SCIENCE CHINA PRESS

SCIENCE CHINA Life Sciences
life.scichina.com link.springer.com
https://doi.org/10.1007/s11427-024-2868-y

RESEARCH PAPER

Genomic and chromosomal architectures underlying fertility maintenance in the testes of intergeneric homoploid hybrids

Li Ren[1,2,3†], Yiyan Zeng[1†], Qizhi Liu[1,2,3†], Xiaolong Tu[4,6†], Fayi Chen[7], Hao Wu[7], Chuan Wang[7], Chang Wu[1,2,3], Mengxue Luo[1], Yakui Tai[1], Hailu Zhou[1], Mengdan Li[1], Ling Liu[1], Dongdong Wu[4,5] & Shaojun Liu[1,2,3*]

[1]Engineering Research Center of Polyploid Fish Reproduction and Breeding of the State Education Ministry, College of Life Sciences, Hunan Normal University, Changsha 410081, China
[2]Yuelushan Laboratory, Changsha 410081, China
[3]Hunan Yuelu Mountain Science and Technology Co. Ltd. for Aquatic Breeding, Changsha 410081, China
[4]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China
[5]Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[6]Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming 650204, China
[7]Wuhan Generead Biotechnologies Co. Ltd., Wuhan 430000, China

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

Received 22 October 2024; Accepted 13 February 2025; Published online 23 May 2025

The remarkable diversity of the Cyprinidae family highlights the importance of hybridization and gene flow in generating genetic variation, adaptation, and even speciation. However, why do cyprinid fish frequently overcome postzygotic reproductive isolation, a mechanism that normally prevents successful reproduction after fertilization? To address this gap in knowledge, we conducted comparative studies using reciprocal F_1 hybrid lineages derived from intergeneric hybridization between the cyprinid species *Megalobrama amblycephala* and *Culter alburnus*. Utilizing long-read genome sequencing, ATAC-seq, Hi-C, and mRNA-seq technologies, we identified rapid genomic variations, chromatin remodeling, and gene expression changes in the testicular cells of F_1 hybrid individuals. By analyzing the distribution of these alterations across three gene categories (allelic genes, orphan genes, and testis-specific genes), we found that changes were less pronounced in allelic and testis-specific genes but significantly more pronounced in orphan genes. Furthermore, we hypothesize that *rnf212b* is a crucial testis-specific gene that regulates spermatogenesis. Our findings suggest that allelic and testis-specific genes potentially mitigate "genomic shock" on reproductive function following hybridization. This research offers potential insights into the formation mechanisms of homoploid hybridization by demonstrating the coordinated interplay of genomic variations, chromatin remodeling, and gene expression changes during testicular development and spermatogenesis.

homoploid hybridization | testis | structural variation | chromatin remodeling

INTRODUCTION

Inland rivers and lakes experience frequent habitat alterations that disrupt ecological relationships and facilitate hybridization among freshwater fish (Hu et al., 2021; Janko et al., 2018; Meier et al., 2019). Recent studies suggest that rather than reproductive isolation, geographical or ecological isolation might be the driving force behind diversification in certain freshwater cyprinid lineages (Nicol et al., 2017; Takamura and Nakahara, 2015). Moreover, the successful production of fertile hybrid offspring in some freshwater cyprinid fish demonstrates the absence of postzygotic reproductive isolation, a factor generally considered critical in limiting gene flow in animals (Wang et al., 2019). However, cyprinid fish also undergo rapid genetic differentiation within their reproductive systems during speciation, a phenomenon observed in other vertebrates (Capel, 2017). This rapid genetic variation, despite the lack of postzygotic reproductive isolation in hybrid progenies, raises a fascinating question: how do the underlying genetic mechanisms maintain fertility in the gonads of these hybrid cyprinid fish?

Homoploid hybrids exhibit extremely high genetic variance, including phenotypes that are more extreme than either parent (Mallet, 2007), due in part to subgenome rearrangements from different parental species (Bozdag et al., 2021; Ren et al., 2019). These genomic changes can be accompanied by bursts of 3D genome restructuring (Lin et al., 2024) and gene expression variation (Cox et al., 2014; Gao et al., 2021; Ren et al., 2019; Romero et al., 2012; Yoo et al., 2014), which rapidly accumulate in early hybrid generations. These rapid genomic changes facilitate the emergence of novel variant phenotypes and enable homoploids to occupy new ecological niches distinct from those parental species populations (Buerkle et al., 2000; Rieseberg et al., 2003). Moreover, these genetic variants can be introduced back into parental species populations through further hybridi-

Citation: Ren, L., Zeng, Y., Liu, Q., Tu, X., Chen, F., Wu, H., Wang, C., Wu, C., Luo, M., Tai, Y., et al. Genomic and chromosomal architectures underlying fertility maintenance in the testes of intergeneric homoploid hybrids. Sci China Life Sci, https://doi.org/10.1007/s11427-024-2868-y

zation events, providing new avenues for adaptive evolution (Meier et al., 2019). However, it remains unclear whether these genomic alterations also harm the reproductive system, leading to hybrid sterility and postzygotic reproductive barriers. Investigating the impact of genomic alterations on fertility in homoploid hybrids of cyprinid fish can shed light on this critical question.

Chromatin remodeling, the dynamic reorganization of chromatin structure, plays a crucial role in regulating gene expression by modulating DNA accessibility. Mediated by chromatin remodelers, histone modifications, and nucleosome repositioning, it responds to diverse signals, including developmental cues, environmental changes, and stresses (Lin et al., 2024; Sokpor et al., 2017; Weaver et al., 2017). In reproductive biology, chromatin remodeling is pivotal for testicular development and spermatogenesis, ensuring precise gene regulation during spermatogonial differentiation and meiosis (Murat et al., 2023). In fish, it regulates pathways essential for gonadal differentiation and germ cell maturation (Hou et al., 2024). Disruption of these pathways has been linked to male infertility and testicular dysfunction. Despite its significance, the role of chromatin remodeling in hybrid fish remains unexplored. Investigating chromatin remodeling in hybrid testicular development may reveal key mechanisms underlying hybrid sterility and reproductive isolation.

To elucidate the molecular mechanisms underlying testicular development and spermatogenesis in hybrid cyprinid fish, we focused on genomic variations, chromatin structural changes, and gene expression alterations in the hybrid lineages derived from two ecologically isolated fish species, *Megalobrama amblycephala* and *Culter alburnus*. These species, which exhibit significant differences in diet and body morphology (Chen et al., 2022; Sun et al., 2021), are naturally isolated in the wild. However, through laboratory hybridization and self-crossing, we successfully produced reciprocal hybrid lineages (F_1–F_3). Previous studies demonstrated that the hybrid offspring exhibited gonadal development, gamete production, fertilization, and hatching success comparable to those of their parental species (Wu et al., 2020; Xiao et al., 2014).

Hybrid males in XY systems are particularly prone to sterility, a phenomenon explained by Haldane's Rule, due to genetic incompatibilities between divergent sex chromosomes that disrupt spermatogenesis and reduce reproductive fitness (Hurst and Pomiankowski, 1991; McDermott and Noor, 2010). Furthermore, hybrid sterility often varies between reciprocal crosses, driven by factors such as maternal inheritance of mitochondria, differences in sex chromosomes, and epigenetic modifications (Polovina et al., 2020). By examining hybrid males from reciprocal lineages, we aim to gain a comprehensive understanding of the genetic and epigenetic mechanisms underlying hybrid fertility.

RESULTS

Structural variations in the testis of F_1 hybrid populations

To obtain high-quality genomes, we generated 111.8 Gb of long-read genome data (28× read depth) using testicular cells from reciprocal F_1 hybrids of *Megalobrama amblycephala* and *Culter alburnus* (BT-1 and TB-1) (Tables S1 and S2). These data produced genome assemblies with contig N50 lengths of 13.26 Mb (BT-1) and 12.40 Mb (TB-1) (Table S3). Using 844.63 Gb of Hi-C reads from the same individuals, we anchored the assembled sequences to 48 pseudochromosomes with integration efficiencies of 97.35% and 95.79% for BT and TB, respectively (Figure S1, Tables S3 and S4). Genome collinearity analyses for the 24 chromosomes of the parental species *M. amblycephala* (blunt snout bream, BSB) and *C. alburnus* (topmouth culter, TC) confirmed that the assembled BT and TB hybrid genomes retained sequences from both parental genomes (Figure 1A).

Genome collinearity analysis also revealed the presence of genomic variations in the testicular cells of F_1 hybrids. To characterize the distribution of these variations within the hybrid population, we performed PacBio HiFi sequencing on three F_1 hybrid progenies of BT and TB lineages, respectively. A total of 105 Gb of sequencing data with an average 9× read depth was obtained (Table S2). Analysis of the mapped files revealed a total of 193,978 structural variations (SVs) in all eight hybrid individuals, including 58,391 insertions, 131,204 deletions, 3,149 duplications, 1,080 translocations, and 154 inversions (Figure 1B; Table S5). SV counts among individuals ranged from 26,749 to 82,522 (Figure S2 and Table S6). Hybrid offspring exhibited distinct SVs, particularly in translocation events between the two subgenomes, with 11–156 events identified in these eight individuals (Figure 1B; Table S7). Our study demonstrates that rapid structural variations, primarily involving genomic deletions, occur in the testicular cells of F_1 progeny following hybridization.

Genomic variations for different genes

To investigate the potential impact of genomic variations on physiological functions, we conducted a statistical analysis of the distribution of SVs in three categories of genes: allelic genes, orphan genes (OGs), and testis-specific genes (TSGs). Allelic genes, representing conserved orthologs from the parental species BSB and TC, are likely to play a more critical role in maintaining core biological processes due to their functional similarity (Bell et al., 2013). OGs are unique to each parental species and may be more involved in species-specific adaptations and evolutionary divergence (Tautz and Domazet-Lošo, 2011). TSGs are among the most highly expressed genes in the testis, playing an indispensable role in testicular development and spermatogenesis (Ong and Corces, 2011). We identified 23,940 allelic genes, 1,090 OGs, and 2,718 TSGs in the hybrid lineages (Table S8).

