

ORIGINAL ARTICLE

Molecular characteristics and transcriptional regulatory of spermatogenesis-related gene *RFX2* in adult Banna mini-pig inbred line (BMI)

Zhipeng Liu^{1#}, Hongmei Dai^{1#}, Hailong Huo^{2#}, Weizhen Li³, Yun Jiang¹, Xia Zhang^{4*}, Jinlong Huo^{1,5*}

¹College of Animal Science and Technology, Yunnan Agricultural University, Kunming, Yunnan, China

²Yunnan Vocational and Technical College of Agriculture, Kunming, Yunnan, China

³College of Veterinary Medicine, Yunnan Agricultural University, Kunming, Yunnan, China

⁴College of Life Science, Lyuliang University, Lvliang, Shanxi, China

⁵Department of Biology, University of Rochester, Rochester, New York, USA

#These authors contributed equally to this work.

How to cite: Liu Z, Dai H, Huo H, Li W, Jiang Y, Zhang X, Huo J. Molecular characteristics and transcriptional regulatory of spermatogenesis-related gene *RFX2* in adult Banna mini-pig inbred line (BMI). *Anim Reprod.* 2023;20(1):e20220090. <https://doi.org/10.1590/1984-3143-AR2022-0090>

Abstract

RFX2 plays critical roles in mammalian spermatogenesis and cilium maturation. Here, the testes of 12-month-old adult boars of Banna mini-pig inbred line (BMI) were subjected to whole-transcriptome sequencing. The results indicated that the average expression (raw count) of *RFX2* gene in BMI testes was 16138.25, and the average expression value of the corresponding transcript ENSST0000043271.2 was 123.1898. The CDS of *RFX2* obtained from BMI testes was 2,817 bp (GenBank accession number: OL362242). Gene structure analysis showed that *RFX2* was located on chromosome 2 of the pig genome with 19 exons. Protein structure analysis indicated that *RFX2* contains 728 amino acids with two conserved domains. Phylogenetic analysis revealed that *RFX2* was highly conserved with evolutionary homologies among mammalian species. Other analyses, including PPI networks, KEGG, and GO, indicated that BMI *RFX2* had interactions with 43 proteins involving various functions, such as in cell cycle, spermatid development, spermatid differentiation, cilium assembly, and cilium organization, etc. Correlation analysis between these proteins and the transcriptome data implied that *RFX2* was significantly associated with *FOXJ1*, *DNAH9*, *TMEM138*, *E2F7*, and *ATR*, and particularly showed the highest correlation with *ATR*, demonstrating the importance of *RFX2* and *ART* in spermatogenesis. Functional annotation implied that *RFX2* was involved in 17 GO terms, including three cellular components (CC), six molecular functions (MF), and eight biological processes (BP). The analysis of miRNA-gene targeting indicated that BMI *RFX2* was mainly regulated by two miRNAs, among which four lncRNAs and five lncRNAs competitively bound *ssc-miR-365-5p* and *ssc-miR-744* with *RFX2*, respectively. Further, the dual-luciferase report assay indicated that the *ssc-miR-365-5p* and *ssc-miR-744* significantly reduced luciferase activity of *RFX2* 3'UTR in the 293T cells, suggesting that these two miRNAs regulated the expression of *RFX2*. Our results revealed the important role of *RFX2* in BMI spermatogenesis, making it an intriguing candidate for follow-up studies.

Keywords: Banna mini-pig inbred line (BMI), whole-transcriptome sequencing, *RFX2*, functional annotation, transcriptional regulatory.

*Corresponding author: xiazhang1425@163.com; jinlonghuo973@163.com

Received: September 27, 2022. Accepted: December 31, 2022.

Financial support: This work was supported by funds from the National Natural Science Foundation of China (no. 32060733 to J.L.H.) and China Scholarship Council (CSC) (no. 201808535037 to J.L.H.).

Conflicts of interest: The authors have no conflict of interest to declare.



Copyright © The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Banna mini-pig inbred (BMI) line is an important laboratory animal resource and promising donor for porcine-to-human xenotransplantation (Wang et al., 2020; Huo et al., 2022). The extreme inbreeding method, i.e., full-sibling or parent-offspring mating has been performed since 1980 in an isolated environment at Xishuangbanna, Southwest China (Huo et al., 2012). In recent years, we found that the decreased reproductivity in boars had severely restricted the population expansion of BMI (Huo et al., 2022). Therefore, investigating the molecular characteristics and transcriptional regulations of genes expressed in BMI testis will fuel future projects clarifying the expression regulation mechanism of BMI spermatogenesis.

