

Genome-wide DNA Methylation and Gene Expression Patterns of Androgenetic Haploid Tiger Pufferfish (*Takifugu Rubripes*) Provide Insights into Haploid Syndrome

He Zhou

Dalian Ocean University

Qian Wang

Chinese Academy of Fishery Sciences

Zi-Yu Zhou

Dalian Ocean University

Xin Li

Dalian Ocean University

Yu-Qing Sun

Dalian Ocean University

Gu Shan

Dalian Ocean University

Xin-Yi Zheng

Dalian Ocean University

Qi Chen

Dalian Ocean University

Hai-Jin Liu

Dalian Days Industrial Co., LTD

Wei Wang

Dalian Ocean University

Chang-Wei Shao (✉ shaochangwei303@163.com)

Chinese Academy of Fishery Sciences

Research Article

Keywords: Takifugu rubripes, androgenesis, haploid, whole genome bisulfite sequencing, RNA-seq

Posted Date: June 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-601807/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Genome-wide DNA methylation and gene expression patterns of
androgenetic haploid tiger pufferfish (*Takifugu rubripes*) provide insights
into haploid syndrome**

He Zhou^{1,†}, Qian Wang^{2,3,†}, Zi-Yu Zhou¹, Xin Li¹, Yu-Qing Sun¹, Gu Shan¹, Xin-Yi Zheng¹, Qi Chen¹, Hai-Jin Liu⁴, Wei Wang¹, Chang-Wei Shao^{2,3,*}

¹ Key Laboratory of Mariculture, Agriculture Ministry, PRC, Dalian Ocean University, Dalian 116023, China.

² Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China.

³ Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture, Qingdao 266071, China.

⁴ Dalian Days Industrial Co., LTD, Dalian 116011, China.

[†] These authors contributed equally to this work.

* Correspondence:

Changwei Shao

Nanjing Road 106, Qingdao, 266071, China

Email: shaochangwei303@163.com

Abstract

Androgenesis is an important chromosome set manipulation technique used in sex control in aquaculture. Haploid embryos exhibit haploid syndrome and show abnormalities and even die during early embryonic development. In this study, we used whole genome bisulfite sequencing (WGBS) to investigate the genome-wide DNA methylation profiles in haploid females (1n-X) and males (1n-Y), and diploid females (2n-XX) and males (2n-XY) of tiger pufferfish (*Takifugu rubripes*), an economically important fish in China. A total of 96.32 Gb clean data was produced. Analysis of differentially methylated regions (DMRs) showed that haploids had more hyper-methylated regions than diploids, which may be related to abnormal development and early embryonic death in haploids. There were 7,838 hyper-methylated differentially methylated genes (DMGs) and 4,755 hypo-methylated DMGs in haploid vs. diploid comparisons in both females and males. These DMGs were mainly related to genomic stability maintenance and cell cycle regulation. *slf1*, *actr8*, *gas2*, and *pbrm1* genes were detected to validate the methylation sequencing. After combining the methylation data with the corresponding transcriptome data, we identified several genes, including *guca2a*, *myoc*, *fezf2*, *rprml*, *telo2*, *s100a1*, and *marveld1*, which exhibited differential expression levels modulated by DNA methylation. In conclusion, our study revealed different methylation and expression profiles between haploid and diploid *T. rubripes*

for the first time. Several DMGs were identified between different ploidy levels, which may be related to haploid syndrome formation. The results expand the understanding of the effects of ploidy on the early development of teleosts and provide knowledge about target genes and networks to improve the survival rate of haploids.

Key words *Takifugu rubripes*; androgenesis; haploid; whole genome bisulfite sequencing; RNA-seq

Introduction

Sex control is important because of the different features that discriminate between male and female fish. Different features, such as external body color, shape and size are associated with reproduction. A system that includes chromosome set and genome manipulation has been developed to achieve mono-sex biased populations by altering the number or combination of the chromosome set to maximize productivity. Like most vertebrates, male teleosts undergo complete meiosis and spermatogenesis to produce haploid sperms. However, females ovulate physiologically fertile mature eggs arrested at the metaphase of the second meiotic division (MII arrest). Eggs resume the MII arrest and release the second polar body with a haploid set of chromosomes to complete meiosis while activated by sperms. Subsequently, female and male pronuclei are fused to form zygotic nuclei, and mitosis is initiated to continue cell reproduction and differentiation. These

features of teleosts facilitate the chromosome set manipulation to produce artificial polyploid and unisexual progeny.

Androgenesis is defined as a chromosome set manipulation technique leading to a uniparental reproduction without any genetic contribution from the maternally derived nucleus [1]. Spontaneous androgenetic teleosts have never been reported in nature [2]. Androgenesis has significant applications in aquaculture, such as rapid generation of inbred lines, all-paternal inheritance, recovering genotypes from cryopreserved sperm of unique or endangered species and studying the effects of mitochondrial genotype on performance. The preparation for artificial androgenesis has two major steps, such as genetic inactivation of egg nucleus and duplication of sperm-derived paternal chromosomes using temperature or hydrostatic pressure shock at an optimal time for inhibiting the cleavage division. Egg nucleus has been genetically inactivated by irradiation with X-ray, γ -ray or ultraviolet (UV) ray [2-4]. These techniques are simple, but egg nuclei, cytoplasm, and various cell organelles can be damaged to a certain extent, thereby causing abnormal development. Recently, Morishima *et al.* [5] reported that the cold-shock treatment of fertilized eggs of pond loach (*Misgurnus anguillicaudatus*) could successfully induce haploid androgenesis, and then doubled haploid loach were successfully produced based on this technique [6, 7]. The feasibility of cold-shock-induced androgenesis has also been reported in large-scale loach (*Paramisgurnus dabryanus*) [8], zebrafish (*Danio rerio*) [9], and olive flounder

(*Paraichthys olivaceus*) [10].

Tiger pufferfish (*Takifugu rubripes*) is an economically important fish in China. The mature male has better consumer appeal and economic benefits due to the meat quality and taste. Therefore, it is necessary to develop a safe and efficient method without any irradiation or hormone for all-male production. In our previous study, we used a cold-shock technique to produce androgenetic haploid tiger pufferfish with a proportion of 86.7%. However, the haploid embryos exhibited haploid syndrome, which displayed abnormalities and even died during early embryonic development. As previously shown in pond loach and zebrafish, the egg nucleus could be released along with the second polar body under cold treatment. The sperm nucleus remained in the egg to initiate androgenetic development [5, 9]. However, no study reported the mechanism of the haploid syndrome in tiger pufferfish. In the present study, we cold-shocked the fertilized eggs of tiger pufferfish to induce haploid androgenesis and conducted cytological observation to get an insight into this biological process. Furthermore, we applied whole genome bisulfite sequencing (WGBS) and RNA-seq on androgenesis-derived haploid and zygote tiger pufferfish to investigate the gene expression and methylation changes.

Materials and Methods

Ethics statement

All experiments in this study were approved by the Animal Study Ethical Committee of Dalian Ocean University, and the experiments were performed according to the Guide for the Care and Use of Laboratory Animals in Dalian Ocean University, Dalian, China. According to the editorial policy for scientific reports, the study was carried out in compliance with the ARRIVE guidelines.

Haploid androgenic tiger pufferfish production

Parental tiger pufferfish were injected with human chorionic gonadotropin (HCG) 24 h prior to artificial insemination (females: 2,500-3,000 IU/kg; males: 1,250-1,500 IU/kg). For haploid androgenic tiger pufferfish production, 8-minute-post-fertilization embryos were cold-treated in 4°C seawater for 60 min and transferred to 20°C aerated seawater for hatching [11]. The diploid group maintained in 20°C aerated seawater was prepared as control group.