By examining the distribution of SVs in allelic genes, OGs, and TSGs, we found that the occurrence probability of SVs in allelic genes and TSGs is significantly lower than that in the other genes (*t*-test: $P<0.001$; Pearson chi-squared test: $P<0.001$) (Figure 1C). Specifically, SVs occurred in the *plekhg4* gene in 3 out of 8 hybrid individuals, while SVs were observed in the *gpr179* and *patl2* genes in 2 out of 8 hybrid individuals (Figure S3). This suggests that a high level of SVs in these genes may lead to deleterious mutations in individual development and testicular function. In contrast, we discovered that the occurrence probability of SVs in OGs was significantly higher than the occurrence frequency of SVs in the other genes (*t*-test: $P<0.001$; Pearson chi-squared test: $P<0.001$) (Figure 1C). Our analysis of SVs revealed a distinct distribution pattern across different genes (Figure 1D). SVs derived from OGs were predominantly located downstream of genes, whereas those arising from allelic genes

https://doi.org/10.1007/s11427-024-2868-y

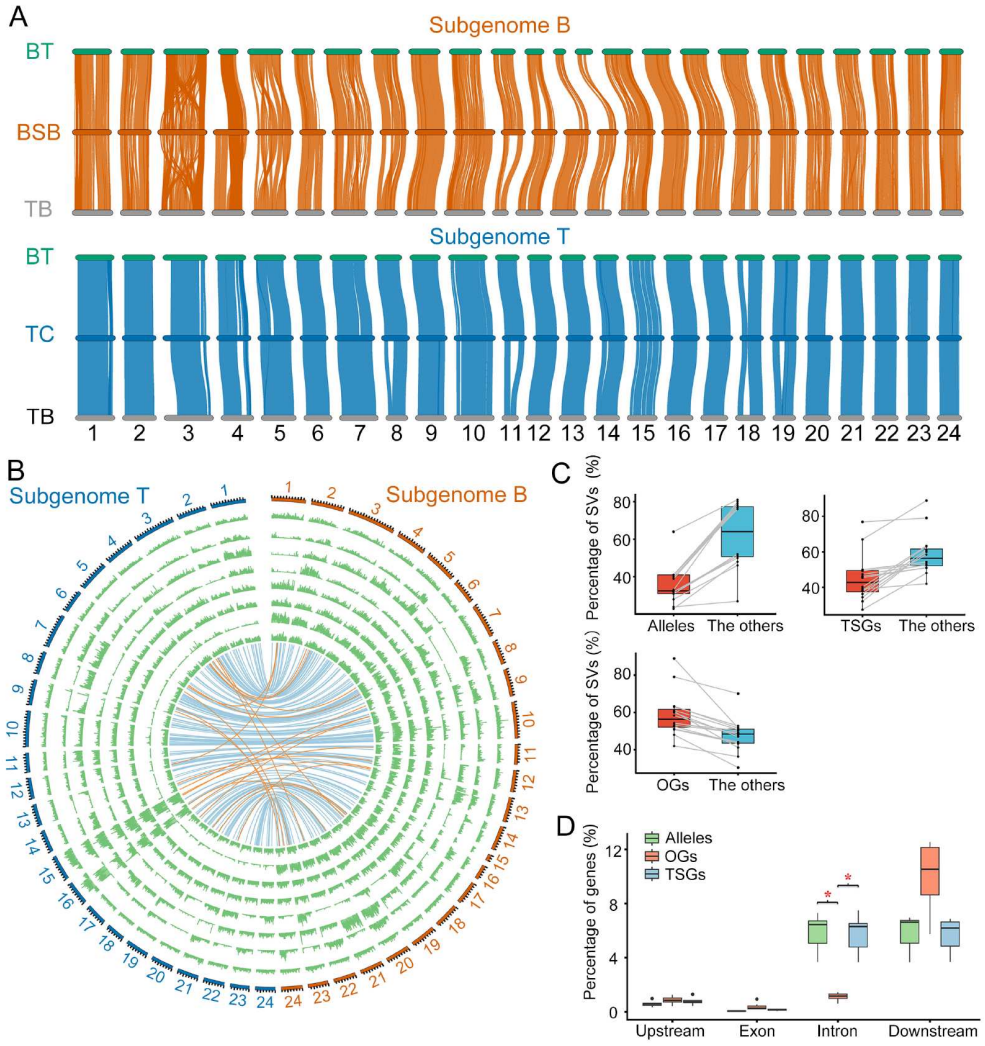

Figure 1. Genome synteny and genetic variations in testes of hybrids. A, Genome synteny of the parental species *M. amblycephala* (BSB) and *C. alburnus* (TC), and their reciprocal F_1 hybrid (BT and TB). B, Genomic variations across different chromosomes. SVs (from the outer to the inner ring, the samples are TB-1 to TB-4 and BT-1 to BT-4). Translocation events between the subgenomes are highlighted with color lines in the central region of the Circos plot. The blue line represents a translocation event between homologous chromosomes, while the orange line represents a translocation event between non-homologous chromosomes. C, Comparative analyses of SV frequencies between allelic, orphan, and testis-specific genes and other genes. The gray line represents the percentages of SVs found in either subgenome B (originated from species BSB) or T (originated from species TC) within the same individual. "The others" refers to all genes other than the specific genes currently being discussed. D, SV distribution across gene categories (allelic genes, OGs, and TSGs) and genomic regions (upstream, exon, intron, and downstream). "*" represents significant differences in intronic regions (*t*-test: $P<0.001$).

and TSGs were found in both downstream and intronic regions. This lower frequency of SVs within orphan gene introns suggests a potential selective pressure against genomic variations in these regions.

Gene expression changes across different tissues and organs

To investigate the effects of hybridization on various tissues and organs, we collected 72 samples (three biological replicates each from six tissues or organs) from BSB and TC, the reciprocal F_1 hybrids (BT and TB) (Tables S1 and S9). Gene expression clustering analysis revealed that hybrid progenies (BT and TB) exhibited similar expression patterns across different organs or tissues due to their similar genome composition (Figure 2A; Table S10). This pattern is distinct from that observed in the parents (BSB and TC). In muscle tissues, the hybrid progenies displayed expression patterns more similar to the parental BSB. This finding suggests a potential for incomplete dominance of the subgenome B (originating from species BSB) in the muscle tissue of both reciprocal hybrid progenies (Figure 2A).

To elucidate the gene expression patterns following hybridization, we conducted differential expression analyses between hybrid progenies and their parents (Figure 2B). Our analysis across six tissues (muscle, testis, intestine, liver, kidney, and brain) revealed a significantly higher proportion of differentially expressed genes (DEGs) in subgenome T (originated from species TC) (246–2,015 genes, 1.76%–18.06%) compared with subgenome B (76–500 genes, 0.62%–3.32%) (*t*-test: $P<0.001$) (Figure 2B; Figure S4). Furthermore, we investigated the distribution of DEGs within allelic genes and OGs. Our analysis showed that allelic genes were less likely to be differentially expressed across different tissues (Pearson chi-squared test: $P<0.01$) (Figure 2C). Conversely, OGs exhibited a propensity for

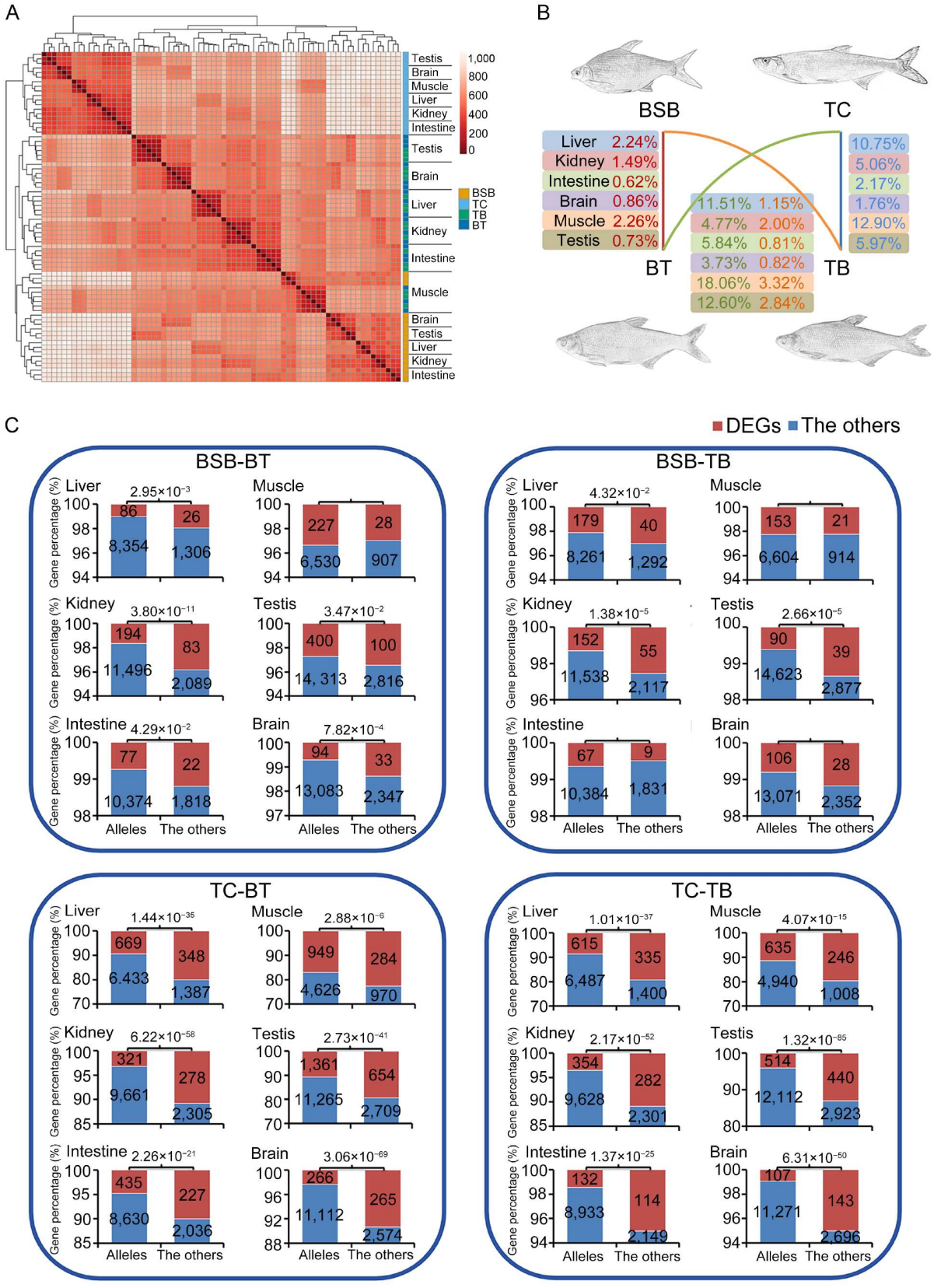

Figure 2. Gene expression analyses across six tissues and organs. A, Gene expression profiles of the parental BSB and TC and their hybrid progenies. B, DEGs identified in four comparisons between the parental BSB and TC and the corresponding subgenome in hybrids. C, Pearson chi-squared test investigating the association between allelic genes and DEGs. "BSB-BT" represents a comparison between the parental BSB genome and the subgenome B of the hybrid BT. Figures display *P*-values if the value is less than 0.05. "The others" refers to all genes other than the specific genes currently being discussed.

https://doi.org/10.1007/s11427-024-2868-y

occasionally increased differential expression within the intestine, liver, kidney, and brain tissues compared with other genes (Pearson chi-squared test: $P<0.05$) (Figure S5).

