
A20/B20	1,165/1,069	115/105	32,792,066/33,067,088	2.87/0.77
A21/B21	1,099/873	97/79	31,705,930/25,983,384	2.71/0.69
A22/B22	1,305/811	288/88	53,241,296/24,947,274	2.54/0.74
A23/B23	777/969	56/73	28,964,060/27,799,969	2.81/0.89
A24/B24	695/817	81/69	25,511,874/26,957,012	2.73/0.71
A25/B25	1,093/911	112/63	30,904,218/22,927,566	2.30/0.54
Total anchored sequences in A/B	28,133/26,141	3093/2288	918,286,741/808,803,224	-
Unclustered sequences	1,977(3.51%)	507	12,154,105	-
Total Sequences Clustered	-	5,381(91.50%)	1,727,089,965(99.30%)	-
Total Sequences ordered and Oriented	-	4,168(77.46%)	1,653,618,961(95.75%)	-

Table S6. Completeness of the genome based on CEGMA.

Description	Fully mapped CEGs	Fully+partially mapped CEGs
Number of CEGs present in the assembly	220	246
Completeness of the genome(%)	88.71	99.19
Average number of othologs per CEG	2.43	2.83
CEGs with more than one ortholog(%)	80.45	87.40

Table S7. BUSCO analysis of genome assembly.

Description	Number	Percentage(%)
Complete BUSCOs	4,344	94.7
Complete and single-copy BUSCOs	2,064	45.0
Complete and duplicated BUSCOs	2,280	49.7
Fragmented BUSCOs	92	2.0
Missing BUSCOs	148	3.3
Total BUSCO groups searched	4,584	100

Table S8. Classification of noncoding transcripts.

Intergenic						Genic					
Divergent		Convergent		Same strand		Overlapping		Containing		Nested	
U	D	U	D	U	D	E	I	E	I	E	I
1,269	0	0	1,057	1,228	1,074	1,366	944	561	464	607	771
4,628						**4,713**					

U, upstream; D, downstream; E, exonic; I, intronic.

Table S9. Summary of transposable elements and repeat annotation.

	Length(Mb)	% of Repeat	% of Genome	Genome_A	Genome_B
Total Repeat Fractions	574.83	100	33.04	15.97	14.89
Class Ⅰ:Retrotransposon	144.93	25.21	8.33	3.89	3.86
LTR Retrotransposon					
Ty3/Gypsy	28.65	4.98	1.65	0.75	0.78
Ty1/Copia	0.17	0.03	0.01	0.01	0.00
Other	38.02	6.61	2.19	1.03	1.02
Non-LTR Retrotransposon					
LINE	68.56	11.93	3.94	1.85	1.78
SINE	9.53	1.66	0.55	0.26	0.26
Class Ⅱ:DNA transposon	274.86	47.82	15.80	7.83	7.11
Autonomous					
CACTA	40.00	6.96	2.30	1.16	1.00
Mutator	1.84	0.32	0.11	0.08	0.03
hAT	56.37	9.81	3.24	1.59	1.47
PIF/Harbinger	10.43	1.81	0.60	0.29	0.27
Tc/Mariner	39.11	6.80	2.25	1.08	1.06
Other	116.24	20.22	6.68	3.35	3.02
Helitron	10.86	1.89	0.62	0.28	0.26
Tandem Repeat	44.21	7.69	2.54	1.15	1.02
Unknown	110.83	19.28	6.37	3.10	2.90

Table S10. Identification of centromeres in goldfish genome.

Chromosome	Length(Mb)	Position(Mb)		Length(b)	Chromosome	Length(Mb)	Position(Mb)		Length(b)
		start	end				start	end	
A01	42.439	37.207	37.210	2,452	B01	40.237	31.609	31.675	66,272
A02	34.494	3.003	3.005	2,037	B02	33.491	21.872	21.874	1,836
A03	60.771	48.385	48.679	294,564	B03	40.822	6.042	6.047	5,434
A04	41.400	29.351	29.385	33,842	B04	32.100	21.605	21.826	220,890
A05	46.420	29.067	29.067	1,915	B05	43.227	10.776	10.778	2,172
A06	34.955	0.844	0.850	6,037	B06	32.270	NA	NA	NA
A07	49.001	47.798	47.800	1,690	B07	47.434	22.938	22.940	1,560
A08	31.842	NA	NA	NA	B08	33.012	0.300	0.302	2,074
A09	35.860	15.487	15.489	2,036	B09	34.473	32.890	32.092	3,222
A10	27.473	NA	NA	NA	B10	25.974	11.834	11.836	1,741
A11	27.182	NA	NA	NA	B11	26.910	24.072	24.074	2,080
A12	29.890	29.117	29.118	1,959	B12	27.090	0.812	0.865	53,545
A13	31.410	3.163	3.165	1,913	B13	32.776	2.559	2.563	4,445
A14	31.722	NA	NA	NA	B14	30.071	1.789	1.790	1,422
A15	29.634	28.306	28.309	3,630	B15	24.975	0.157	0.160	3,602
A16	34.058	31.672	31.677	4,332	B16	30.899	12.544	12.546	2,148
A17	32.280	NA	NA	NA	B17	27.735	NA	NA	NA
A18	29.442	17.159	17.160	1,241	B18	30.994	30.587	30.988	401,185
A19	31.328	NA	NA	NA	B19	39.306	27.999	28.002	2,288
A20	31.311	NA	NA	NA	B20	31.926	28.215	28.216	1,365
A21	29.813	NA	NA	NA	B21	25.228	NA	NA	NA
A22	49.504	23.452	23.454	2,446	B22	23.971	23.617	23.624	6,129
A23	27.470	NA	NA	NA	B23	26.326	9.492	9.494	1,942
A24	24.392	17.289	17.294	4,703	B24	26.540	17.239	17.247	8,619
A25	29.951	7.849	7.856	5,916	B25	22.386	19.051	19.291	239,562

Table S11. Comparison of Subgenome notations in the studies of various goldfish and common carp

Current assembly	Kon *et al.* 2020 (42)	Luo *et al.* 2020 (2)	Xu *et al.* 2019 (40)
Subgenome A	Subgenome L	Subgenome M	Subgenome B
Subgenome B	Subgenome S	Subgenome P	Subgenome A

Table S12. Assessment of genome consistency.