DNA relative content determination

Embryos were developed until eyespots appeared, and thirty embryos were randomly sampled from the cold-treated and control groups, respectively. Each embryo was dissected into three parts, and one-third of each embryo was used for DNA relative content analysis. DNA relative content of each embryo was measured using Partec PA Flow Cytometry (Munich, Germany) as previously reported [12]. After removing chorion and yolk, the embryos were placed into Cystain DNA 1 Step Staining Solution (Partec, Germany) at 4°C without exposure to light. The samples were ground, filtered, and mixed with 500 µL distilled water, and the DNA contents were

then measured and recorded.

Genetic sex identification

For genetic sex identification, one-third of each embryo was dissected for genomic DNA (gDNA) extraction using the traditional phenol extraction method. Single nucleotide polymorphism (SNP) in the sex-determining gene, anti-Müllerian hormone receptor (*amhr2*), was used to identify the genetic sex. The locus on the X chromosome was C, and the locus on the Y chromosome was G [13]. Primers, including *amhr2*-F and *amhr2*-R, are listed in Table S1. The PCR conditions were as follows: 95°C for 3 min, 35 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 40 s, and 72°C for 5 min. PCR products were gel extracted and ligated to pMD18-T vector for sequencing. The remaining one-third of each fish was used for WGBS and RNA-seq.

WGBS

Two biological replicates were used for each of the four samples, and each replicate was pooled by three fish. The samples, including haploid females (1n-X) and males (1n-Y), and diploid females (2n-XX) and males (2n-XY) were detected. Total DNA was extracted using QIAamp Fast DNA Tissue Kit (Qiagen, Dusseldorf, Germany) according to the manufacturer's instructions. The quantity of DNA was measured by reading A260/280 ratios by spectrophotometer Agilent 2100 bioanalyzer (Thermo, SC, USA). The fragmented DNA samples were subjected to bisulfite conversion. The Accel-NGS Methyl-Seq DNA Library Kit (Swift, MI, USA) was used for attaching

adapters to single-stranded DNA fragments. Briefly, the Adaptase step is a highly efficient, proprietary reaction that was performed for end-repair, tailing of 3' ends and ligation of the first truncated adapter complement to 3' ends. The extension step was used to incorporate truncated adapter 1 by a primer extension reaction. The ligation step was used to add the second truncated adapter to the bottom strand. The indexing PCR step increased yield and incorporated full-length adapters. Bead-based SPRI clean-ups (Beckman, IN, USA) were used to remove both oligonucleotides and small fragments and change enzymatic buffer composition. Finally, pair-end 2×150 bp sequencing was performed on an Illumina HiSeq 4000 platform (LC Sciences, Houston, TX, USA).

RNA-seq

Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The total RNA quantity and purity were analyzed using Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent, CA, USA) with RNA integrity number (RIN) >7.0. Ten µg of total RNA representing a specific adipose type was subjected to isolate Poly (A) mRNA with poly-T oligo-attached magnetic beads (Invitrogen, Carlsbad, CA, USA). Following purification, the poly(A)- or poly(A)+ RNA fraction was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were reverse-transcribed to create the final cDNA library using mRNA-Seq Sample Preparation Kit (Illumina, San

Diego, CA, USA) according to the manufacturer's instructions, and the average insert size for the paired-end libraries was 300 bp (± 50 bp). The paired-end sequencing was performed on an Illumina HiSeq 4000 (LC-Bio, Hangzhou, China) following the vendor's recommended protocol.

Bioinformatics analysis

Cutadapt [14] and Perl scripts in house were used to remove the reads that contained adapter contamination and low-quality and undetermined bases. Then, sequence quality was verified using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), including the Q20, Q30 and GC-content of the clean data. All downstream analyses were based on clean data of high quality. For WGBS, the reads were mapped to the reference genome using WALT [15]. After alignment, the reads were deduplicated using SAMTool [16]. For each cytosine site (or guanine corresponding to cytosine on the opposite strand) in the reference genome sequence, the DNA methylation level was determined by the ratio of the number of reads supporting C (methylated) to that of total reads (methylated and unmethylated) using per script in house and MethPipe [17]. Differentially methylated regions (DMRs) were calculated by R package-MethylKit [18] with default parameters (1000 bp slide windows, 500 bp overlap, p value <0.05). For RNA-seq, *de novo* assembly of the transcriptome was performed with Trinity 2.4.0 [19]. Trinity grouped transcripts into clusters based on shared sequence content. A transcript cluster was referred to gene. The longest transcript in

the cluster was chosen as the gene sequence (aka unigene). All assembled unigenes were aligned against the non-redundant (NR) protein database (<http://www.ncbi.nlm.nih.gov/>), and Gene Ontology (GO) (<http://www.geneontology.org>), SwissProt (<http://www.expasy.ch/sprot/>), Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>), and EggNOG (<http://eggnogdb.embl.de/>) databases using DIAMOND [20] with a threshold of Evalue < 0.00001. Salmon 1.4.0 was used to determine the expression levels for unigenes by calculating transcripts per million mapped reads (TPM) [21, 22]. The differentially expressed unigenes were selected with log₂ (fold change) >1 or log₂ (fold change) <-1 and statistical significance (*p* value <0.05) by R package edgeR [23]. Then, GO and KEGG enrichment analyses were performed on the differentially expressed unigenes by Perl scripts in house.

Bisulfite sequencing-PCR (BS-PCR)

To check the reliability of WGBS, parts of DMRs were selected to perform BS-PCR verification, such as SMC5-SMC6 complex localization factor 1 (*slf1*), actin-related protein 8 (*actr8*), growth arrest-specific protein 2 (*gas2*) and polybromo 1 (*pbrm1*). Pooled DNA samples from the same family were sodium bisulfite-modified following the manufacturer's instructions. The primers for BS-PCR were identified by online MethPrimer design software (<http://www.urogene.org/methprimer/>) (Table S1). PCR was performed using TaKaRa EpiTaq HS (Takara, Shiga, Japan) in accordance with the

manufacturer's instructions. The amplified products were purified and cloned into pEASY-T1 vector, and at least 20 clones were randomly selected for sequencing.

Quantitative real time-PCR (qRT-PCR)

To check the reliability of RNA-seq, nine DEGs were randomly selected to perform qRT-PCR verification, including transcription factor AP-2 epsilon (*tfap2e*), MAF BZIP transcription factor K (*mafk*), proline-rich transmembrane protein 1 (*prrt1*), adenylate cyclase 7 (*adcy7*), Rho-associated protein kinase 2 (*rock2*), oligodendrocyte transcription factor 2 (*olig2*), WD repeat domain 5 (*wdr5*), cAMP-regulated phosphoprotein 19 (*arpp19*), and WD repeat domain 1 (*wdr1*). The primers for qRT-PCR are shown in Table S1. The qRT-PCR was performed using SYBR Premix Ex TaqTM II (Tli RNaseH Plus) (Takara, Shiga, Japan) according to the manufacturer's instructions. The PCR conditions were as follows: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s and 60°C for 30 s. All samples were run with triplicate. The relative expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method. Statistical significance was determined by one-way analysis of variance (ANOVA). Significance was set at $p < 0.05$.

Results

Determination of diploid and cold-treated haploid tiger pufferfish

The embryos in the control group showed normal appearance (Figure 1A).