Chromatin accessibility shifts in allelic and orphan genes

To understand how chromatin accessibility changes after hybridization, we performed ATAC-seq to profile chromatin accessibility in the parents (BSB and TC) and their reciprocal F_1 hybrids (BT and TB). These samples coincided with the parts of testis samples used in RNA-seq analyses (Table S1). The high read depth and preferential mapping to transcription start sites (TSSs) and transcription end sites (TESs) indicated high data quality (Tables S11, S12 and Figure S6). Comparative analysis between the parents and their F_1 hybrids revealed a relatively small subset (BT: 2.82%, 4,578 regions; TB: 4.90%, 7,942 regions) of differentially accessible regions (DARs) (Table S13). This observation suggests that most chromatin accessibility patterns are conserved following hybridization. DARs were primarily distributed in subgenome T compared with subgenome B (68.40% in TB, 53.80% in BT), further highlighting differential chromatin accessibility between subgenomes (Figure 3A; Table S13 and Figure S7). Furthermore, a larger proportion of DARs exhibited decreased chromatin accessibility in the hybrid offspring (83.32% in TB and 77.30% in BT) compared with their parental species (Table S13). This finding suggests that hybridization might lead to a predominantly repressive chromatin landscape in these hybrid progenies.

To further investigate the relationship between chromatin accessibility changes and different gene categories, we first identified genes associated with DARs, finding that 5,502 genes in TB and 3,175 genes in BT were potentially regulated by these DARs (Table S14). Next, we performed Pearson correlation analyses to explore how chromatin accessibility changes differed across two gene categories: allelic genes and OGs. The analysis revealed that allelic genes displayed a lower propensity for chromatin accessibility changes (Pearson chi-squared test: $P<0.001$) (Figure 3B). This suggests that genes shared by the parental species (orthologs) tend to exhibit more conserved chromatin accessibility patterns in the hybrid progenies. In contrast, OGs were more prone to chromatin accessibility changes (Pearson chi-squared test: $P<0.001$) (Figure 3C).

Altered chromatin architecture in allelic and orphan genes

To investigate the 3D genome architecture changes following hybridization, we performed *in situ* Hi-C among the parents (BSB and TC) and their F_1 hybrid progenies (BT and TB), which were the same as the testis samples of eight individuals in RNA-seq analyses (Table S1). Approximately 2.81 TB of *in situ* Hi-C data and 3.01 billion valid Hi-C contacts were obtained from the eight samples (Table S15). Euchromatin (A compartments) is generally transcriptionally active, while heterochromatin (B compartments) is repressed (Jerković et al., 2020). We detected an average of 1,733 compartments in them (Figure 3A; Table S16). The analyses on compartmental rearrangements (A/B switches) showed that there were 34.21% changes in comparison between TB and the parents and 31.84% changes in comparison between BT and their parents (Table S17). GO analyses on spermatogenesis (GO: 0007283) reflected that 14 genes in subgenome B and 9 genes in subgenome T exhibited compartmental rearrangements in the hybrid progenies (Table S18).

To further investigate the relationship between compartmental rearrangements and the two gene categories (allelic and orphan genes), we employed Pearson's chi-squared test to compare hybrid progeny and their parents. Transcriptionally active OGs exhibited a significantly lower propensity for rearrangement compared with the parents (Pearson chi-squared test: P-value<0.001) (Figure 3D). This suggests that the specific expression of orphan genes in the ancestral parental species plays a crucial role in testicular development. Any alterations in their expression patterns could lead to abnormal testicular development and disrupted spermatogenesis. Interestingly, transcriptionally repressed OGs demonstrated a higher susceptibility to rearrangement in hybrid progeny compared with the parents (Pearson chi-squared test: P-value<0.001) (Figure 3D). This phenomenon could be attributed to post-hybridization modifications in the genome and transcriptional regulatory systems, potentially rendering certain orphan genes prone to transcriptional activation. However, these changes might not have a significant impact on testicular development and spermatogenesis. In contrast, our analysis of the relationship between compartmental rearrangements and allelic genes revealed a distinct pattern. Transcriptionally active allelic genes were more likely to undergo rearrangement (Pearson chi-squared test: P-value<0.001) (Figure 3D), while transcriptionally repressed allelic genes tended to maintain their repressed state. These findings highlight the importance of maintaining the transcriptional activation status of orphan genes and the transcriptional repression status of allelic genes for normal testicular development and spermatogenesis.

Differential expression of testis-specific genes

Despite exhibiting lower genomic variation than other genes following hybridization, investigating expression changes in TSGs sheds light on the molecular mechanisms underlying testicular development and spermatogenesis in hybrid individuals. A total of 2,503 TSGs were expressed in the subgenomes B and T of hybrids (1,373 in B and 1,130 in T). Analysis of the gene expression revealed distinct patterns in reciprocal hybrids compared with their parents. The BT hybrid displayed minimal differential expression (DE) of TSGs in subgenome B, with a negative correlation observed between TSGs and DEGs (Pearson chi-squared test: $P=0.02$) (Figure S8). Conversely, the TB hybrid exhibited a significant increase in TSGs within subgenome T, showing a positive correlation with DEGs (Pearson chi-squared test: $P<0.001$). These findings demonstrate that TSGs in subgenome B undergo conserved gene expression changes following hybridization.

To elucidate the key regulatory genes regulating reproductive capacity in the hybrid lineages, we further detected three reproductive phenotypes (semen volume, testicular weight, and area percentage) in 19 individuals (24 months after hatching) of TB (Figure 4A and B; Tables S1 and S19). A positive correlation was observed among these three phenotypes (Figure S9). Next, we performed testicular transcriptome sequencing and calculated gene expression levels. Subsequently, we employed weighted gene correlation network analysis (WGCNA) to identify reproductive-related key regulatory genes among TSGs. This analysis yielded 75 and 12 genes exhibiting positive and negative