Items	Statistics
Number of reads	442,861,726
Data size (Gb)	66
Mapped bases (Gb)	62
Map rate (%)	93.59
Genome Length (Mbp)	1,739
Mean Depth	46.4
Coverage Rate (%)	99.01

Table S13. Summary of sequenced *C. auratus* and wild *C. carp.*

Sample Acc.	Sample Name	Subpopulation*	Cluster	Location	Morphological characteristics												
					Body	TwinTails	DorsalFin	CaudalFin	TelescopeEye	CelestialEye	BubbleEye	HumpBack	Hood	Pompon	Scale	Pearlscale	Coloration
GF001	red white matt crown-pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Pearlscale	NA
GF002	black telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	black
GF003	black telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	black
GF004	black telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	black
GF005	chocolate broad-tail ryukin	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	HumpBack	Non-hooded	None	Normal	Normal	chocolate
GF006	blue chocolate long-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	long-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	bluechocolate
GF007	calico pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Pearlscale	NA
GF008	red white matt crown-pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Pearlscale	NA
GF009	chocolate pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Pearlscale	NA
GF010	broad-tail redcap-oranda	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white
GF011	red white broad-tail ryukin	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	HumpBack	Non-hooded	None	Normal	Normal	redwhite
GF012	calico short-tail egg-pompon	EG	Egg-pompon	Jiangsu province, China	ShortEgg	Twin	Absent	short-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Transparent	Normal	NA
GF013	calico short-tail egg-pompon	EG	Egg-pompon	Jiangsu province, China	ShortEgg	Twin	Absent	short-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Transparent	Normal	NA
GF014	calico short-tail egg-pompon	EG	Egg-pompon	Jiangsu province, China	ShortEgg	Twin	Absent	short-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Transparent	Normal	NA
GF015	chocolate celestial	EG	Egg-pompon	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	CelestialEye	NA	Normal	Non-hooded	None	Normal	Normal	chocolate

GF016	single-tail calico oranda	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	NA	Present	single-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF017	red egg-pheonix	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Phoenix Tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	red
GF018	red white egg-pheonix	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Phoenix Tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite
GF019	blue white bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Normal	Normal	bluewhite
GF020	blue bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Normal	Normal	blue
GF021	calico bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Transparent	Normal	NA
GF022	calico bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Transparent	Normal	NA
GF023	black bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Normal	Normal	black
GF024	black bubble-eye	EG	Eggfish-Lionhead	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	NA	NA	BubbleEye	Normal	Non-hooded	None	Normal	Normal	NA
GF025	blue long-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	long-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	blue
GF026	calico pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Pearlscale	NA
GF027	red chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF028	black red chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blackred
GF029	black chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF030	blue chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blue
GF031	white chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white
GF032	red white chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF033	black butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	NA	Normal	NA
GF034	black red butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	blackred

GF035	red white butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	redwhite
GF036	red butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	red
GF037	chocolate butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	chocolate
GF038	blue butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	blue
GF039	calico butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Transparent	Normal	NA
GF040	white butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	white
GF041	calico butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Transparent	Normal	NA
GF042	black red butterfly-tail telescope	WG	Telescope	Jiangsu province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	blackred
GF043	calico chinese-ranchu	EG	Chinese-ranchu	Jiangsu province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF044	red white matt crown-pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Pearlscale	NA
GF045	red white matt long-tail pearlscale	WG	Wenyu-Oranda	Jiangsu province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Pearlscale	NA
GF046	black pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	black
GF047	black pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	black
GF048	black pompon-	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	black

	telescope-oranda																	
GF049	black telescope-oranda	WG	Telescope	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	NA	None	Normal	Normal	black	
GF050	black pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	black	
GF051	chocolate pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	chocolate	
GF052	chocolate pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	chocolate	
GF053	black red pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	blackred	
GF054	red white matt telescope-oranda	WG	Telescope	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	NA	None	Transparent	Normal	NA	
GF055	red white matt telescope-oranda	WG	Telescope	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	NA	None	Transparent	Normal	NA	
GF056	white pompon-telescope-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Hooded	Pompon	Normal	Normal	white	
GF057	red white matt telescope-oranda	WG	Telescope	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	NA	None	Transparent	Normal	NA	
GF058	blue chocolate chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	bluechocolate	
GF059	lilac chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	lilac	
GF060	chocolate chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	chocolate	
GF061	black red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blackred	
GF062	white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white	