Most embryos in the cold-shock group showed abnormalities, including spinal curvature, tail malformations, insufficiency of blood circulation, etc. (Figure 1B). The DNA content of cold-treated haploid tiger pufferfish was half of that of diploid tiger pufferfish (Figure 1C). The specific SNP locus in the *amhr2* gene was detected. The genotypes of female and male diploid tiger pufferfish were CC and CG, respectively. In haploid tiger pufferfish, the genotypes of female and male haploid were C and G, respectively (Figure 1D).

WGBS data

To evaluate the DNA methylation patterns across different chromosome ploidy and sex of tiger pufferfish, we performed WGBS on genomic DNA extracted from haploid females (1n-X) and males (1n-Y), and diploid females (2n-XX) and males (2n-XY). A total of 96.32 Gb clean data were produced from the eight libraries, which yielded an average depth of 33 × per strand for each sample (Table S2). The data were uploaded to National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA accession: PRJNA535391).

An average of 33.29% of genomic cytosines (Cs) was covered, and the average percent of mCs was 10.80%. Among the mCs, 90.95%, 2.58%, and 6.47% were mCG, mCHG, and mCHH types, respectively (Figure 2A, Table S3). The mCG sites were focused on the subsequent analyses. Then, the methylation status of CGs was analyzed in various genomic elements. The

hyper-methylated CGs (methylation level > 0.75) were enriched in exons and introns, whereas hypo-methylated CGs (methylation level < 0.25) were enriched in gene promoters and intergenic regions (Figure 2B). The chromosome methylation status of mCGs was analyzed. The highest mCG level was observed on chromosome 1 (74.32%, 72.15%, 72.72%, and 71.41% for 1n-X, 1n-Y, 2n-XX, and 2n-XY, respectively), and the lowest mCG level was observed on chromosome 18 (28.89%, 28.00%, 28.15%, and 27.55% for 1n-X, 1n-Y, 2n-XX, and 2n-XY, respectively) (Figure 2C).

Analysis of DMRs

Pair-wise comparisons were applied to detect DMRs between haploid and diploid tiger pufferfish (1n-X vs. 2n-XX and 1n-Y vs. 2n-XY). In 1n-X vs. 2n-XX comparison, there were 470,641 haploid hyper-methylated and 240,996 haploid hypo-methylated DMRs, corresponding to 11,513 and 9,238 DMGs, respectively. In 1n-Y vs. 2n-XY comparison, there were 560,100 haploid hyper-methylated and 194,484 haploid hypo-methylated DMRs corresponding to 10,000 and 7,384 DMGs, respectively (Figure 3A & 3B). The number of DMRs of each group showed that haploids had more hyper-methylated regions than diploids, which may be related to abnormal development and early embryonic death in haploids.

To evaluate the functions of these DMGs, GO and KEGG pathway enrichment analyses were used. DMGs between haploid and diploid females

were significantly enriched in GO terms of molecular function, protein kinase activity, cellular component and positive regulation of transcription from RNA polymerase II promoter, cilium and multicellular organism development ($p < 0.05$) (Figure S1A). In males, the significantly enriched GO terms were protein kinase activity, cilium, cellular component, multicellular organism development, positive regulation of transcription from RNA polymerase II promoter, transcription, DNA-templated and cellular response to interferon-gamma ($p < 0.05$) (Figure S1B). Also, DMGs between haploid and diploid were significantly enriched in adipocytokine signaling pathway in females and mitogen-activated protein kinase (MAPK) signaling pathway in males ($p < 0.05$) (Figure S1C, D).

When compared DMGs between 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY, 7,838 DMGs were hyper-methylated in haploid vs. diploid comparisons in females and males, and 4,755 DMGs were hypo-methylated in haploid males and females (Figure 3B). GO enrichment analysis revealed that the haploid hyper-methylated DMGs were significantly enriched in GO terms, such as G protein-coupled receptor signaling pathway, cilium assembly and anatomical structure formation involved in morphogenesis, translation and heart development. Haploid hypo-methylated DMGs were significantly enriched in GO terms, such as intracellular protein transport, positive regulation of transcription by RNA polymerase II, positive regulation of multicellular organismal process, tissue morphogenesis and anatomical structure

formation involved in morphogenesis (Figure 3C). The GO terms of regulation of mitotic sister chromatid segregation and regulation of mitotic cell cycle phase transition were also significantly enriched in haploid hyper-methylated DMGs ($p < 0.05$) (Table S4). Several DMGs related to genomic stability maintenance and cell cycle changed the methylation levels between haploids and diploids. For example, *slf1* and *actr8* were hyper-methylated in haploid females and males. *gas2* and *pbrm1* were hypo-methylated in haploid females and males (Figure S2).

Transcriptome sequencing and assembly

To identify the differentially expressed genes (DEGs) between haploids and diploids and assess the relationship between DNA methylation and gene expression, gene expression profiles were measured by transcriptome sequencing using the same samples in the WGBS study. The data were uploaded to NCBI SRA (SRA accession no.: PRJNA540016). A total of 60.97 G raw bases were obtained. After quality control, a total of 54.00 G clean reads (10.71 G in 1n-X, 13.52 G in 1n-Y, 15.27 G in 2n-XX and 14.50 G in 2n-XY, respectively) with an average Q20 percentage of 97.24% were generated and used for subsequent analysis (Table S5). The *de novo* assembled transcriptomes included 74,033 unigenes with an N50 length of 1,507 bp and an average length of 716 bp (Figure 4A, Table S6). The reference transcriptome of tiger pufferfish unigenes was annotated by NCBI, NR, GO,

KEGG, Pfam, Swissprot, and EggNOG databases (Table S7). DEG analysis showed that there were 411 DEGs significantly up-regulated and 1,077 were down-regulated in 1n-X compared with 2n-XX. Compared with 2n-XY, there were 156 significantly up-regulated and 182 down-regulated DEGs in 1n-Y (Figure 4B). Venn diagram identified that 87 DEGs were shared both in 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons (Figure 4C). To identify transcriptional events occurred during cold-treated haploid induction, DEGs shared in 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons were subjected to GO enrichment analysis. The GO annotation suggested that single-organism process, cellular process and metabolic process were the most abundant GO functions in the biological process. Cell, cell part and membrane were most abundant in the cellular component. In molecular function, GO terms were most enriched for binding, catalytic activity and antioxidant activity (Figure 4D). In KEGG enrichment analysis, 85 KEGG pathways were enriched, of which 6 pathways were significantly enriched ($p < 0.05$), including synthesis and degradation of ketone bodies, retinol metabolism, protein digestion and absorption, extracellular matrix (ECM)-receptor interaction, dorso-ventral axis formation and prion diseases (Figure 4E).

To validate the RNA-seq results, relative mRNA levels for nine DEGs (*tfap2e*, *mafk*, *prrt1*, *adcy7*, *rock2*, *olig2*, *wdr5*, *arpp19*, and *wdr1*) were measured by qRT-PCR (Figure S3). The RNA-seq results were in agreement with the transcriptome data.