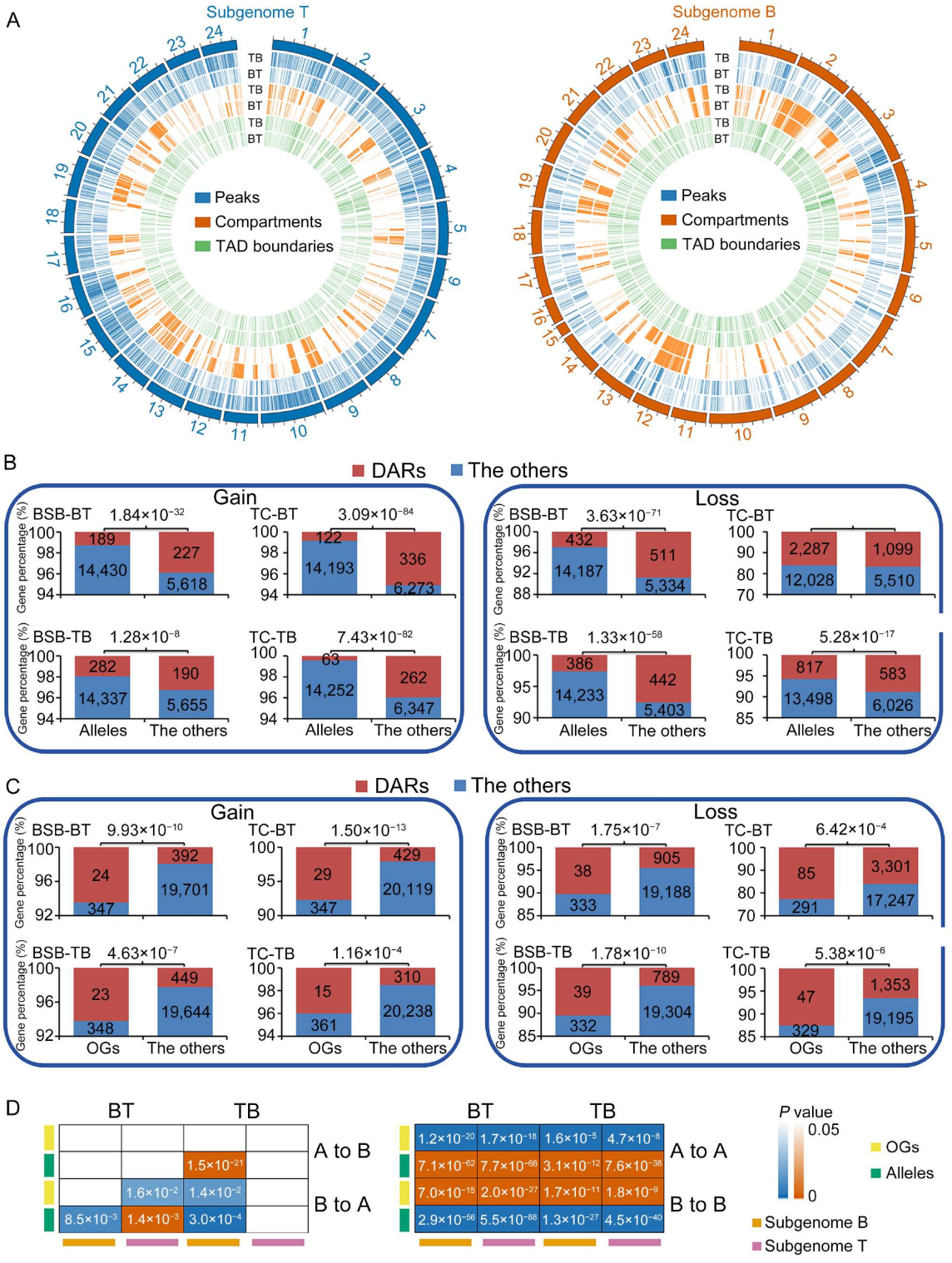

Figure 3. Chromatin remodeling of allelic and orphan genes. A, Density of differential chromatin accessibility regions (red), compartmental rearrangements (blue), and variable TAD boundaries (green) between the parental species and their hybrid progeny. B, Pearson chi-squared test between OGs and genes associated with DARs. "BSB-BT" represents a comparison between the parental BSB genome and the subgenome B of the hybrid BT. "Gain" represents increased chromatin accessibility in subgenome B or T in hybrid offspring (latter) compared with the parental BSB genome (former). Conversely, "Loss" represents decreased chromatin accessibility in hybrid offspring (latter). "DARs" (red) represents genes located within or near DARs, while "The others" (blue) represents all other genes not associated with DARs. The *Y*-axis represents the percentage of genes. *P*-values are shown within the figure if they are less than 0.05. C, Pearson's correlation coefficients between alleles and genes associated with DARs. D, Pearson's correlation coefficients comparing alleles or OGs with the genes associated with compartments A and B. *P*-values in orange indicate positive correlations, while those in blue represent negative correlations. "A to B" represents A compartments in the parental species transitioning to B compartments in the hybrid progeny. Similarly, "B to A" indicates a switch from parental B compartments to hybrid A compartments. "A to A" and "B to B" represent no change in transcriptional state between parental and hybrid compartments. "The others" refers to all genes other than the specific genes currently being discussed.

https://doi.org/10.1007/s11427-024-2868-y

correlations with reproductive capacity, respectively (Pearson correlation coefficient: $r>0.5$, $P<0.05$) (Figure 4C). Among these, we identified *rnf212b*, a gene exhibiting strong correlations with all three reproductive phenotypes (Figure 4D). Furthermore, we validated the expression of *rnf212b* in 19 TB individuals using qRT-PCR and compared the results with transcriptome data. A strong correlation ($R=0.68$, $P=0.0014$) was observed between the two datasets, indicating the reliability of our transcriptome data in assessing *rnf212b* expression (Figure S10). Moreover, the expression level of *rnf212b* was also positively correlated with three reproductive phenotypes: semen volume, testicular weight, and area percentage (Figure S10). Previous studies have implicated *rnf212b* as a pivotal gene in homologous chromosome pairing during vertebrate meiosis (Condezo et al., 2024; Reynolds et al., 2013). Although we did not conduct gene editing experiments to directly verify its role in reproductive capacity within the hybrid lineages, our findings, coupled with the existing literature, suggest a significant contribution of *rnf212b* to testicular function and fertility in the hybrids.

Chromatin remodeling of testis-specific genes

Further exploration revealed a divergence in the correlation patterns between chromatin accessibility changes and TSGs between subgenomes. In subgenome B, a negative correlation was observed between DAR-associated genes and TSGs (Figure S11). This indicates that TSGs in subgenome B exhibited a lower tendency for decreased accessibility in both BT and TB hybrids (Pearson chi-squared test: $P<0.01$) (Figure S11). Conversely, no significant correlation was observed in subgenome T (Figure S11). Our results suggest that TSGs in subgenome B underwent conserved changes in chromatin accessibility following hybridization.

We investigated the relationship between 3D genome architecture changes and TSGs. Compartmental rearrangements between the hybrids and their parents were associated with 602 TSGs in BT and 477 TSGs in TB. Transcriptionally active TSGs (near and within A compartments) displayed a higher propensity for no alteration compared with other genes (Pearson chi-squared test: $P<0.05$) (Figure S12). Meanwhile, non-transcribed TSGs located near or within B compartments in subgenome T displayed a lower propensity for no alterations compared with other genes (Figure S13). Our analysis of topologically associating domain (TAD) boundaries using Hi-C data revealed an average of 1,553 genes located near these boundaries in each sample (Table S20). We found a higher proportion of stable boundaries within subgenome B compared with subgenome T in both BT and TB hybrids (Figure 3A; Table S21). Our findings suggest that continued expression of TSGs may be functionally important for testis function after hybridization.

We used Hi-C data to identify DNA loop anchors. Due to limited sequencing depth, we detected 18,163 loops in the parental BSB genome, 26,126 loops in the parental TC genome, 30,415 loops in the hybrid offspring TB, and 43,638 loops in the hybrid offspring BT (Table S22). Subsequently, we analyzed 87 key genes involved in reproductive regulation, which were identified through WGCNA analysis (Figure 4C). Among these, 26 genes exhibited identical loops in both the parents and offspring, which further supports the accuracy of the identified loops. Interestingly, 16 of these genes showed potential *cis*-interactions (Table S23). Notably, two genes annotated as *klh10* are located on chromosome 3 of the subgenome T (Table S23). Additionally, three unannotated genes located on chromosome 9 of the subgenome B exhibit potential interactions with each other (Figure 5). We hypothesize that these gene interactions, identified through chromatin interaction data, may influence the reproductive system of the hybrid individuals. However, further studies are needed to confirm the functional significance of these gene interactions.

Influence of SVs on TAD boundaries and allelic gene expression

SVs are considered a potential driving force behind alterations in TAD boundaries (Spielmann et al., 2018). Our study demonstrates the co-existence of SVs and variable TAD boundaries within the testes of hybrid individuals. To elucidate the potential impact of SVs on TAD boundaries, we performed an overlap analysis of SV and TAD data from a single individual. This analysis revealed that 1,702 variable TAD boundaries overlapped with 2,251 SVs in TB hybrids and 1,552 variable TAD boundaries overlapped with 2,011 SVs in BT hybrids. Further analysis identified 1,061 shared SVs and 773 shared variable TAD boundaries across all hybrids (TB and BT) (Figure S14). Additionally, 2,087 genes were identified in BT hybrids and 2,323 genes in TB hybrids, all located in proximity to variable TAD boundaries associated with SVs (Tables S24 and S25). These genes were designated as potential targets for SV-mediated variable TAD boundaries.

To delve deeper into the potential influence of SV-mediated variable TAD boundaries on gene expression patterns, we employed Pearson's chi-squared tests to investigate which genes (alleles, OGs, and TSGs) from our previous analysis might be affected. Our results revealed a significantly lower proportion of allelic genes within regions overlapping both SVs and variable TAD boundaries, compared with the overall proportion of allelic genes in the hybrid individual (Pearson chi-squared test: P-value<0.001) (Figure 6). This observation suggests that allelic gene expression is less susceptible to disruptions caused by SV-mediated variable TAD boundaries. In contrast, no significant correlation was observed for OGs and TSGs (Figures S15 and S16).

DISCUSSION

Reproductive isolation is a fundamental driver of speciation, often associated with increased genetic divergence among species. However, in freshwater fish, ecological or geographical isolation can sometimes outweigh genetic divergence in driving speciation (Cohen et al., 2016). Understanding the genetic mechanisms underlying the absence of postzygotic reproductive isolation in species with high genetic divergence can provide critical insights into hybridization's role in evolution and speciation. This study focuses on hybrid lineages derived from the intergeneric hybridization of *M. amblycephala* and *C. alburnus*, two species that diverged approximately 12.74 million years ago but exhibit ecological isolation (Ren et al., 2019; Xiao et al., 2016). Using advanced genomic and epigenomic tools, we characterized structural variations, chromatin organization, and gene expression to explore the molecular mechanisms that sustain fertility in these hybrids.