Testes are the gonads producing sperm and hormones in male mammals (Shima et al., 2004). Spermatogenesis is a complex and developmental process regulated delicately, including the proliferation and differentiation of spermatogonium, the meiosis of spermatocyte, and the spermiogenesis process of spermatids transforming into spermatozoa (Cho et al., 2001; Pan et al., 2005). Spermatogenesis depends on the precise regulation of thousands of testis-specific genes whose defection can lead to spermatogenesis disorder or male infertility (Cho et al., 2001; Pan et al., 2005). Screening important genes related to BMI spermatogenesis and studying their expression characteristics and regulatory roles are not only crucial to uncover molecular mechanisms underlying BMI spermatogenesis disorder but also promote gene therapy to solve decreased reproductivity in BMI boars.

Cilia are microtubule organelles in vertebrates and play multiple roles in sensory reception, signal transduction, and movement. There are two major patterns in mammalian cilia, including nonmotile "9+0" primary cilia and motile "9+2" cilia, among which the sperm cilia, named flagella, is a specially motile cilium with a typical "9+2" axonemes with the nine doublet microtubules surrounding a pair of microtubules (Satir and Christensen, 2007). Cilia in mammal testis play important roles specifically at the time of postnatal development, and the formation of the flagella is essential to sperm motility and fertilization (Girardet et al., 2019). *RFX2*, a member of *RFX2* factors, is a crucial factor of sperm flagellar assembly in mice (Emery et al., 1996a; Shawlot et al., 2015). *RFX2* contains a winged-helix DNA-binding domain (DBD) that recognizes the X-Box promoter motif (Choksi et al., 2014). *RFX2* is highly expressed in spermatocytes, especially round spermatids, and regulates the expression of hundreds of genes during spermatogenesis (Horvath et al., 2004). In cells, *RFX2* can control the surface expansion of multi-ciliated epithelial cells (MCCs) and coordinate the expression of some genes, and regulate cell movement, ciliogenesis, and ciliary function (Chung et al., 2014). *RFX2* can not only regulate ciliogenesis but also lead to left-right asymmetry through fluid flow in mice (Bisgrove et al., 2012). The deficiency of *RFX2* in mouse testis leads to the difficulty of sperm flagellar production and the failure of spermatid to differentiate from round into elongated, resulting in spermatogenesis disorder and male sterility (Kistler et al., 2015; Wu et al., 2016; Yu et al., 2018). In rats, *RFX2* binds to the *H1t* promoter, which is specifically expressed in testis and activate the transcription of *H1t* during spermatogenesis (Wolfe et al., 2006; Wolfe et al., 2008).

In this study, we performed transcriptome sequencing to obtain the expression level of *RFX2* mRNA, miRNAs, and lncRNAs of BMI testes, and carried out RT-PCR for amplifying the full-length coding sequence of *RFX2*. We analyzed the molecular characteristics of the *RFX2* gene and corresponding protein functions, and carried out protein-protein interaction and correlation analyses. We constructed the ceRNA transcriptional regulatory network of *RFX2* via annotating the *RFX2* gene and obtained the GO terms, miRNAs, and lncRNAs that interact with *RFX2*. To further validate whether the enriched miRNAs had targeting roles on *RFX2*, we characterized the reliability of results using the dual-luciferase report assay. Accordingly, this study underlines the importance of the high expression of *RFX2* in BMI testis, providing a valuable resource for further investigation of potential mechanisms and functions of *RFX2* gene in the process of BMI spermatogenesis.

Methods

Sample collection

Three 12-month-old adult BMI boars were randomly chosen, and their testis samples were obtained by surgical castration and washed with PBS buffer. After removing the fat and fascia tissues, small aliquots were crosscut from the middle area of the parenchyma, including the seminiferous tubules and immediately immersed into liquid nitrogen and transferred to -80°C freezer for storage. All animal procedures were approved by the Research Ethics Committee of Yunnan Agricultural University (No. YNAUREC2020224), and conducted according to the guideline for care and use of laboratory animals established by the National Research Council (NRC).