Correlation between methylation and gene expression

In order to analyze the relationship between DNA methylation regions and DEGs, the whole genome DNA methylation and transcriptional group were combined to screen the genes having a positive relationship between DNA methylation and expression level. The thresholds were set as fold-change > 1.2 and p value < 0.05 for WGBS and \log_2 |fold-change| > 2 and p value < 0.05 for transcriptomes. In females, guanylate cyclase activator 2A (*guca2a*) had hyper-methylation level with down-regulated mRNA expression level and myocilin (*myoc*) had hypo-methylation level with up-regulated mRNA expression level in 1n_X vs. 2n_XX comparison. In males, *fez* family zinc finger protein 2 (*fezf2*), reprimin-like protein (*rprm1*), telomere length regulation protein TEL2 homolog (*telo2*) and S-100 protein alpha chain (*s100a1*) had hyper-methylation level with down-regulated mRNA expression level, and MARVEL domain-containing protein 1 (*marveld1*) had hypo-methylation level with up-regulated mRNA expression level in 1n_Y vs. 2n_XY comparison (Figure 5).

Discussion

Androgenesis is an important manipulation in fish aquaculture for the acquisition of mono-sex fish having sex-biased characteristics and the conservation of fish germplasm [24]. Androgenetic embryos may exhibit

haploid syndrome, which includes developmental defects and extensive lethality [25]. DNA methylation is an essential epigenetic mechanism during embryonic development. In order to evaluate methylation and gene expression changes after androgenesis, the whole genome methylation and transcription profiles in haploid and diploid tiger pufferfish were examined. To the best of our knowledge, it is the first WGBS on haploids in aquaculture fish.

Our analysis on the haploid and diploid methylome profiles of tiger pufferfish found that about 10.80% of C was methylated in both genetic female and male genomes, which is consistent with our previous reports in adult tiger pufferfish [26] and Chinese tongue sole (*Cynoglossus semilaevis*) [27]. The global DNA methylation level in haploid fish was higher than that in diploid females and males. There was limited information on the methylation state in haploid and diploid fish. A study on rice (*Oryza sativa* L.) found that the relative level of DNA methylation in diploid fish was higher than that in haploid fish. The methylation in CG sequence context did not differ between haploid and diploid fish [28]. In animals, DNA methylation occurs in the CG context, and it commonly occurs within all sequence contexts (CG, CHG and CHH) in plants [29]. More information on methylation changes among different ploidy levels in fish species is still needed.

Several DMRs were identified between haploid and diploid in both sexes, which indicated that DNA methylation might participate in abnormal development and early embryonic death in haploids. We found GO terms,

namely regulation of mitotic cell cycle phase transition and regulation of mitotic sister chromatid segregation, which were enriched. Cell cycle is important for the growth of an organism and plays important roles in cell proliferation, cell-fate decision and many other cell functions [30-32]. According to the previous findings, the duration of mitosis in haploid ESCs is significantly longer than that in diploid embryonic stem cells (ESCs), especially at the metaphase stage [33]. Besides, accelerate G2/M transition could stabilize haploid ESCs in mice [34]. These findings suggest the relationship between the cell cycle and ploidy in the organism. In the present study, *slf1* and *actr8* were hyper-methylated in haploid females and males. SLF1 acts with SLF2 as a bridge to link the SMC5/6 complex to RAD18, which is important for higher-order chromatin structure maintenance in mammals. SLF1 depletion can significantly increase abnormalities in anaphase cells [35]. ACTR8 is a key component of the INO80 complex, which has critical functions in DNA replication, repair and recombination, transcription and heterochromatin maintenance [36]. On the other hand, *gas2* and *pbrm1* were hypo-methylated in haploid females and males. GAS2 plays an important role in mitogenesis, cell cycle, microfilament alterations and cellular morphological variation during apoptosis [37-39]. *pbrm1* gene encodes BAF180 protein, which is a subunit of the ATP - dependent complex SWItch/Sucrose NonFermentable (SWI/SNF) regulating chromatin remodeling [40]. Different methylation levels in these genes between haploid and diploid

fish may reflect the effects of ploidy on methylation regulators related to cell cycle and mitosis, which may explain the different development situations between ploidy levels.

Transcriptome analysis revealed that KEGG pathways of synthesis and degradation of ketone bodies, retinol metabolism, protein digestion and absorption, ECM-receptor interaction and dorso-ventral axis formation were significantly enriched ($p < 0.05$) in haploid vs. diploid comparison, indicating that the difference between haploid and diploid is mainly related to energy metabolism and body axis formation. When combined the DNA methylation and transcription data, limited genes showed a negative correlation. *guca2a* gene was hyper-methylated with down-regulated mRNA expression in 1n_X vs. 2n_XX comparison, while *myoc* showed the opposite pattern in this study. It was reported that the down-regulation of GUCA2A was correlated with inflammatory bowel disease and disrupted intestinal homeostasis in mammals [41, 42]. Besides, the abnormal intracellular accumulation of myocilin could deteriorate the function of the trabecular meshwork cells and elevate intraocular pressure, which can lead to retinal ganglion cell death [43, 44]. In males, *fezf2* and *telo2* were hyper-methylated with down-regulated messenger RNA (mRNA) expression levels in 1n_Y vs. 2n_XY comparison. *fezf2* gene is a highly conserved gene that encodes a zinc finger transcriptional repressor. It is an important regulator to control the development of forebrain neuronal differentiation by promoting local

Wnt/ β -catenin signaling [45]. *TELO2* plays a vital role in the cell cycle by regulating S-phase checkpoint protein. Knockout of *telo2* causes embryonic lethality and S phase cell-cycle arrest in mice [46, 47]. *marveld1* gene was hypo-methylated with up-regulated mRNA expression level in 1n_Y vs. 2n_{XY} comparison. MARVELD1 is a microtubule-associated protein, which plays an important role in cell cycle progression and migration. In mice, overexpression of MARVELD1 can lead to remarkable inhibition of cell proliferation, G1-phase arrest and reduced cell migration [48]. The significant change of these genes indicated abnormal cell cycle, deterioration of visual function and disordered metabolism and development in haploid, which may cause a high fatality rate during early embryonic development.

Conclusion

Our study revealed different methylation and expression profiles between haploid and diploid teleosts for the first time. Several genes were identified showing different performances between different ploidy levels that may participate in haploid syndrome formation. These results may expand the understanding of the effects of ploidy on early development in teleosts and provide knowledge about target genes and networks to improve the survival rate of haploids. However, future studies are needed to determine the roles of these target genes during early embryonic development.

Acknowledgments

This work was supported by the Dalian High-Level Talent Innovation Support Program (Youth Star of Science and Technology, NO.: 2019RQ110), the Key Technology R&D Program of Dalian (NO.: 2019YF23SN050), the National Natural Science Foundation of China (NO.: 41606178), the Central Public-interest Scientific Institution Basal Research Fund, CAFS (NO.: 2020TD19), and the National Marine Genetic Resource Center.

Author Contributions

Experimental design and implementation: He Zhou. Low-temperature treatment: He Zhou, Qi Chen and Zi-Yu Zhou. Parentage identification and DNA and RNA extraction: Xin Li. BS-PCR verification: Zi-Yu Zhou. Quantitative real time-PCR (qRT-PCR): Xin-Yi Zheng. Ploidy identification: Yu-Qing Sun. Genotypic sex identification: Gu Shan. Cold-shock androgenesis and offspring breeding: Hai-Jin Liu and Wei Wang. Thesis writing: He Zhou and Qian Wang. Guidance and revision of the manuscript: Chang-Wei Shao.

Reference

1. Thorgaard, G.H., *Ploidy manipulation and performance*. Aquaculture, 1986. 57(1): p. 57-64.
2. Komen, H. and G.H. Thorgaard, *Androgenesis, gynogenesis and the production of clones in fishes: A review*. Aquaculture, 2007. 269(1-4): p. 150-173.