GF063	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF064	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF065	black chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF066	black red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blackred
GF067	red white short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF068	red white pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	redwhite
GF069	black pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	NA
GF070	red white pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	redwhite
GF071	chocolate oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	chocolate
GF072	black red long-tail pompon-celestial	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	NA	CelestialEye	NA	Normal	Non-hooded	Pompon	Normal	Normal	blackred
GF073	red long-tail pompon-celestial	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	NA	CelestialEye	NA	Normal	Non-hooded	Pompon	Normal	Normal	red
GF074	red white pompon-celestial	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	NA	CelestialEye	NA	Normal	Non-hooded	Pompon	Normal	Normal	redwhite
GF075	red short-tail egg-pompon	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	short-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Normal	Normal	red
GF076	lilac chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	lilac
GF077	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	NA	None	Transparent	Normal	NA
GF078	red long-tail pompon-lionhead	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	NA	Pompon	Normal	Normal	red
GF079	red matt long-tail chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA

GF080	black short-tail wen-pompon	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Normal	Normal	NA
GF081	black red pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	NA
GF082	red cat-lionhead	EG	Eggfish-Lionhead	Anhui province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF083	black red cat-lionhead	EG	Eggfish-Lionhead	Anhui province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blackred
GF084	black red pearlscale	WG	Wenyu-Oranda	Anhui province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Pearlscale	blackred
GF085	red white pearlscale	WG	Wenyu-Oranda	Anhui province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Pearlscale	redwhite
GF086	oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	NA
GF087	wen-butterflytail	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Butterfly Tail	Normal	Normal	Normal	Normal	NA	None	NA	Normal	NA
GF088	chocolate wen-butterflytail	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Butterfly Tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	chocolate
GF089	black short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF090	red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF091	white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white
GF092	red white matt long-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	NA	None	Transparent	Normal	NA
GF093	red white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF094	black short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF095	calico short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF096	black chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black

GF097	calico short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF098	calico broad-tail ryukin	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	HumpBack	Non-hooded	None	Transparent	Normal	NA
GF099	red white telescope	WG	Telescope	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	redwhite
GF100	calico short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF101	red white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF102	red oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF103	black chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF104	red egg-pompon	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Normal	Normal	red
GF105	red white matt long-tail chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF106	red long-tail highhead-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF107	red broad-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	broad-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF108	red pompon-celestial	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	NA	CelestialEye	NA	Normal	Non-hooded	Pompon	Normal	Normal	red
GF109	black red broad-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	broad-tail	Normal	Normal	Normal	Normal	NA	None	Normal	Normal	blackred
GF110	redcap-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF111	red dorsalfin-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	NA	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	NA
GF112	white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white
GF113	red white matt chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA

GF114	red shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	NA	None	Normal	Normal	red
GF115	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF116	red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF117	red white matt chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF118	red long-tail pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	red
GF119	white butterfly-tail telescope	WG	Telescope	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	white
GF120	blue long-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blue
GF121	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF122	red long-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	NA	None	Normal	Normal	red
GF123	chocolate chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	chocolate
GF124	black butterfly-tail telescope	WG	Telescope	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	black
GF125	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF126	black chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF127	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	NA	None	Normal	Normal	NA
GF128	calico oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF129	calico long-tail chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF130	red white jikin	CG	Common.goldfish	Fuzhou, Fujian province, China	Long	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite

GF131	chocolate dorsalfin-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	NA	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	NA
GF132	red broad-tail ryukin	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	HumpBack	Non-hooded	None	Normal	Normal	red
GF133	black short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	black
GF134	white dorsalfin-ranchu	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	white
GF135	red long-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	NA	None	Normal	Normal	red
GF136	calico wenyu	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Normal	NA
GF137	long-tail redcap-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF138	red white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF139	red long-tail highhead-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF140	calico oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF141	calico long-tail chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF142	blue long-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	blue
GF143	red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF144	red matt long-tail highhead-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF145	black butterfly-tail telescope	WG	Telescope	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	butterfly-tail	TelescopeEye	NA	NA	Normal	Non-hooded	None	Normal	Normal	black
GF146	chocolate pompon-lionhead	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	NA	Pompon	Normal	Normal	chocolate
GF147	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA

GF148	black red pompon-oranda	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	Pompon	Normal	Normal	blackred
GF149	red white matt wen-butterflytail	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Butterfly Tail	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Normal	NA
GF150	red chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF151	red broad-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	broad-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF152	red broad-tail shouxin-lionhead	EG	Eggfish-Lionhead	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	broad-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF153	red white matt wen-butterflytail	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Butterfly Tail	Normal	Normal	Normal	Normal	NA	None	Transparent	Normal	NA
GF154	chocolate long-tail chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	chocolate
GF155	chocolate long-tail wen-pompon	WG	Wen-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Non-hooded	Pompon	Normal	Normal	chocolate
GF156	red white fu-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF157	calico broad-tail ryukin	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	broad-tail	Normal	Normal	Normal	HumpBack	Non-hooded	None	Transparent	Normal	NA
GF158	red long-tail pompon-celestial	EG	Egg-pompon	Fuzhou, Fujian province, China	ShortEgg	Twin	Absent	long-tail	NA	CelestialEye	NA	Normal	Non-hooded	Pompon	Normal	Normal	red
GF159	red white long-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	long-tail	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF160	red matt wenyu	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Normal	NA
GF161	red white jikin	CG	Common.goldfish	Fuzhou, Fujian province, China	Long	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite
GF162	red white matt short-tail oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	short-tail	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF163	red wenyu	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEgg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	red