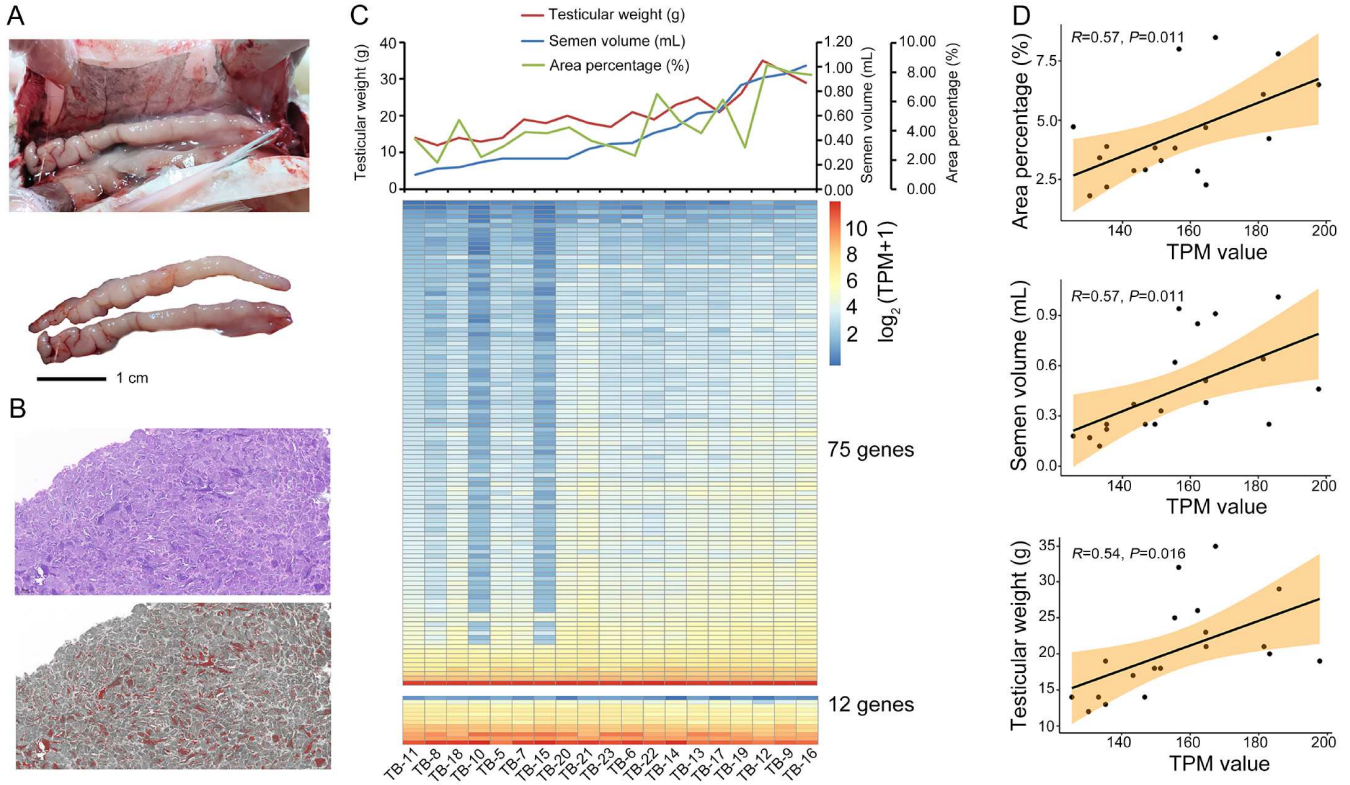

Figure 4. Comprehensive analysis of testicular phenotype and potential regulatory genes for reproductive capacity in 19 TB individuals. A, Appearance of the testes in individual TB-6. B, Hematoxylin and eosin (HE) staining of the testis in TB-6 (upper image) and the corresponding image with sperm marked in red (lower image). C, WGCNA analysis reveals reproductive gene expression in 19 individuals of TB. The relationships between three reproductive phenotypes (semen volume, testicular weight, and area percentage) and the expression of 87 genes potentially involved in regulating reproductive capacity. D, A positive correlation between the expression level of the gene *rnf212b* (transcriptome data) and the three reproductive phenotypes.

Structural and functional integrity of allelic genes

Allelic genes, conserved between the parental genomes, exhibited minimal SVs, stable chromatin accessibility, and conserved expression patterns in the hybrid testes. These characteristics likely reflect selective constraints to preserve essential biological processes. The reduced genomic variation and conserved chromatin architecture of allelic genes likely mitigate the impact of deleterious mutations, thereby supporting normal testicular development and spermatogenesis in hybrids. These findings align with previous studies suggesting that allelic genes play a foundational role in maintaining cellular and developmental processes (Powell et al., 2020; Ren et al., 2019). The observed stability in allelic genes highlights their critical role in counterbalancing the genomic instability introduced by hybridization. By ensuring the functional continuity of essential genes, hybrids can maintain reproductive capabilities despite substantial genomic divergence. These results directly address the scientific question of how hybridization can avoid postzygotic reproductive isolation despite genetic divergence.

Orphan genes and their adaptive potential

Orphan genes, unique to each parental lineage, showed higher susceptibility to SVs and chromatin accessibility changes compared with allelic genes and TSGs. This susceptibility is likely due to reduced functional constraints and greater evolutionary plasticity, which may allow hybrids to explore novel adaptive landscapes. The increased genomic variation in orphan genes may reflect their role in species-specific functions, as previously suggested (Tautz and Domazet-Lošo, 2011). Furthermore, the distinct expression patterns of orphan genes across tissues suggest their potential involvement in tissue-specific adaptations, adding to the hybrid's phenotypic diversity.

The higher variability in orphan gene expression and chromatin remodeling indicates that these genes are more prone to hybridization-induced genomic shocks. However, the ability of orphan genes to tolerate genomic and regulatory disruptions without severely impacting survival suggests an evolutionary mechanism that prioritizes adaptability over stability in these regions. This finding contributes to our understanding of how hybrids harness genetic divergence to maintain viability and adaptability.

Testis-specific genes and hybrid fertility

TSGs, essential for spermatogenesis and testicular development, exhibited remarkable genomic and transcriptional stability in fertile hybrids (Liu et al., 2011). This stability underscores their critical role in maintaining reproductive capacity. Despite the pervasive genomic variations introduced by hybridization, TSGs were among the least affected gene categories, demonstrating lower SV frequencies, minimal chromatin remodeling, and conserved expression levels. These findings strongly suggest that TSGs are under stringent selective pressures to preserve their

https://doi.org/10.1007/s11427-024-2868-y

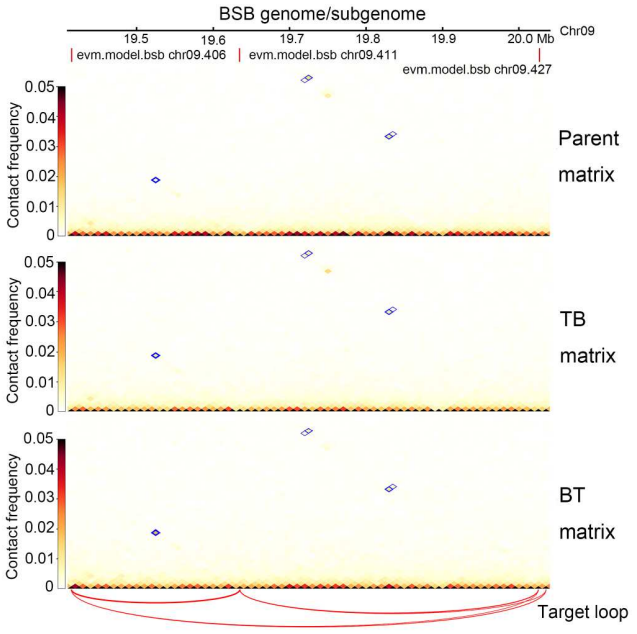

Figure 5. Shared Hi-C loops among the paternal species (BSB and TC) and their hybrid offspring (BT and TB). The genes "evm.model.bsb_chr09.406", "evm.model.bsb_chr09.411", and "evm.model.bsb_chr09.427" are located on chromosome 9 of the subgenome B. Blue boxes indicate interaction signal points.

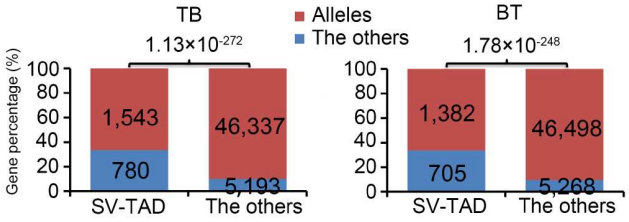

Figure 6. Genomic variations affecting TAD boundaries and allelic gene expression. Pearson chi-squared test investigating the association between allelic genes and the genes associated with SV-mediated variable TAD boundaries. "The others" refers to all genes other than the specific genes currently being discussed.

functional integrity, as disruptions to these genes could lead to infertility.

One of the most significant findings in our study was the identification of *rnf212b* as a key regulator of reproductive traits in hybrids. The expression of *rnf212b* was strongly correlated with semen volume, testicular weight, and sperm area, highlighting its potential role in ensuring meiotic chromosome pairing and fertility. Previous studies have implicated *rnf212b* in meiosis (Condezo et al., 2024; Reynolds et al., 2013), and our findings extend its significance to hybrid systems. Future functional validation using gene editing could confirm its role in maintaining hybrid fertility and provide further insights into its broader role in reproductive isolation mechanisms.

Epigenomic mechanisms: balancing stability and plasticity

Our analysis of chromatin accessibility and 3D genome architecture revealed a delicate balance between stability and plasticity in the hybrid genome. While chromatin accessibility patterns were largely conserved, a subset of regions exhibited differential accessibility, primarily in subgenome T. The decreased accessibility in these regions reflects a shift towards a repressive chromatin state, potentially serving as a compensatory mechanism to mitigate the effects of genomic instability.

Similarly, compartmental rearrangements in the hybrid genome showed distinct patterns between allelic genes and orphan genes. While allelic genes maintained their transcriptional states, orphan genes were more prone to chromatin remodeling and compartmental shifts, particularly in transcriptionally repressed regions. This differential behavior suggests a regulatory mechanism that prioritizes the stability of essential genes while allowing flexibility in less critical regions.

Implications for hybrid evolution and reproductive isolation

The findings of this study underscore the unique genetic and epigenetic strategies employed by cyprinid hybrids to maintain fertility despite substantial genetic divergence. The selective stabilization of allelic genes and TSGs ensures the preservation of core biological processes and reproductive function. In contrast, the higher variability observed in orphan genes reflects an adaptive mechanism that allows hybrids to harness genetic diversity for phenotypic plasticity.

These insights contribute to a broader understanding of how hybridization drives evolutionary processes, including species diversification and adaptation. The ability of hybrids to balance genomic instability with functional stability may provide a template for studying hybrid systems across other taxa. Furthermore, this study highlights the potential role of hybridization in facilitating adaptive evolution by reshaping gene regulatory networks and chromatin architecture.

MATERIALS AND METHODS

Sample collection

After 24 months of being raised in identical, controlled environments following hatching, sexually mature male fish with diploid levels were randomly collected from each group: three individuals of *Megalobrama amblycephala* (blunt snout bream, BSB); three individuals of *Culter alburnus* (topmouth culter, TC); four F_1 hybrid individuals resulting from the hybridization of female BSB and male TC (BT); 23 F_1 hybrid individuals resulting from the hybridization of female TC and male BSB (TB) (Ren et al., 2019; Xiao et al., 2016). All fish were obtained from the Engineering Center of Polyploid Fish Breeding of the National Education Ministry in Changsha, Hunan, China.