3. Pandian, T.J. and R. Koteeswaran, *Ploidy induction and sex control in fish*. *Hydrobiologia*, 1998. 384(1): p. 167-243.
4. Fujimoto, T., et al., *Evaluation of different doses of UV irradiation to loach eggs for Genetic inactivation of the maternal genome*. 2007. 449–462.
5. Morishima, K., et al., *Cold-shock eliminates female nucleus in fertilized eggs to induce androgenesis in the loach (*Misgurnus anguillicaudatus*), a teleost fish*. *BMC Biotechnol*, 2011. 11: p. 116.
6. Hou, J., et al., *Production of androgenetic diploid loach by cold-shock of eggs fertilized with diploid sperm*. *Theriogenology*, 2013. 80(2): p. 125-30.
7. Hou, J., et al., *Androgenetic doubled haploids induced without irradiation of eggs in loach (*Misgurnus anguillicaudatus*)*. *Aquaculture*, 2014. 420-421(1): p. S57-S63.
8. Wang, Y., et al., *Research on low-temperature induced androgenesis haploid of *Paramisgurnus dabryanus**. *Journal of Fisheries of China*, 2014. 38(2): p. 161-169.
9. Hou, J., et al., *Generation of clonal zebrafish line by androgenesis without egg irradiation*. *Scientific Reports*, 2015. 5.
10. Hou, J., et al., *Cold-shock induced androgenesis without egg irradiation and subsequent production of doubled haploids and a clonal line in Japanese flounder, *Paralichthys olivaceus**. *Aquaculture*, 2016. 464: p. 642-646.
11. Zhou, H., et al., *Optimization of inducing conditions for cold shocked androgenetic haploid in redfin puffer *Takifugu rubripes**. *Journal of Dalian Ocean University*, 2017. 32(3): p. 316-322.
12. Zhou, H., et al., *Androgenetic haploid Japanese pufferfish (*Takifugu rubripes*) induced by cold shock*. *Aquaculture Research*, 2019. 50(12): p. 3802-3811.
13. Kamiya, T., et al., *A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (*fugu*)*. *PLoS Genet*, 2012. 8(7): p. e1002798.
14. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. *EMBnet.journal*; Vol 17, No 1: Next Generation Sequencing Data Analysis DO - 10.14806/ej.17.1.200, 2011.

15. Chen, H., A.D. Smith, and T. Chen, *WALT: fast and accurate read mapping for bisulfite sequencing*. *Bioinformatics*, 2016. 32(22): p. 3507-3509.
16. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. *Bioinformatics*, 2009. 25(16): p. 2078-9.
17. Song, Q., et al., *A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics*. *PLoS One*, 2013. 8(12): p. e81148.
18. Akalin, A., et al., *methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles*. *Genome Biol*, 2012. 13(10): p. R87.
19. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome*. *Nature Biotechnology*, 2011. 29(7): p. 644-652.
20. Buchfink, B., C. Xie, and D.H. Huson, *Fast and sensitive protein alignment using DIAMOND*. *Nature Methods*, 2015. 12(1): p. 59-60.
21. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression*. *Nat Methods*, 2017. 14(4): p. 417-419.
22. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq*. *Nat Methods*, 2008. 5(7): p. 621-8.
23. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. *Bioinformatics*, 2010. 26(1): p. 139-40.
24. Pandian, T.J. and S. Kirankumar, *Androgenesis and conservation of fishes*. *Current Science*, 2003. 85(7): p. 917-931.
25. Luo, C. and B. Li, *Diploid-dependent regulation of gene expression: a genetic cause of abnormal development in fish haploid embryos*. *Heredity*, 2003. 90(5): p. 405-409.
26. Zhou, H., et al., *Changes in DNA methylation during epigenetic-associated sex reversal under low temperature in *Takifugu rubripes**. *PLoS One*, 2019. 14(8): p. e0221641.
27. Shao, C., et al., *Epigenetic modification and inheritance in sexual reversal of fish*. *Genome Res*, 2014. 24(4): p. 604-15.
28. Zhang, H., et al., *Effects of ploidy variation on promoter DNA methylation and gene expression in rice (*Oryza sativa* L.)*. *BMC plant biology*, 2018. 18(1):

p. 314-314.

29. Law, J.A. and S.E. Jacobsen, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals*. Nature reviews. Genetics, 2010. 11(3): p. 204-220.

30. Boward, B., T. Wu, and S. Dalton, *Concise Review: Control of Cell Fate Through Cell Cycle and Pluripotency Networks*. Stem Cells, 2016. 34(6): p. 1427-1436.

31. Dalton, S., *Linking the Cell Cycle to Cell Fate Decisions*. Trends Cell Biol, 2015. 25(10): p. 592-600.

32. Pauklin, S. and L. Vallier, *The cell-cycle state of stem cells determines cell fate propensity*. Cell, 2013. 155(1): p. 135-47.

33. Guo, A., et al., *Single-Cell Dynamic Analysis of Mitosis in Haploid Embryonic Stem Cells Shows the Prolonged Metaphase and Its Association with Self-diploidization*. Stem cell reports, 2017. 8(5): p. 1124-1134.

34. Takahashi, S., et al., *Induction of the G2/M transition stabilizes haploid embryonic stem cells*. Development, 2014. 141(20): p. 3842-7.

35. Raschle, M., et al., *DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links*. Science, 2015. 348(6234): p. 1253671.

36. Choe, S.-H., et al., *A single mutation in the ACTR8 gene associated with lineage-specific expression in primates*. BMC Evolutionary Biology, 2020. 20(1): p. 66.

37. Brancolini, C., S. Bottega, and C. Schneider, *Gas2, a growth arrest-specific protein, is a component of the microfilament network system*. J Cell Biol, 1992. 117(6): p. 1251-61.

38. Huang, W., et al., *Decreased calpain activity in chronic myeloid leukemia impairs apoptosis by increasing survivin in myeloid progenitors and xiap1 in differentiating granulocytes*. Oncotarget, 2017. 8(31): p. 50629-50641.

39. Zhu, R.X., et al., *Growth arrest-specific gene 2 suppresses hepatocarcinogenesis by intervention of cell cycle and p53-dependent apoptosis*. World J Gastroenterol, 2019. 25(32): p. 4715-4726.

40. da Costa, W.H., et al., *Polybromo-1 (PBRM1), a SWI/SNF complex subunit is a prognostic marker in clear cell renal cell carcinoma*. (1464-410X

(Electronic)).

41. Brenna, Ø., et al., *The guanylate cyclase-C signaling pathway is down-regulated in inflammatory bowel disease*. Scandinavian journal of gastroenterology, 2015. 50(10): p. 1241-1252.

42. Zhang, H., et al., *Integrated Analysis of Oncogenic Networks in Colorectal Cancer Identifies GUCA2A as a Molecular Marker*. Biochemistry Research International, 2019. 2019: p. 6469420.

43. Zhou, Y., O. Grinchuk, and S.I. Tomarev, *Transgenic Mice Expressing the Tyr437His Mutant of Human Myocilin Protein Develop Glaucoma*. Investigative Ophthalmology & Visual Science, 2008. 49(5): p. 1932-1939.

44. Liu, Y. and D. Vollrath, *Reversal of mutant myocilin non-secretion and cell killing: implications for glaucoma*. Human Molecular Genetics, 2004. 13(11): p. 1193-1204.

45. Zhang, S., et al., *Fezf2 promotes neuronal differentiation through localised activation of Wnt/ β -catenin signalling during forebrain development*. Development, 2014. 141(24): p. 4794.