GF164	red highhead-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF165	chocolate highhead-oranda	WG	Wenyu-Oranda	Fuzhou, Fujian province, China	ShortEg	Twin	Present	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	chocolate
GF166	red white chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF167	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF168	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF169	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF170	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF171	chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	NA	Normal	NA
GF172	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF173	calico chinese-ranchu	EG	Chinese-ranchu	Fuzhou, Fujian province, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Transparent	Normal	NA
GF174	red wang-tigerhead	EG	Eggfish-Lionhead	Beijing, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	red
GF175	red nankin	EG	Eggfish-Lionhead	Beijing, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	red
GF176	goosehead	EG	Eggfish-Lionhead	Beijing, China	ShortEg	Twin	Absent	Normal	Normal	Normal	Normal	Normal	Hooded	None	Normal	Normal	redwhite
GF177	red white matt egg-pheonix	EG	Eggfish-Lionhead	Beijing, China	ShortEg	Twin	Absent	Phoenix Tail	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Normal	NA
GF178	red broad-tail nankin	EG	Eggfish-Lionhead	Beijing, China	ShortEg	Twin	Absent	broad-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	red
GF179	red common goldfish	CG	Common.goldfish	Fuzhou, Fujian province, China	Long	Single	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	NA	Normal	red
GF180	red common goldfish	CG	Common.goldfish	Fuzhou, Fujian province, China	Long	Single	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	NA	Normal	red
GF182	red white wakin	CG	Common.goldfish	Suzhou, Jiangsu province, China	Long	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite
GF183	black red long-tail common goldfish	CG	Common.goldfish	Suzhou, Jiangsu province, China	Long	Single	Present	long-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	blackred
GF184	red white common goldfish	CG	Common.goldfish	Suzhou, Jiangsu province, China	Long	Single	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite

GF185	calico common goldfish	CG	Common.goldfish	Suzhou, Jiangsu province, China	Long	Single	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Transparent	Normal	NA
GF186	red white wakin	CG	Common.goldfish	Guangzhou, Guangdong province, China	Long	Twin	Present	Normal	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	redwhite
GF187	crucian carp ZJPH1	CC	Crucian.carp	Pinghu, Zhejiang province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF188	crucian carp ZJPH2	CC	Crucian.carp	Pinghu, Zhejiang province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF189	crucian carp JSNJ1	CC	Crucian.carp	Nanjing, Jiangsu province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF190	crucian carp JXSR1	CC	Crucian.carp	Shangrao, Jiangxi province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF191	crucian carp JXXY1	CC	Crucian.carp	Xinyu, Jiangxi province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF192	crucian carp HNXY1	CC	Crucian.carp	Xinyang, Henan province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF193	crucian carp HNZMD1	CC	Crucian.carp	Zhumadian, Henan province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF194	crucian carp HLJHEB1	CC	Crucian.carp	Haerbin, Heilongjiang province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF195	crucian carp SDBZ1	CC	Crucian.carp	Binzhou, Shandong province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF196	crucian carp FJSM1	CC	Crucian.carp	Sanming, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF197	crucian carp FJSM2	CC	Crucian.carp	Sanming, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	NA
GF198	crucian carp FJLY1	CC	Crucian.carp	Longyan, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF199	crucian carp FJZN1	CC	Crucian.carp	Zhounin, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF200	crucian carp FJZN2	CC	Crucian.carp	Zhounin, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	grey
GF201	crucian carp BJY01	CC	Crucian.carp	Fuzhou, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	NA
GF202	crucian carp BJY02	CC	Crucian.carp	Fuzhou, Fujian province, China	Long	Single	Present	single-tail	Normal	Normal	Normal	Normal	Non-hooded	None	Normal	Normal	NA

* WG: wen-goldfish ; EG: egg-goldfish; CG: common goldfish; CC: crucian carp, All information for the detail were listed in Supplementary Data 2.

Table S14. Summary resequencing statistics for the 202 resequenced genomes.

Sample ID	Sequence coverage(x)	Sequencing mode	Clean Bases Number(Gbp)
GF001	15.6	paired end 150 nt	28.0
GF002	16.6	paired end 150 nt	30.0
GF003	10.6	paired end 150 nt	19.1
GF004	11.7	paired end 150 nt	21.0
GF005	12.2	paired end 150 nt	22.0
GF006	11.1	paired end 150 nt	19.9
GF007	12.2	paired end 150 nt	22.0
GF008	9.9	paired end 150 nt	17.7
GF009	13.8	paired end 150 nt	24.9
GF010	13.2	paired end 150 nt	23.8
GF011	10.6	paired end 150 nt	19.1
GF012	11.1	paired end 150 nt	20.0
GF013	12.1	paired end 150 nt	21.8
GF014	10.1	paired end 150 nt	18.2
GF015	19.8	paired end 150 nt	35.6
GF016	12.3	paired end 150 nt	22.1
GF017	14.5	paired end 150 nt	26.1
GF018	10.0	paired end 150 nt	18.0
GF019	12.7	paired end 150 nt	22.8
GF020	9.9	paired end 150 nt	17.8
GF021	9.9	paired end 150 nt	17.8
GF022	12.9	paired end 150 nt	23.2
GF023	13.6	paired end 150 nt	24.4
GF024	12.6	paired end 150 nt	22.8
GF025	13.2	paired end 150 nt	23.8
GF026	11.5	paired end 150 nt	20.8
GF027	11.1	paired end 150 nt	20.1
GF028	10.1	paired end 150 nt	18.2
GF029	12.9	paired end 150 nt	23.2
GF030	14.3	paired end 150 nt	25.7
GF031	10.2	paired end 150 nt	18.3
GF032	10.1	paired end 150 nt	18.2
GF033	12.2	paired end 150 nt	22.0
GF034	9.9	paired end 150 nt	17.7
GF035	10.9	paired end 150 nt	19.5
GF036	9.6	paired end 150 nt	17.3
GF037	9.9	paired end 150 nt	17.8
GF038	10.0	paired end 150 nt	18.0
GF039	13.9	paired end 150 nt	25.0
GF040	11.4	paired end 150 nt	20.6
GF041	10.4	paired end 150 nt	18.8
GF042	11.5	paired end 150 nt	20.7
GF043	9.1	paired end 150 nt	16.3
GF044	11.0	paired end 150 nt	19.8
GF045	10.4	paired end 150 nt	18.6
GF046	10.9	paired end 150 nt	19.7
GF047	10.9	paired end 150 nt	19.7
GF048	10.0	paired end 150 nt	18.0
GF049	12.4	paired end 150 nt	22.2
GF050	10.6	paired end 150 nt	19.1
GF051	10.1	paired end 150 nt	18.1