Before tissue collection, the ploidy level of each individual was determined using flow cytometry. The DNA content of erythrocytes of BSB, TC, BT, and TB was measured using flow cytometry (Cell Counter Analyzer, Partec, Germany). About 0.5–1 mL of blood from each fish was collected from the caudal vein using a syringe containing 100–150 units of sodium heparin. The blood samples were treated according to the method described by Xiao et al. (2014). The DNA contents of BSB and TC were used as controls. To facilitate tissue collection, all individuals were deeply anesthetized with 300 mg L^{-1} tricaine methanesulfonate (Sigma-Aldrich, USA) for 30 min at 20°C in a separate tank. Death was confirmed according to standard protocols before dissection. Six tissues (muscle, brain, liver, intestine, kidney, and testis) were then collected from each individual. To stimulate sperm release,

the male fish's abdomen was gently massaged. The ejaculate was then collected in a container for semen volume assessment. For testis collection, all testicular tissues were harvested and weighed using a weighing scale.

Hematoxylin and eosin staining

A 10-mm-thick section of testis was dissected from each of the 19 hybrid individuals of TB. The tissue samples were fixed in Bouin's solution for 24 h, followed by washing with distilled water for 4 h. After fixation, the tissues were dehydrated through a graded ethanol series (70%, 80%, 90%, and 100%) and embedded in paraffin blocks. The paraffin-embedded tissues were sectioned into 10-μm-thick slices using a microtome. Hematoxylin and eosin (HE) staining was performed on the tissue sections following the manufacturer's protocol with an HE staining kit. Digital images of the stained sections were captured using a DX8 microscope (Olympus, Japan). To quantify the sperm content in the HE-stained testis sections, ImageJ software was used. The sperm content index was calculated as the ratio of the area occupied by sperm to the total testis area in each section.

DNA isolation and whole-genome sequencing

High-quality genomic DNA was isolated from the testis of eight individuals (four individuals for each BT and TB) using the QIAGEN® Genomic Kit following the manufacturer's protocol. The extracted DNA was assessed for degradation and contamination by gel electrophoresis on 1% agarose gels. Additionally, DNA purity (260/280 and 260/230 ratios) and concentration were measured using a NanoDrop™ One UV-Vis spectrophotometer (Thermo Fisher Scientific, USA) and a Qubit® 4.0 Fluorometer (Invitrogen, USA), respectively.

High-quality genomic DNA from the eight individuals was used to construct SMRTbell HiFi libraries following PacBio's standard protocol (15 kb prep solutions) (Pacific Biosciences, USA). The library preparation process involved the following steps: (i) shearing: the genomic DNA was fragmented using g-TUBEs (Covaris, USA). (ii) A-tailing: an A-tailing reaction was performed to add overhangs to the fragmented DNA. (iii) Adaptor ligation: the fragments were ligated with the hairpin adaptor using the SMRTbell Express Template Prep Kit 2.1 (Pacific Biosciences). (iv) Purification: the library was treated with nuclease to remove unwanted fragments and then purified using AMPure PB beads. (v) SMRTbell library purification: the purified library was further purified using PB beads. (vi) Quality control: the final high-quality library was assessed for fragment size using the Agilent 2100 Bioanalyzer (Agilent Technologies, USA). (vii) Following library construction, sequencing was performed on the PacBio Revio platform with a loading concentration of 120 pmol L^{-1}. SMRTlink (v. 7.0) was used to filter and process the raw sequencing data, removing reads shorter than 50 bp.

We performed whole-genome re-sequencing on the genomic DNA of the two individuals (BT-1 and TB-1) (Table S1). High-quality DNA samples were used to prepare single-stranded circular libraries. These libraries were then transformed into DNA nanoballs (DNBs), each containing millions of copies of the circular DNA template. The DNBs were loaded onto patterned nanoarrays for combinatorial probe anchor synthesis (CPAS) sequencing. Finally, DNBSEQ-T7 sequencing was conducted using a paired-end 150 bp×2 read length following the manufacturer's protocol. Adapter sequences and low-quality bases were filtered from the reads using Fastp (v. 0.21.0) (Chen et al., 2018) before assembly.

Genome assembly and chromosomal organization

Highly accurate long sequencing reads (HiFi reads) from the two individuals (BT-1 and TB-1) were used for genome assembly with hifiasm software (v0.15.4-r347), respectively (Cheng et al., 2022). Hi-C libraries were constructed from testicular cells of BT and TB (the same sample with the genome assembly analyses). Briefly, the protocol involved formaldehyde fixation, cell lysis, and cross-linked DNA digestion with MobI enzyme. Biotinylated sticky ends were then generated, followed by proximity ligation to create chimeric junctions. These junctions were enriched, size-selected (300–700 bp), and processed into paired-end sequencing libraries. Clean Hi-C reads were obtained by trimming adapter sequences and low-quality reads. Further processing involved truncating reads at putative Hi-C junctions and aligning them to the assembled genomes using BWA (v. 0.7.17) (Li and Durbin, 2010). HiC-Pro (v. 2.8.1) (Servant et al., 2015) was employed to filter invalid read pairs (dangling ends, self-cycles, religations, and dumped products). Finally, LACHESIS (Release: 2017-12-21) (Burton et al., 2013) was used to correct scaffolds, cluster them, and orient them into chromosomes. Following this step, manual adjustments were made to address any placement or orientation errors identified through chromatin interaction patterns.

Structural variation detection

The data obtained from PacBio HiFi was aligned to the combined genome files of BSB and TC using NGMLR (v. 0.2.1) (https://github.com/philres/nextgenmap-lr.gitcdngmlr) with the default parameters. SVs were detected based on the aligned data for each sample using two established SV callers: Sniffles (v. 2.2) and cuteSV (v. 2.1.1). The SURVIVOR module within PBSV (v. 1.0.7) was employed to merge the SV calls from Sniffles and cuteSV, generating a high-confidence consensus set of SVs. The consensus set of SVs with a minimum length of 50 bp was subjected to comprehensive genomic annotation using ANNOVAR (Yang and Wang, 2015) (released in 2019), providing valuable insights into the functional and regulatory impact of the identified variants.

RNA isolation and library preparation

Total RNA of BSB, TC, BT, and TB was isolated and purified from the brain, liver, intestine, muscle, kidney, and testis (three biological replicates) using the TRIzol extraction method (Rio et al., 2010), respectively. The RNA concentration was measured using NanoDrop technology. To remove contaminating genomic DNA, samples were treated with DNase I (Invitrogen). The purified RNA was then quantified using a 2100 Bioanalyzer system (Agilent Technologies). For transcriptome sequencing, isolated mRNA was fragmented using a fragmentation buffer, and the resulting short fragments were reverse transcribed and amplified to cDNA. Transcriptome data were obtained for 72 samples, representing three biological replicates each from six organs or tissues collected from four groups (BSB, TC, BT, and TB). The DNBSEQ-T7 DNA nanoball technology was employed

following the standard protocol described by Patterson et al. (2019). The library preparation process involved the following key steps: (i) single-stranded circular library preparation: MGI Library Prep Kits were used for this step. (ii) CPAS: following DNA anchor hybridization, a fluorescent probe was attached to the DNA nanoball. (iii) High-resolution imaging: a high-resolution imaging system captured the fluorescent signal. (iv) Digital processing and sequencing: after digital processing of the optical signal, the sequencer generated high-quality and accurate sequencing information. (v) Low-quality bases and adapter sequences were trimmed from the reads using Fastp (v. 0.21.0) (Chen et al., 2018).

Gene expression analysis

Clean reads from BSB, TC, BT, and TB were mapped to the combined genome files of BSB and TC using STAR (v. 2.7.10) (Dobin et al., 2013) with default parameters. Mapped reads were processed with SAMtools (v. 1.10) (Li et al., 2009) to obtain uniquely mapped reads, which were then quantified using htseq-count (v. 0.12.4) (Srinivasan et al., 2020). Gene expression values for mRNA-seq were normalized and calculated as transcripts per million (TPM). Genes with fewer than five mapped reads in any sample were excluded from further analysis. DE analysis was performed using the DESeq2 package in R with a significance threshold of P-value<0.001. To compare gene expression between the parents (BSB or TC) and their hybrid progenies (BT or TB), DE analyses were restricted to the corresponding subgenome. For instance, in the DE analysis between BSB and BT, only read counts mapped to subgenome B (originating from species BSB) were used for calculations. Gene Ontology (GO) enrichment analysis was performed with a significance threshold of Benjamini-Hochberg false discovery rate (FDR)<0.05.

Identification of allelic, orphan, and testis-specific genes in hybrids

In F_1 hybrids resulting from the hybridization between BSB and TC, alleles were derived from orthologs of the parental species. Allelic sequences for the subgenomes B (BSB-derived) and T (TC-derived) in the hybrid were obtained using an all-against-all reciprocal BLASTP (v 2.8.1) search with a stringent e-value cutoff of 1×10^{-6} and a minimum sequence alignment of 70% at the protein level. Subsequently, transcripts shorter than 300 bp from allelic genes were excluded. Some genes, lacking detectable homologs in other known species or lineages, are classified as orphan genes. Orphan genes were identified using stringent BLASTx (against protein databases) and tBLASTx (against nucleotide databases) searches against public databases with an e-value threshold of 1×10^{-5}. Genes lacking significant BLAST hits were considered potential orphan genes. Among these potential orphan genes, only those with a transcript abundance (TPM) greater than 10 in the corresponding organ were classified as expressed orphan genes. TSGs are genes with significantly higher expression levels in the testis compared with other tissues and organs. To identify TSGs, we employed a rigorous criterion. Genes were considered TSGs if their expression in the testis was significantly higher compared with all five other analyzed tissues and organs (DESeq2 package, P-value<0.001).

Weighted gene correlation network analysis on testis-specific genes

Total RNA was isolated and purified from the testes of 19 individuals using the TRIzol extraction method (Rio et al., 2010). We performed library preparation, DNBSEQ-T7 sequencing, and gene expression analysis for all 19 individuals as described above. To identify key genes involved in testicular development and spermatogenesis, we performed weighted gene correlation network analysis (WGCNA, v. 1.67) on TSGs from all 19 samples. An unsupervised coexpression network was constructed using the following steps: (i) Correlation Matrix Generation: a Pearson correlation matrix was generated based on the gene expression values. (ii) Adjacency Matrix Construction: an adjacency matrix representing the connection strength between genes was built by raising the correlation matrix to a soft-threshold power. (iii) Topological Overlap Matrix Calculation: the adjacency matrix was used to calculate a topological overlap matrix, which captures the shared network connections between genes. (iv) Module Identification: genes with similar coexpression patterns were clustered using hierarchical clustering based on the topological overlap matrix. (v) Module-Phenotype Association: Pearson correlations were calculated between the eigengene (a representative gene for each module) and the expression levels of individual genes within the same module. Additionally, Pearson correlations were performed to assess the strength and direction of association between modules and three reproductive phenotypes: semen volume, testicular weight, and area percentage. Key genes potentially involved in testicular development and spermatogenesis were identified from the two most significant modules based on two criteria: module membership >0.5 and gene significance >0.5.