46. Hurov, K.E., C. Cotta-Ramusino, and S.J. Elledge, *A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability*. Genes & development, 2010. 24(17): p. 1939-1950.

47. Takai, H., et al., *Tel2 Regulates the Stability of PI3K-Related Protein Kinases*. Cell, 2007. 131(7): p. 1248-1259.

48. Zeng, F., et al., *Identification of mouse MARVELD1 as a microtubule associated protein that inhibits cell cycle progression and migration*. Molecules and Cells, 2011. 31(3): p. 267-274.

Figure legends

Figure 1. Morphological observation of tiger pufferfish (*T. rubripes*). (A) Diploid group and (B) Haploid group produced by cold-shock. Bar = 0.5 mm. (C) Flow cytometry results of diploid (left panel) and cold-treated haploid (right panel) tiger pufferfish. (D) Genetic sex identification of tiger pufferfish. SNP genotypes of the *amhr2* gene in female and male diploids (left panel) were CC and CG, respectively. SNP genotypes in female and male haploids (right panel) were C and G, respectively.

Figure 2. DNA methylation of haploid and diploid tiger pufferfish (*T. rubripes*). (A) Percentages of mCs in the mCG, mCHG and mCHH contexts (1n-X was shown). (B) Methylation level of genome regions (1n-X was shown). (C) Distribution of mCG at the chromosome level in haploid and diploid tiger pufferfish. The red, yellow, blue, and green lines represent haploid female (1n-X) and male (1n-Y), and diploid female (2n-XX), and male (2n-XY), respectively.

Figure 3. 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons in tiger pufferfish (*T. rubripes*). (A) Length of DMRs between haploid and diploid female and male tiger pufferfish. (B) Venn diagram of haploid hyper-methylated/haploid hypo-methylated DMGs between males and females, respectively. (C) Top 5 enriched GO biological processes of haploid hyper-methylated/ haploid hypo-methylated DMGs shared in female and male tiger pufferfish.

Figure 4. Transcriptome analysis of haploid and diploid tiger pufferfish (*T. rubripes*). (A) Length distribution of assembled genes. (B) Number of DEGs in 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons. (C) Venn diagram of 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons. (D). GO classification of shared DEGs. (E) KEGG enrichment of shared DEGs.

Figure 5. Correlation between methylation and expression levels of target genes. The left Y-axis represents the expression level (TPM) of DEGs, and the right Y-axis represents methylation level of promoter of corresponding DEGs.

Figures

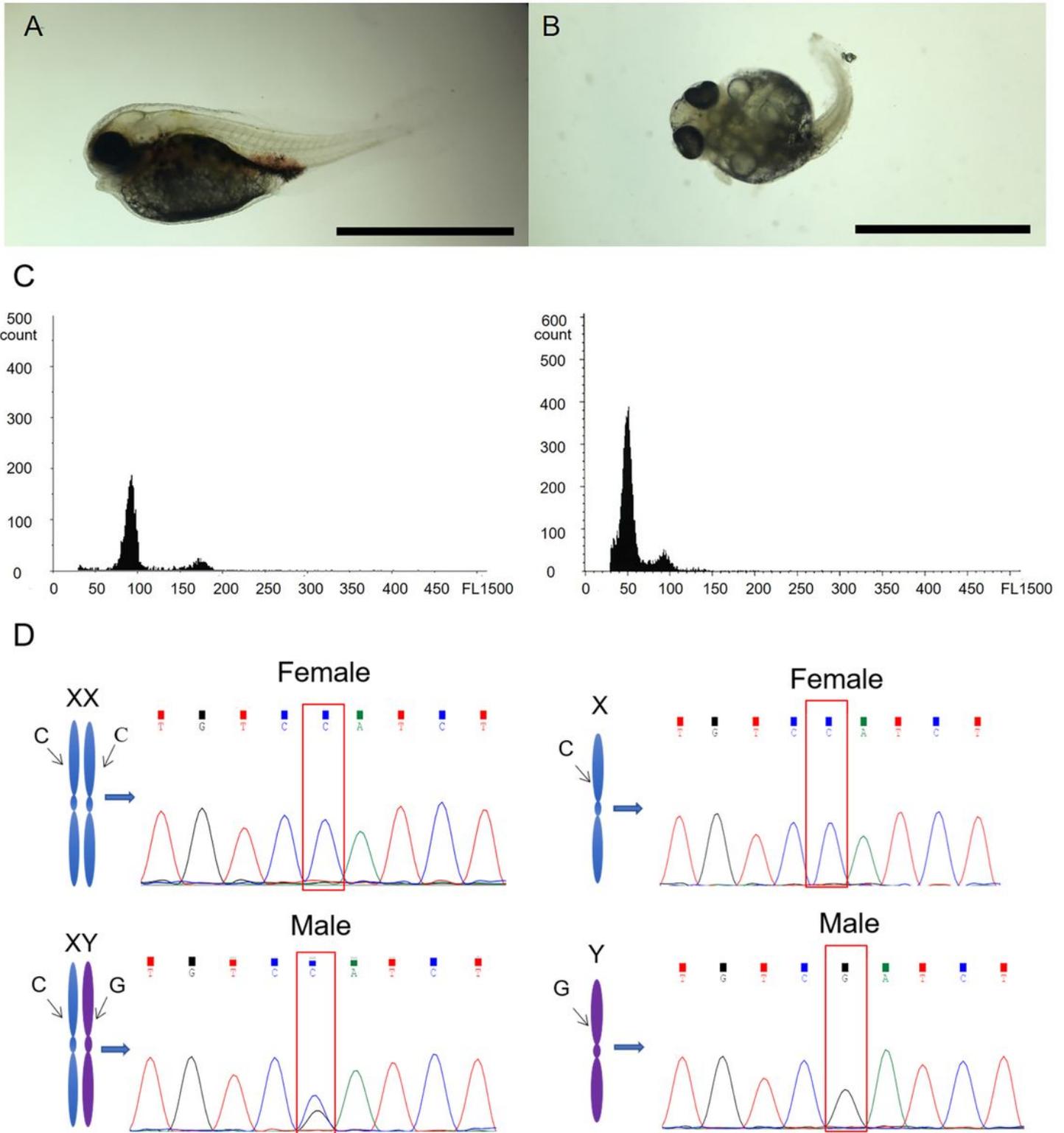


Figure 1

Morphological observation of tiger pufferfish (*T. rubripes*). (A) Diploid group and (B) Haploid group produced by cold-shock. Bar = 0.5 mm. (C) Flow cytometry results of diploid (left panel) and cold-treated haploid (right panel) tiger pufferfish. (D) Genetic sex identification of tiger pufferfish. SNP genotypes of

the *amhr2* gene in female and male diploids (left panel) were CC and CG, respectively. SNP genotypes in female and male haploids (right panel) were C and G, respectively.