GF052	10.0	paired end 150 nt	18.0
GF053	10.5	paired end 150 nt	18.9
GF054	10.3	paired end 150 nt	18.5
GF055	10.2	paired end 150 nt	18.4
GF056	11.6	paired end 150 nt	20.8
GF057	10.0	paired end 150 nt	18.0
GF058	12.3	paired end 150 nt	22.1
GF059	11.0	paired end 150 nt	19.8
GF060	11.6	paired end 150 nt	20.8
GF061	10.7	paired end 150 nt	19.2
GF062	12.5	paired end 150 nt	22.4
GF063	11.0	paired end 150 nt	19.8
GF064	13.0	paired end 150 nt	23.3
GF065	11.7	paired end 150 nt	21.1
GF066	11.5	paired end 150 nt	20.7
GF067	11.8	paired end 150 nt	21.2
GF068	13.3	paired end 150 nt	23.9
GF069	11.1	paired end 150 nt	20.0
GF070	11.1	paired end 150 nt	20.0
GF071	8.8	paired end 150 nt	15.8
GF072	10.7	paired end 150 nt	19.3
GF073	10.3	paired end 150 nt	18.5
GF074	12.9	paired end 150 nt	23.3
GF075	11.7	paired end 150 nt	21.1
GF076	11.6	paired end 150 nt	20.9
GF077	10.7	paired end 150 nt	19.2
GF078	10.3	paired end 150 nt	18.6
GF079	10.0	paired end 150 nt	17.9
GF080	15.4	paired end 150 nt	27.7
GF081	13.1	paired end 150 nt	23.6
GF082	11.2	paired end 150 nt	20.2
GF083	11.6	paired end 150 nt	20.8
GF084	11.1	paired end 150 nt	19.9
GF085	10.1	paired end 150 nt	18.2
GF086	16.7	paired end 150 nt	30.0
GF087	15.8	paired end 150 nt	28.5
GF088	14.6	paired end 150 nt	26.2
GF089	13.2	paired end 150 nt	23.7
GF090	11.4	paired end 150 nt	20.5
GF091	10.0	paired end 150 nt	18.0
GF092	15.6	paired end 150 nt	28.1
GF093	13.3	paired end 150 nt	23.9
GF094	12.4	paired end 150 nt	22.3
GF095	12.3	paired end 150 nt	22.2
GF096	12.4	paired end 150 nt	22.3
GF097	13.8	paired end 150 nt	24.8
GF098	11.1	paired end 150 nt	20.0
GF099	9.9	paired end 150 nt	17.8
GF100	11.1	paired end 150 nt	19.9
GF101	11.7	paired end 150 nt	21.1
GF102	11.2	paired end 150 nt	20.2
GF103	14.7	paired end 150 nt	26.5
GF104	11.0	paired end 150 nt	19.8
GF105	13.0	paired end 150 nt	23.4
GF106	11.0	paired end 150 nt	19.9
GF107	10.7	paired end 150 nt	19.2

GF108	11.5	paired end 150 nt	20.8
GF109	10.0	paired end 150 nt	18.1
GF110	11.4	paired end 150 nt	20.5
GF111	11.0	paired end 150 nt	19.7
GF112	11.4	paired end 150 nt	20.6
GF113	10.0	paired end 150 nt	18.1
GF114	12.9	paired end 150 nt	23.1
GF115	12.0	paired end 150 nt	21.6
GF116	13.1	paired end 150 nt	23.7
GF117	12.1	paired end 150 nt	21.8
GF118	12.2	paired end 150 nt	22.0
GF119	11.6	paired end 150 nt	20.9
GF120	12.2	paired end 150 nt	22.0
GF121	10.7	paired end 150 nt	19.3
GF122	12.2	paired end 150 nt	21.9
GF123	11.3	paired end 150 nt	20.3
GF124	11.6	paired end 150 nt	20.9
GF125	11.4	paired end 150 nt	20.5
GF126	11.6	paired end 150 nt	20.9
GF127	13.5	paired end 150 nt	24.3
GF128	11.7	paired end 150 nt	21.1
GF129	12.0	paired end 150 nt	21.7
GF130	11.3	paired end 150 nt	20.4
GF131	13.2	paired end 150 nt	23.8
GF132	11.6	paired end 150 nt	20.8
GF133	12.3	paired end 150 nt	22.2
GF134	12.3	paired end 150 nt	22.2
GF135	11.8	paired end 150 nt	21.3
GF136	11.7	paired end 150 nt	21.0
GF137	11.2	paired end 150 nt	20.1
GF138	13.3	paired end 150 nt	24.0
GF139	11.8	paired end 150 nt	21.2
GF140	12.1	paired end 150 nt	21.8
GF141	12.9	paired end 150 nt	23.3
GF142	12.8	paired end 150 nt	23.1
GF143	12.7	paired end 150 nt	22.9
GF144	7.1	paired end 150 nt	12.9
GF145	9.1	paired end 150 nt	16.3
GF146	9.0	paired end 150 nt	16.2
GF147	13.5	paired end 150 nt	24.4
GF148	10.5	paired end 150 nt	18.9
GF149	11.6	paired end 150 nt	20.9
GF150	12.3	paired end 150 nt	22.1
GF151	12.6	paired end 150 nt	22.6
GF152	10.9	paired end 150 nt	19.6
GF153	11.9	paired end 150 nt	21.5
GF154	14.5	paired end 150 nt	26.0
GF155	12.8	paired end 150 nt	23.1
GF156	10.5	paired end 150 nt	18.9
GF157	11.0	paired end 150 nt	19.7
GF158	10.8	paired end 150 nt	19.4
GF159	10.9	paired end 150 nt	19.7
GF160	9.2	paired end 150 nt	16.6
GF161	10.3	paired end 150 nt	18.6
GF162	11.3	paired end 150 nt	20.4
GF163	10.4	paired end 150 nt	18.7