ATAC-seq library preparation and sequencing

Eight testis samples (two biological replicates each from BSB, TC, BT, and TB) were used for ATAC-seq analysis, following the same samples used for mRNA-seq. Testes were washed with 0.09% NaCl and cryogenically ground into a powder. The powder was lysed, incubated at 4°C for 10 min, and filtered through a 40 μm cell strainer. Nuclei were washed with cold PBS and verified for purity and integrity under a microscope. Approximately 50,000 nuclei were allocated for tagmentation following established protocols (Corces et al., 2017). Tn5-transposed DNA fragments were purified using AMPure DNA magnetic beads. qPCR was used to determine the optimal number of PCR cycles. The amplified libraries were assessed for quality on an Agilent Tapestation 2200 with a D5000 DNA ScreenTape. The final libraries were sequenced on an Illumina NovaSeq 6000 sequencing platform (150 bp paired-end reads).

ATAC-seq data processing and analysis

Raw sequencing data underwent quality control using Cutadapt (v. 1.9.1). Clean reads were aligned to the combined files of BSB and TC genome data using Bowtie2 (v. 2.2.6) (Langmead and Salzberg, 2012). Mitochondrial reads and PCR duplicates were removed. Two key metrics were used to assess data quality: insert size distribution and enrichment of reads around TSSs. The signal at the center of the insert size distribution served as the TSS enrichment metric after library normalization.

Biological replicates were used for peak calling with MACS2 (--nomodel --extsize 200 --shift -100) to identify regions with high confidence (Zhang et al., 2008). Peaks identified in at least three replicates were considered to have high-to-medium confidence. Stringent irreproducible discovery rate analysis (FDR<0.05) was performed on peaks across replicates. A consensus peak set was obtained by merging peaks from all samples. DARs were identified using the DESeq2 R package with a significance threshold of FDR<0.05 and a $\log_2$ fold change (FC) threshold of >1. Motif analysis of DARs was performed using the MEME suite with default settings (*P*-value<0.01) (Heinz et al., 2010). All sequencing data tracks were visualized with the Integrated Genomic Viewer (IGV, v. 2.3.61).

Hi-C library construction

Eight testis samples (two biological replicates each from BSB, TC, BT, and TB) were used for chromosome interaction analysis, following the same samples used for ATAC-seq and mRNA-seq. The Hi-C library was constructed based on the next steps: (i) the sample was cross-linked with 1% formaldehyde (final concentration) for 10 min, and the cross-linking reaction was quenched with 0.2 mol L^{-1} glycine (final concentration) for 5 min. Then the cross-linked cells were lysed in lysis buffer (10 mmol L^{-1} Tris-HCl pH 8.0, 10 mmol L^{-1} NaCl, 0.2% NP-40, and complete protease inhibitors), and nuclei were isolated by centrifugation and resuspend in 150 μL of 0.1% SDS. (ii) The nuclei were first incubated at 65°C for 10 min, after SDS was quenched with 120 μL of water and 30 μL of 10% Triton X-100, they were incubated at 37°C for 15 min. Then the nuclear DNA was digested with 30 μL of 10× NEB Buffer 2.1 and 150 U of MboI restriction enzyme at 37°C overnight, and the MboI enzyme was inactivated at 65°C for 20 min. (iii) The cohesive ends were filled in with dNTPs and biotin-labeled dCTP using Klenow fragments at 37°C for 2 h, and the blunt ends were ligated using T4 DNA ligase at 16°C for 4 h. After reversing the cross-linking with proteinase K at 65°C overnight, the DNA was purified using a QIAamp DNA Mini Kit following the manufacturer's instructions. At last, the purified DNA was sheared to a fragment size of approximately 400 bp, and Hi-C libraries were constructed for all eight samples for sequencing on the DNBSEQ-T7 sequencing platform (150-bp paired-end reads).

Hi-C data preprocessing and normalization

Following quality filtering with Fastp (v. 0.23.2) (Chen et al., 2018), clean Hi-C reads were processed using the Distiller pipeline (https://github.com/mirnylab/distiller-nf). This pipeline mapped reads to the combined files of BSB and TC genomes, filters dangling ends, and other unusable data. Valid read pairs were then used to assess the correlation efficiency of biological replicates for each sample using HiCRep (Yang et al., 2017). Data from replicates were subsequently pooled for further analysis. Raw Hi-C contact maps often contain biases due to factors like mappability, GC content, and uneven restriction enzyme site distribution. To address these biases, an iterative normalization method was applied to the contact maps. The "matrix resolution" of a Hi-C map refers to the bin size used to construct the contact matrix. We define "map resolution" as the smallest locus size at which at least 80% of the bins contain a minimum of 1,000 contacts.

Comprehensive analysis of chromatin compartments, TADs, and chromatin loops

Compartments are defined as groups of chromosomal domains that exhibit increased interaction with each other, either on the same chromosome or across different chromosomes. In heat maps generated from 100 kb bins, this appears as a distinctive "plaid pattern" with alternating blocks of high and low interaction frequencies representing A and B compartments. Principal component analysis (PCA) was employed to identify these compartments. Each chromosomal bin was assigned to compartment A or B based on the sign (positive or negative) of its first eigenvector (PC1). TADs are contiguous regions with high self-association, separated from neighboring regions by distinct boundaries. TAD locations can be determined by analyzing interaction data binned at 40 kb resolution. We used the insulation score algorithm implemented in Cooltools (v. 0.4.1) to identify TAD boundaries and determine the location and number of TADs within each sample.

The contact matrices were transformed into *Z*-score matrices. This transformation allows for direct comparison between the two matrices. Cooltools was then used to compare TAD boundaries between two samples. TAD boundaries were classified into five categories based on changes in their strength between samples: stable, appeared, disappeared, weakened, and strengthened. Weakened boundaries exhibited a strength reduction of at least twofold, while strengthened boundaries showed a strength increase of at least twofold.

Intra-chromosomal and inter-chromosomal interactions were analyzed at a 5 kb resolution. Interaction data for each sample were processed using Fit-Hi-C software to calculate cumulative probabilities (*P*-values) and FDRs. Interactions were considered significant if both the *P*-values and FDR were less than 0.01 and the contact count exceeded 2.

qRT-PCR experiments and analyses

Total RNA from the testes of 19 TB individuals that underwent transcriptome sequencing was subjected to qRT-PCR experiments. Primers for beta-actin (F: 5′-GGAATCTAGCGGTATCCACGA-3′, R: 5′-GGTCAGCAATGCCAGGGTA-3′) and *rnf212b* (F: 5′-AAAGGCAGAAGGAGCGAGTC-3′, R: 5′-GCGGCTTCTTTAGTTCTTCGT-3′) were designed using Primer Premier (v. 5.0). RNA integrity was assessed by 1.5% agarose gel electrophoresis, and RNA concentration was determined using a microvolume spectrophotometer (Merinton Instrument, Beijing, China). cDNA was synthesized using the Maxima Reverse Transcriptase kit (Thermo Fisher Scientific). Quantitative PCR was performed using Fast qPCR Master Mix (High Rox, China) on a 7500 Fast Real-Time PCR System (Applied Biosystems, USA). Amplification and melting curves were analyzed for both the reference gene (beta-actin) and the target gene (*rnf212b*). Mean C_t values were calculated from three technical replicates. Relative gene expression was determined using the $2^{-\Delta\Delta Ct}$ method.

Data availability

The raw genomic sequencing data have been deposited in the National Center for Biotechnology Information (NCBI) under BioProject PRJNA1134618 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1134618), with accession numbers

https://doi.org/10.1007/s11427-024-2868-y

SRR30291412–SRR30291417 and SRR30280421–SRR30280424, and in the National Genomics Data Center (NGDC) under BioProject PRJCA027450 (https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA027450), with accession numbers CRA017553 and CRA018195. Additionally, mRNA-seq data (NCBI: SRR29824779–SRR29824869; NGDC: CRA017468, CRA017690), Hi-C data (NCBI: SRR30280467–SRR30280474; NGDC: CRA017535), and ATAC-seq data (NCBI: SRR30179897–SRR30179904; NGDC: CRA017508) associated with this project are available under the same BioProjects (NCBI: PRJNA1134618; NGDC: PRJCA027450). The assembled genome and annotation files of the two hybrid fish have been deposited on figshare (DOI: 10.6084/m9.figshare.26161939).

Compliance and ethics

The authors declare that they have no conflict of interest. All procedures performed on animals were approved by the academic committee at Hunan Normal University, Hunan, China.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (32341057, 32293252, U19A2040, 32002372), Hunan Provincial Natural Science Foundation (2022JJ10035), National Key Research and Development Program of China (2023YFD2401602), Special Funds for Construction of Innovative Provinces in Hunan Province (2021NK1010), Earmarked Fund for China Agriculture Research System (CARS-45), and 111 Project (D20007).

Supporting information

The supporting information is available online at https://doi.org/10.1007/s11427-024-2868-y. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

Bell, G.D.M., Kane, N.C., Rieseberg, L.H., and Adams, K.L. (2013). RNA-Seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. Genome Biol Evol 5, 1309–1323.

Bozdag, G.O., Ono, J., Denton, J.A., Karakoc, E., Hunter, N., Leu, J.Y., and Greig, D. (2021). Breaking a species barrier by enabling hybrid recombination. Curr Biol 31, R180–R181.

Buerkle, C.A., Morris, R.J., Asmussen, M.A., and Rieseberg, L.H. (2000). The likelihood of homoploid hybrid speciation. Heredity 84, 441–451.

Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O., and Shendure, J. (2013). Chromosome-scale scaffolding of *de novo* genome assemblies based on chromatin interactions. Nat Biotechnol 31, 1119–1125.

Capel, B. (2017). Vertebrate sex determination: evolutionary plasticity of a fundamental switch. Nat Rev Genet 18, 675–689.

Chen, J., Liu, H., Gooneratne, R., Wang, Y., and Wang, W. (2022). Population genomics of megalobrama provides insights into evolutionary history and dietary adaptation. Biology 11, 186.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890.

Cheng, H., Jarvis, E.D., Fedrigo, O., Koepfli, K.P., Urban, L., Gemmell, N.J., and Li, H. (2022). Haplotype-resolved assembly of diploid genomes without parental data. Nat Biotechnol 40, 1332–1335.

Cohen, M.J., Creed, I.F., Alexander, L., Basu, N.B., Calhoun, A.J.K., Craft, C., D'Amico, E., DeKeyser, E., Fowler, L., Golden, H.E., et al. (2016). Do geographically isolated wetlands influence landscape functions? Proc Natl Acad Sci USA 113, 1978–1986.

Condezo, Y.B., Sainz-Urruela, R., Gomez-H, L., Salas-Lloret, D., Felipe-Medina, N., Bradley, R., Wolff, I.D., Tanis, S., Barbero, J.L., Sánchez-Martín, M., et al. (2024). RNF212B E3 ligase is essential for crossover designation and maturation during male and female meiosis in the mouse. Proc Natl Acad Sci USA 121, e2320995121.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B.J., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14, 959–962.

Cox, M.P., Dong, T., Shen, G.G., Dalvi, Y., Scott, D.B., Ganley, A.R.D., and Stajich, J.E. (2014). An interspecific fungal hybrid reveals cross-kingdom rules for allopolyploid gene expression patterns. PLoS Genet 10, e1004180.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Gao, X., Zhang, H., Cui, J., Yan, X., Zhang, X., Luo, M., Tang, C., Ren, L., and Liu, S. (2021). Interactions between mitochondrial and nuclear genomes and co-regulation of mitochondrial and nuclear gene expression in reciprocal intergeneric hybrids between *Carassius auratus* red var. × *Cyprinus carpio* L. Reprod Breed 1, 213–220.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589.

Hou, M., Wang, Q., Zhao, R., Cao, Y., Zhang, J., Sun, X., Yu, S., Wang, K., Chen, Y., Zhang, Y., et al. (2024). Analysis of chromatin accessibility and DNA methylation to reveal the functions of epigenetic modifications in *Cyprinus carpio* gonads. Int J Mol Sci 25, 321.

Hu, F., Zhong, H., Wu, C., Wang, S., Guo, Z., Tao, M., Zhang, C., Gong, D., Gao, X., Tang, C., et al. (2021). Development of fisheries in China. Reprod Breed 1, 64–79.

Hurst, L.D., and Pomiankowski, A. (1991). Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. Genetics 128, 841–858.

Janko, K., Pačes, J., Wilkinson-Herbots, H., Costa, R.J., Roslein, J., Drozd, P., Iakovenko, N., Rídl, J., Hroudová, M., Kočí, J., et al. (2018). Hybrid asexuality as a primary postzygotic barrier between nascent species: on the interconnection between asexuality, hybridization and speciation. Mol Ecol 27, 248–263.

Jerković, I., Szabo, Q., Bantignies, F., and Cavalli, G. (2020). Higher-order chromosomal structures mediate genome function. J Mol Biol 432, 676–681.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Lin, Y., Li, J., Gu, Y., Jin, L., Bai, J., Zhang, J., Wang, Y., Liu, P., Long, K., He, M., et al. (2024). Haplotype-resolved 3D chromatin architecture of the hybrid pig. Genome Res 34, 310–325.

Liu, F.J., Jin, S.H., Li, N., Liu, X., Wang, H.Y., and Li, J.Y. (2011). Comparative and functional analysis of testis-specific genes. Biol Pharm Bull 34, 28–35.

Mallet, J. (2007). Hybrid speciation. Nature 446, 279–283.

McDermott, S.R., and Noor, M.A.F. (2010). The role of meiotic drive in hybrid male sterility. Phil Trans R Soc B 365, 1265–1272.

Meier, J.I., Stelkens, R.B., Joyce, D.A., Mwaiko, S., Phiri, N., Schliewen, U.K., Selz, O. M., Wagner, C.E., Katongo, C., and Seehausen, O. (2019). The coincidence of ecological opportunity with hybridization explains rapid adaptive radiation in Lake Mweru cichlid fishes. Nat Commun 10, 5391.

Murat, F., Mbengue, N., Winge, S.B., Trefzer, T., Leushkin, E., Sepp, M., Cardoso-Moreira, M., Schmidt, J., Schneider, C., Mößinger, K., et al. (2023). The molecular evolution of spermatogenesis across mammals. Nature 613, 308–316.

Nicol, E., Stevens, J.R., and Jobling, S. (2017). Riverine fish diversity varies according to geographical isolation and land use modification. Ecol Evol 7, 7872–7883.

Ong, C.T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet 12, 283–293.

Patterson, J., Carpenter, E.J., Zhu, Z., An, D., Liang, X., Geng, C., Drmanac, R., and Wong, G.K.S. (2019). Impact of sequencing depth and technology on *de novo* RNA-Seq assembly. BMC Genomics 20, 604.

Polovina, E.S., Parakatselaki, M.E., and Ladoukakis, E.D. (2020). Paternal leakage of mitochondrial DNA and maternal inheritance of heteroplasmy in *Drosophila* hybrids. Sci Rep 10, 2599.

Powell, D.L., García-Olazábal, M., Keegan, M., Reilly, P., Du, K., Díaz-Loyo, A.P., Banerjee, S., Blakkan, D., Reich, D., Andolfatto, P., et al. (2020). Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. Science 368, 731–736.

Ren, L., Li, W., Qin, Q., Dai, H., Han, F., Xiao, J., Gao, X., Cui, J., Wu, C., Yan, X., et al. (2019). The subgenomes show asymmetric expression of alleles in hybrid lineages of *Megalobrama amblycephala* × *Culter alburnus*. Genome Res 29, 1805–1815.

Reynolds, A., Qiao, H., Yang, Y., Chen, J.K., Jackson, N., Biswas, K., Holloway, J.K., Baudat, F., de Massy, B., Wang, J., et al. (2013). RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis. Nat Genet 45, 269–278.

Rieseberg, L.H., Raymond, O., Rosenthal, D.M., Lai, Z., Livingstone, K., Nakazato, T., Durphy, J.L., Schwarzbach, A.E., Donovan, L.A., and Lexer, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301, 1211–1216.

Rio, D.C., Ares Jr, M., Hannon, G.J., and Nilsen, T.W. (2010). Purification of RNA Using TRIzol (TRI Reagent). Cold Spring Harb Protoc 2010(6), pdb.prot5439.
Romero, I.G., Ruvinsky, I., and Gilad, Y. (2012). Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet 13, 505–516.
Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. Genome Biol 16, 259.
Sokpor, G., Xie, Y., Rosenbusch, J., and Tuoc, T. (2017). Chromatin remodeling BAF (SWI/SNF) complexes in neural development and disorders. Front Mol Neurosci 10, 243.
Spielmann, M., Lupiáñez, D.G., and Mundlos, S. (2018). Structural variation in the 3D genome. Nat Rev Genet 19, 453–467.
Srinivasan, K.A., Virdee, S.K., and McArthur, A.G. (2020). Strandedness during cDNA synthesis, the stranded parameter in htseq-count and analysis of RNA-Seq data. Brief Funct Genomics 19, 339–342.
Sun, N., Zhu, D.M., Li, Q., Wang, G.Y., Chen, J., Zheng, F., Li, P., and Sun, Y.H. (2021). Genetic diversity analysis of Topmouth Culter (*Culter alburnus*) based on microsatellites and D-loop sequences. Environ Biol Fish 104, 213–228.
Takamura, K., and Nakahara, M. (2015). Intraspecific invasion occurring in geographically isolated populations of the Japanese cyprinid fish *Zacco platypus*. Limnology 16, 161–170.
Tautz, D., and Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. Nat Rev Genet 12, 692–702.
Wang, S., Tang, C., Tao, M., Qin, Q., Zhang, C., Luo, K., Zhao, R., Wang, J., Ren, L., Xiao, J., et al. (2019). Establishment and application of distant hybridization technology in fish. Sci China Life Sci 62, 22–45.
Weaver, I.C.G., Korgan, A.C., Lee, K., Wheeler, R.V., Hundert, A.S., and Goguen, D. (2017). Stress and the emerging roles of chromatin remodeling in signal integration and stable transmission of reversible phenotypes. Front Behav Neurosci 11, 41.
Wu, C., Huang, X., Chen, Q., Hu, F., Zhou, L., Gong, K., Fu, W., Gong, D., Zhao, R., Zhang, C., et al. (2020). The formation of a new type of hybrid culter derived from a hybrid lineage of *Megalobrama amblycephala* (♀) × *Culter alburnus* (♂). Aquaculture 525, 735328.
Xiao, J., Hu, F., Luo, K., Li, W., and Liu, S. (2016). Unique nucleolar dominance patterns in distant hybrid lineage derived from *Megalobrama amblycephala* × *Culter alburnus*. BMC Genet 17, 1.
Xiao, J., Kang, X., Xie, L., Qin, Q., He, Z., Hu, F., Zhang, C., Zhao, R., Wang, J., Luo, K., et al. (2014). The fertility of the hybrid lineage derived from female *Megalobrama amblycephala* × male *Culter alburnus*. Anim Reprod Sci 151, 61–70.
Yang, H., and Wang, K. (2015). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10, 1556–1566.
Yang, T., Zhang, F., Yardımcı, G.G., Song, F., Hardison, R.C., Noble, W.S., Yue, F., and Li, Q. (2017). HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. Genome Res 27, 1939–1949.
Yoo, M.J., Liu, X., Pires, J.C., Soltis, P.S., and Soltis, D.E. (2014). Nonadditive gene expression in polyploids. Annu Rev Genet 48, 485–517.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137.