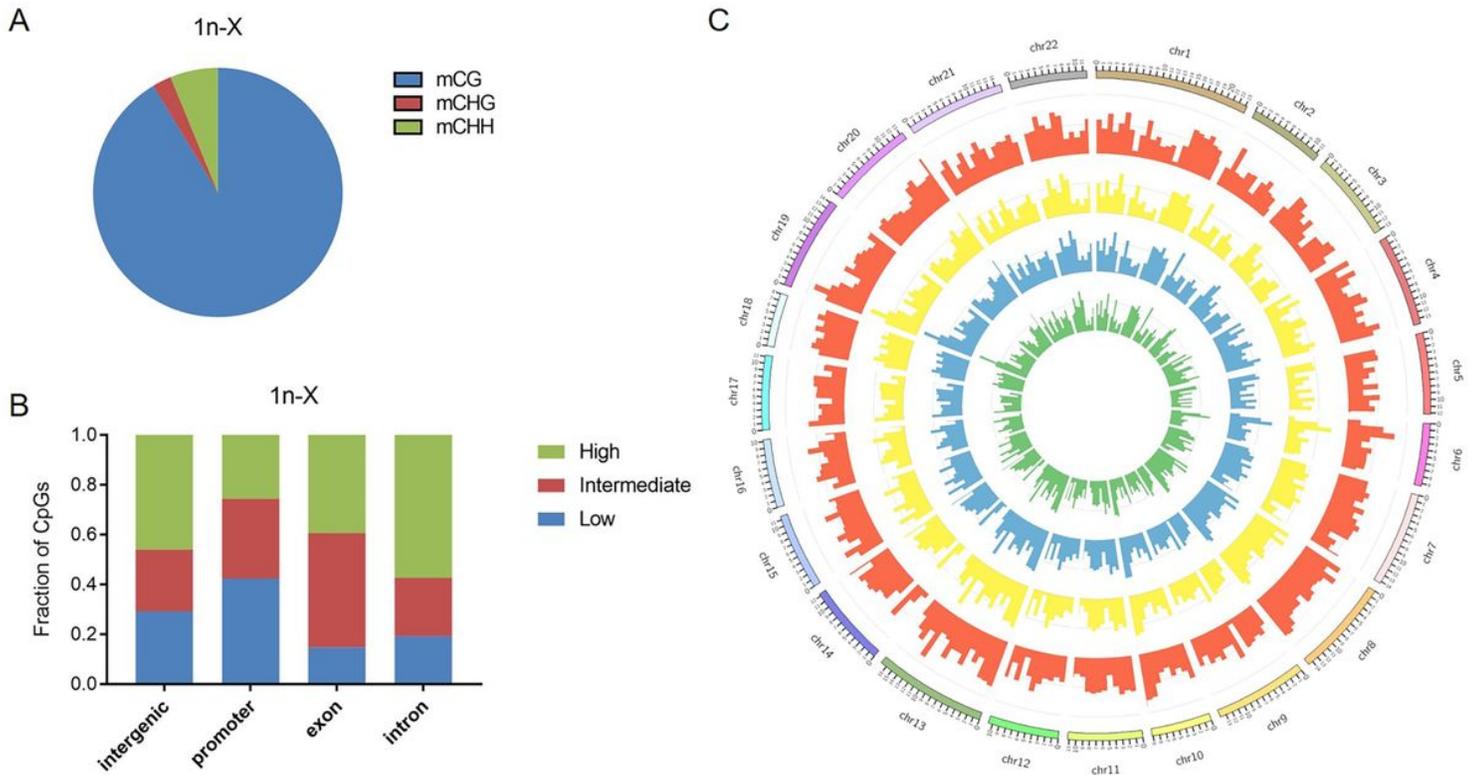


Figure 2

DNA methylation of haploid and diploid tiger pufferfish (*T.rubripes*). (A) Percentages of mCs in the mCG, mCHG and mCHH contexts (1n-X was shown). (B) Methylation level of genome regions (1n-X was shown). (C) Distribution of mCG at the chromosome level in haploid and diploid tiger pufferfish. The red, yellow, blue, and green lines represent haploid female (1n-X) and male (1n-Y), and diploid female (2n-XX), and male (2n-XY), respectively.

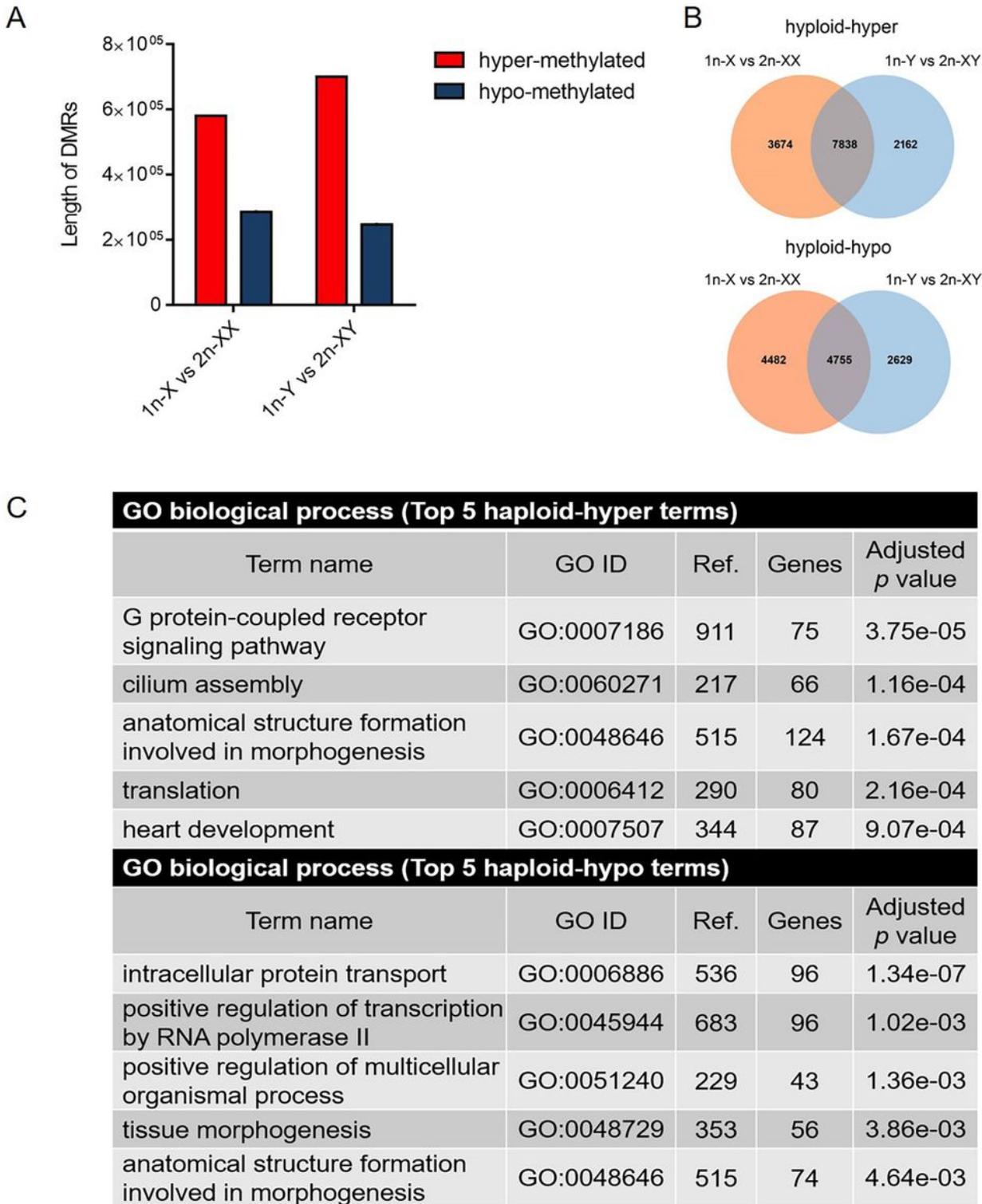


Figure 3

1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons in tiger pufferfish (*T. rubripes*). (A) Length of DMRs between haploid and diploid female and male tiger pufferfish. (B) Venn diagram of haploid hyper-methylated/haploid hypo-methylated DMGs between males and females, respectively. (C) Top 5 enriched GO biological processes of haploid hyper-methylated/ haploid hypo-methylated DMGs shared in female and male tiger pufferfish.