GF164	11.4	paired end 150 nt	20.5
GF165	11.8	paired end 150 nt	21.2
GF166	13.3	paired end 150 nt	24.0
GF167	11.0	paired end 150 nt	19.8
GF168	11.4	paired end 150 nt	20.5
GF169	11.4	paired end 150 nt	20.6
GF170	10.6	paired end 150 nt	19.0
GF171	9.0	paired end 150 nt	16.2
GF172	15.1	paired end 150 nt	27.2
GF173	14.2	paired end 150 nt	25.6
GF174	10.5	paired end 150 nt	18.9
GF175	10.0	paired end 150 nt	18.0
GF176	10.5	paired end 150 nt	18.9
GF177	4.2	paired end 150 nt	7.6
GF178	12.2	paired end 150 nt	22.0
GF179	34.6	paired end 150 nt	62.2
GF180	36.9	paired end 150 nt	66.4
GF182	10.8	paired end 150 nt	19.4
GF183	12.2	paired end 150 nt	21.9
GF184	10.6	paired end 150 nt	19.1
GF185	8.2	paired end 150 nt	14.7
GF186	9.4	paired end 150 nt	17.0
GF187	10.5	paired end 150 nt	18.9
GF188	11.1	paired end 150 nt	19.9
GF189	11.0	paired end 150 nt	19.8
GF190	11.1	paired end 150 nt	20.1
GF191	9.7	paired end 150 nt	17.5
GF192	15.2	paired end 150 nt	27.4
GF193	10.6	paired end 150 nt	19.1
GF194	11.0	paired end 150 nt	19.7
GF195	12.1	paired end 150 nt	21.7
GF196	9.1	paired end 150 nt	16.4
GF197	10.0	paired end 150 nt	18.0
GF198	10.6	paired end 150 nt	19.1
GF199	7.8	paired end 150 nt	14.1
GF200	9.1	paired end 150 nt	16.4
GF201	10.8	paired end 150 nt	19.4
GF202	11.2	paired end 150 nt	20.1

Dataset S1. Photo of 185 goldfishes and 16 crucian carps used for resequencing.

Dataset S2. Information included 185 goldfishes, 16 crucian carps and 33 common carps.

Dataset S3. Sweep regions and candidate genes list for goldfish domestication.

Dataset S4. Genotype heatmap of sweep regions for goldfish domestication.

Dataset S5. Candidate regions and candidate genes list for dorsal fin (Wen and Egg goldfish) trait GWAS.

Dataset S6. Genotype heatmap of candidate regions for dorsal fin trait GWAS.

Dataset S7. Candidate regions and candidate genes list for Blue-Red color trait GWAS.

Dataset S8. Candidate regions and candidate genes list for white-red color trait GWAS.

Dataset S9. Candidate regions and candidate genes list for Black-Red color trait GWAS.

Dataset S10. Homolog gene pairs list for Subgenomes A and B.

Dataset S11. Whole genome gene function annotation based on GO and KEGG database.

Dataset S12. The transcript expression level of ten tissues of goldfish.