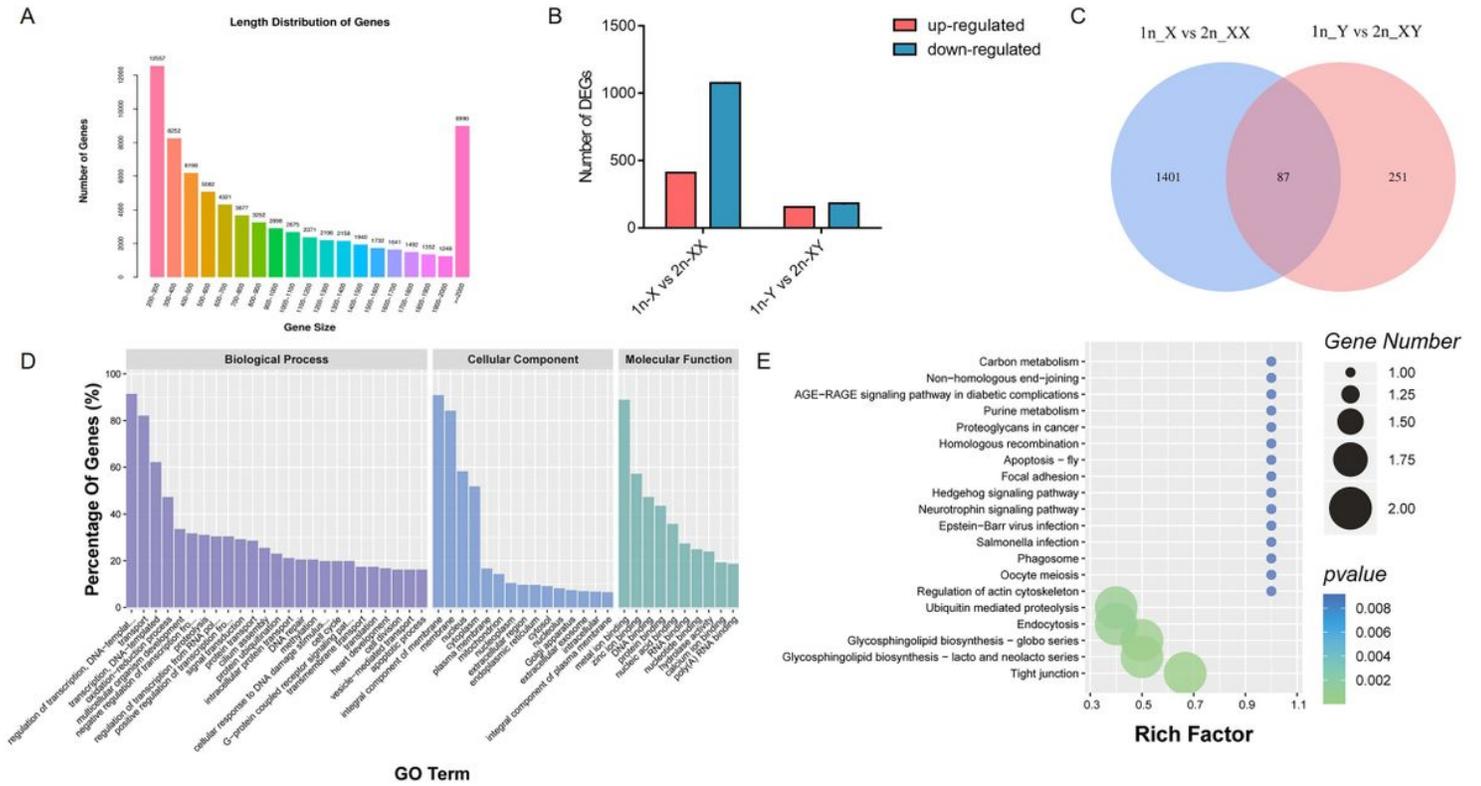


Figure 4

Transcriptome analysis of haploid and diploid tiger pufferfish (*T. rubripes*). (A) Length distribution of assembled genes. (B) Number of DEGs in 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons. (C) Venn diagram of 1n-X vs. 2n-XX and 1n-Y vs. 2n-XY comparisons. (D). GO classification of shared DEGs. (E) KEGG enrichment of shared DEGs.

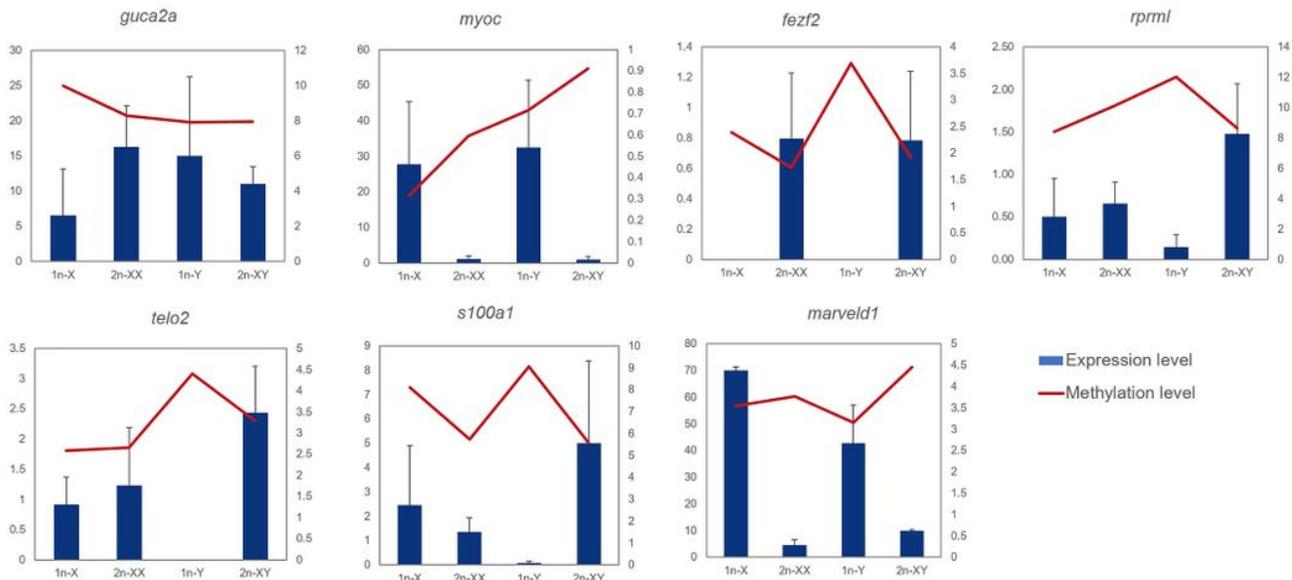


Figure 5

Correlation between methylation and expression levels of target genes. The left Y-axis represents the expression level (TPM) of DEGs, and the right Y-axis represents methylation level of promoter of corresponding DEGs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.jpg](#)
- [FigureS2.jpg](#)
- [FigureS3.jpg](#)
- [SupplementaryFile.docx](#)
- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.xlsx](#)
- [TableS5.docx](#)
- [TableS6.docx](#)
- [TableS7.docx](#)