SI References

1. T. Nakajima, M. J. Hudson, J. Uchiyama, K. Makibayashi, J. Zhang, Publisher Correction: Common carp aquaculture in Neolithic China dates back 8,000 years. *Nature Ecology & Evolution* **3**, 1494 (2019).
2. K. Katoh, D. M. Standley, MAFFT: Iterative Refinement and Additional Methods. *Methods in Molecular Biology* **1079**, 131-146 (2014).
3. S. C. Chen, Transparency and Mottling, a Case of Mendelian Inheritance in the Goldfish CARASSIUS AURATUS. *Genetics* **13**, 434-452 (1928).
4. G. F. Hervey, E. Billardon-Sauvigny. Mus éum national d'histoire, The goldfish of China in the XVIII century, Hervey, Eds. (China Society, 1950),pp.2-3.
5. Y. Matsui, "Physically speaking" in Goldfish, Matsui, Eds. (T.F.H. Publications,1981), pp.30-32.
6. S. Chen, A history of the domestication and the factors of the varietal formation of the common goldfish, Carassius auratus. *Scientia Sinica* **5**, 287-321 (1956).
7. S. C. Chen, The inheritance of blue and brown colours in the goldfish, *Carassius auratus*. *J. Genet.* **29**, 61-74 (1934).
8. J. Smartt, "Introduction" in Goldfish Varieties and Genetics: A Handbook for Breeders, Smartt, Eds. (Blackwell Science,2008), pp.1-10.
9. E. Billardon-Sauvigny, F. N. Martinet, "Introduction" in Histoire naturelle des dorades de la Chine, Billardon-Sauvigny, Eds. (Wentworth Press,2019),pp.19.
10. X. Wang, X. Gan, J. Li, Y. Chen, S. He, Cyprininae phylogeny revealed independent origins of the Tibetan Plateau endemic polyploid cyprinids and their diversifications related to the Neogene uplift of the plateau. *Sci. China Life. Sci.* **59**, 1149-1165 (2016).
11. T. G. Maja, N. Chen, Using RepeatMasker to identify repetitive elements in genomic sequences, *Current Protocols in Bioinformatics* **25,** 10-14 (2009).
12. K. Katoh, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. **30**,3059-3066 (2002).
13. K. Howe *et al.*, The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498-503 (2013).
14. Y. Wang *et al.*, Erratum: The draft genome of the grass carp (Ctenopharyngodon idellus) provides insights into its evolution and vegetarian adaptation. *Nat. Genet.* **47**, 962 (2015).
15. J. L. Anderson, *et al.*, Multiple Sex-Associated Regions and a Putative Sex Chromosome in Zebrafish Revealed by RAD Mapping and Population Genomics, *PLoS One*, **7**, e40701 (2012).
16. M. Wen, *et al.*, Sex chromosome and sex locus characterization in the goldfish, *Carassius auratus. BioRxiv*, 875377 (2019).

17. R. Goto-Kazeto, *et al.*, Temperature-dependent sex differentiation in goldfish: Establishing the temperature-sensitive period and effect of constant and fluctuating water temperatures. *Aquaculture* **254**, 617-624 (2006).
18. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. methods* **9**, 357-359 (2012).
19. T. O. Yamamoto, T. Kajishima, Sex hormone induction of sex reversal in the goldfish and evidence for male heterogamity. *J. Exp. Zool.* **168**, 215-221 (1968).
20. F. Martinon, O. Gaide, V. Petrilli, A. Mayor, J. Tschopp, NALP inflammasomes: a central role in innate immunity. *Semin. immunopathol.* **29**, 213-229 (2007).
21. A. Ferrer-Admetlla *et al.*, Balancing selection is the main force shaping the evolution of innate immunity genes. *J. Immunol.* **181**, 1315-1322 (2008).
22. P. Xu *et al.*, Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat. Genet.* **46**, 1212-1219 (2014).
23. Y. Wang, *et al.*, The draft genome of the grass carp (Ctenopharyngodon idellus) provides insights into its evolution and vegetarian adaptation. *Nat. Genet.* **47**, 625-631 (2015).
24. J. Smartt, "Introduction" in *Goldfish Varieties and Genetics: A Handbook for Breeders*, Smartt, Eds. (Blackwell Science,2008), pp.1-10.
25. V. Wucher, *et al.*, FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res.* **45**, e57 (2017).
26. A. Choudhuri, T. Evans, U. Maitra, Non-core subunit eIF3h of translation initiation factor eIF3 regulates zebrafish embryonic development. *Dev. Dyn.* **239**, 1632-1644 (2010).
27. A. Docquier, *et al.*, eIF3f depletion impedes mouse embryonic development, reduces adult skeletal muscle mass and amplifies muscle loss during disuse. *The J. physiol.* **597**, 3107-3131 (2019).
28. E. Carlsson, *et al.*, Potential role of a navigator gene NAV3 in colorectal cancer. *Br. J. Cancer* **106**, 517-524 (2012).
29. A. A. George, S. Hayden, G. R. Stanton, S. E. Brockerhoff, Arf6 and the 5'phosphatase of synaptojanin 1 regulate autophagy in cone photoreceptors. *BioEssays* **38**, S119-S135 (2016).
30. A. Rissone, *et al.*, Adenylate Kinase 2 Regulates Zebrafish Primitive and Definitive Hematopoiesis. *Blood* **120**, 1208-1208 (2012).
31. S. Volpi, *et al.*, EXTL3 mutations cause skeletal dysplasia, immune deficiency, and developmental delay. *J. Exp. Med.* **214**, 623-637 (2017).
32. J. C. Moore *et al.*, T cell immune deficiency in zap70 mutant zebrafish. *Molecular and cellular biology* **36**, 2868-2876 (2016).

33. R. Maeda, I. V. Pacentine, T. Erickson, T. Nicolson, Functional analysis of the transmembrane and cytoplasmic domains of Pcdh15a in zebrafish hair cells. *J. Neurosci.* **37**, 3231-3245 (2017).
34. S. J. Wanner, I. Saeger, S. Guthrie, V. E. Prince, Facial motor neuron migration advances. *Curr. Opin. Neurobiol.* **23**, 943-950 (2013).
35. Y. Matsui, Genetical studies on gold-fish of Japan. 2. On the Mendelian inheritance of the telescope eyes of gold-fish. *Journal of the Imperial Fisheries Institute* **30**, 37-46 (1934).
36. G. F. Hervey, J. Hems, The goldfish, Hervey, Eds. (Batchworth Press, London, 1948), pp.232.
37. C. Darwin, "Maldonado" in The Voyage of the Beagle. Darwin, Eds. (Penguin Books, 1839).pp.61-70.
38. A. A. Sharov, *et al.*, Bone morphogenetic protein (BMP) signaling controls hair pigmentation by means of cross-talk with the melanocortin receptor-1 pathway. *P. Natl. Acad. Sci. U. S. A.* **102**, 93-98 (2005).
39. C. B. Kaelin, *et al.*, Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science* **337**, 1536-1541 (2012).
40. V. A. Kottler ,*et al.*, Adenylate cyclase 5 is required for melanophore and male pattern development in the guppy (Poecilia reticulata). *Pig. Cell Melanoma R.* **28**, 545-558 (2015).
41. P. Elena, D. Fredj-Reygrobellet, G. Moulin, P. Lapalus, Pharmacological characteristics of β-adrencrgic-sensitive adenylate cyclase in non pigmented and in pigmented cells of bovine ciliary process. *Curr. Eye Res.* **3**, 1383-1389 (1984).
42. J. D'Orazio, D. E. Fisher, Central role for cAMP signaling in pigmentation and UV resistance. *Cell cycle* **10**, 8-9 (2011).
43. K. G. Daly, *et al.*, Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Science* **361**, 85-88 (2018).
44. F. J. Alberto, *et al.*, Convergent genomic signatures of domestication in sheep and goats. *Nature communications* **9**, 813 (2018).
45. N. Chen, *et al.*, Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nature communications* **9**, 2337 (2018).
46. J. Leno-Colorado, N. J. Hudson, A. Reverter, M. Perez-Enciso, A Pathway-Centered Analysis of Pig Domestication and Breeding in Eurasia. *G3-Genes Genomes Genetics* **7**, 2171-2184 (2017).
47. S. C. Chen, Transparency and mottling, a case of Mendelian inheritance in the goldfish *Carassius auratus*. *Genetics* **13**, 434 (1928).
48. T. Kajishima, Genetic and developmental analysis of some new color mutants in the goldfish, *Carassius auratus*. *Genetics* **86**, 161-174 (1977).

49. T. O. Yamamoto, A YY male goldfish from mating estrone-induced XY female and normal male. *J. Hered.* **66**, 2-4 (1975).
50. S. Ohno, "The conservative nature of chromosomal evolution" in Evolution by Gene Duplication, Ohno, Eds. (Springer, 1970),pp.41-46.
51. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
52. G. Benson, Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573-580 (1999).
53. G. Abrusan, N. Grundmann, L. DeMester, W. Makalowski, TEclass--a tool for automated classification of unknown eukaryotic transposable elements. *Bioinformatics* **25**, 1329-1330 (2009).
54. Z. Xu, H. Wang, LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, 265-268 (2007).
55. D. Ellinghaus, S. Kurtz, U. Willhoeft, LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC bioinformatics* **9**, 18 (2008).
56. S. Ou, N. Jiang, LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. *Plant Physiol.* **176**, 1410-1422 (2018).
57. N. Chen, *et al.*, Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nat. Commun.* **9**, 2337 (2018).
58. M. G. Grabherr, *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644-652 (2011).
59. I. Korf, Gene finding in novel genomes. *BMC bioinformatics* **5**, 59 (2004).
60. A. Lomsadze, V. Ter-Hovhannisyan, Y. O. Chernoff, M. Borodovsky, Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* **33**, 6494-6506 (2005).
61. M. Stanke, O. Schoffmann, B. Morgenstern, S. Waack, Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC bioinformatics* **7**, 62 (2006).
62. B. L. Cantarel, *et al.*, MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188-196 (2008).
63. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
64. F. Cabanettes, C. Klopp, D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ.* **6**, e4958 (2018).
65. Z. Hao, *et al.*, RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *Peer J.* **6**,e251(2019).

66. B. Langmead, C. Wilks, V. Antonescu, R. Charles, Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35**, 421-432 (2019).
67. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
68. B. J. Haas, *et al.*, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494-1512 (2013).
69. D. Kim, *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biol.* **14**, R36 (2013).
70. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
71. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
72. G. G. Faust, I. M. Hall, SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503-2505 (2014).
73. A. McKenna, *et al.*, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303 (2010).
74. S. Summa, *et al.*, GATK hard filtering: tunable parameters to improve variant calling for next generation sequencing targeted gene panel data. *BMC Bioinformatics* **18**,119 (2017).
75. P. Danecek, *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
76. B. L. Browning, Y. Zhou, S. R. Browning, A One-Penny Imputed Genome from Next-Generation Reference Panels. *Am. J. Hum. Genet.* **103**, 338-348 (2018).
77. P. Cingolani, *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92 (2012).
78. P. Xu *et al.*, Genome sequence and genetic diversity of the common carp, Cyprinus carpio. *Nat. Genet.* **46**, 1212-1219 (2014).
79. C. C. Chang, *et al.*, Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
80. V. Lefort, R. Desper, O. Gascuel, FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. *Mol. Biol Evol.* **32**, 2798-2800 (2015).
81. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, 242-245 (2016).

82. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
83. M. DeGiorgio, C. D. Huber, M. J. Hubisz, I. Hellmann, R. Nielsen, SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics* **32**, 1895-1897 (2016).
84. F. Supek, M. Bosnjak, N. Skunca, T. Smuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One* **6**, e21800 (2011).
85. P. Shannon, *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498-2504 (2003).
86. M. Gautier, A. Klassmann, R. Vitalis, rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mole. Ecol. Res.* **17**, 78-90 (2017).