Transcriptome sequencing and expression analysis of *RFX2* gene

Total RNAs of testicular samples were extracted with oligo(dT) beads (NEB, USA #E7490S). The libraries of RNA-seq and miRNAs were constructed by Novogene (Tianjin, China) and sequenced on Illumina Hiseq 4000 and Novaseq6000 platforms, respectively. The whole RNA-seq analysis included quality control of raw data using the fastp software, filtering low-quality data, and removing adaptors, sequences with N ratio greater than 10% and with all A bases. The filtered data were aligned with the rRNA reference sequence (NCBI, 2022) using Bowtie2 (V.2.1.0), and then removed the sequences that were matched with reference sequences. The pig reference genome (*Sus_scrofa*.Sscrofa11.1.dna.toplevel.fa) and the annotations file (*Sus_scrofa*.Sscrofa11.1.102.gtf) were downloaded from Ensembl, and the genomic index was constructed using STAR (v.2.5.2a). The data removed rRNA sequences were aligned with the pig reference genome. The original expression and corrected TPM values were calculated with FeatureCounts (V.2.0.1) and Salmon (V.1.5.1), respectively. The expression of *RFX2* transcript ENSST00000043271.2 in BMI was obtained and visualized using the Gviz 1.40.1 package of R. The quality control analysis of Small RNA-seq was similar to the above analysis of RNA-Seq, and then removed the reads matching with porcine rRNAs, tRNAs, snRNAs, and snoRNAs of the RFAM14.8 database (<http://rfam.xfam.org/>), and the remaining sequences were aligned to miRBase22.1, and miRNAs were quantified using miRdeep2 (V2.0.1.3).

Gene amplification of *RFX2*

Two pairs of primers of the *RFX2* gene were synthesized based on the transcript sequence NSSST00000043271.2, including F1: GATCCGGTTCCTAATAACTGAGCA, R1: ACCAGCCTTCCAAGCTCT; F2: GAGCATCACACTGCAGGACGTCA, R2: CGCGGTGGTCACAACACTGTTT). The coding region sequence of *RFX2* gene was amplified using cDNA of BMI testis, and the full-length CDS sequence was spliced using Lasergene7.0. The reaction system 25 µL: Premix Taq™ 12.5 µL; 10 µM primer F and primer R 1 µL, respectively; 25 ng/µL cDNA 1 µL; H₂O 9.5 µL. Amplification program: 95 °C 5 min; 95 °C 30 sec, 57 °C 45 sec, 72 °C 80 sec, 35 cycles; 72 °C 10 min. The products were sequenced by Kunming Tsingke Biotechnology Co., Ltd.

Functional analysis of *RFX2* protein

The complete coding region sequence of *RFX2* was obtained by splicing the amplification sequences of F1/R1 and R2/R2 using Lasergene7.1, and the open reading frame (ORF) of

RFX2 was obtained using NCBI's ORFfinder. The molecular weight (Mw), molecular formula, isoelectric point (PI), and numbers of positively and negatively charged residues of the *RFX2* protein were analyzed using the ProtParam. The secondary structure, hydrophobic residues, functional site, transmembrane helix, signal peptide and tertiary structure *RFX2* protein were analyzed using SOPMA, ProtScale, Prosite, TMHMM 2.0, SignalP 5.0 and I-TASSER, respectively. The phylogenetic tree of *RFX2* protein was generated using MEGA-X and visualized using ITOL (V6). The conserved domain of the multi-species amino acid sequence of *RFX2* protein was analyzed using WebLogo3. The protein-protein interaction analysis was carried out using String11.5, and the items with $P < 0.05$ were subjected to GO and KEGG enrichment analysis using clusterProfiler4.0 package of R. The correlation between the proteins enriched above and the expression levels of *RFX2* obtained from RNA-seq was calculated using cor.test function of R.

Transcriptional regulation analysis of *RFX2*

The GO function annotation of *RFX2* was carried out using EggNOG-mapper (V2) and Uniprot. miRNAs and lncRNAs regulating *RFX2* were obtained by analyzing RNA-seq data using miRanda 3.3 and RNAhybrid 2.1.2, and the ceRNA network of transcriptional regulation was visualized using Cytoscape 3.8.2.

Dual-luciferase reporter assay

To verify the reliability of the miRNAs obtained above acting on the 3'UTR region of the *RFX2*, we constructed the plasmids including pMIR-REPORT Luciferase-H306, pMIR-REPORT Luciferase-*RFX2* 3'UTR(WT)-H22978 and pMIR-REPORT Luciferase-*RFX2* 3'UTR(MUT1)-H22980. The 293T cells subcultured to 70% confluence were inoculated into 96-well cell culture plates, and cultured in DMEM+10% FBS, 37°C and 5% CO₂ for 24 h. The constructed wild-type and mutant recombinant dual-luciferase reporter vectors were co-transfected into 293T cells respectively, according to the instructions of Lipofectamine™2000 kit (11668019, Invitrogen). The luciferase activity was detected by the dual-luciferase reporter assay kit (E1910, Promega) in the dark. The data were calculated using the ratio of firefly luciferase activity to the renilla luciferase activity. A two-tailed t-test was performed on 3 replicates for each experiment.

Results

Expression characteristics of *RFX2* gene

The average expression (raw count) level of *RFX2* gene in BMI testis identified using RNA-seq was 16,138.25, and the average expression (transcripts per million, TPM) value of the corresponding transcript ENSSSCT00000043271.2 was 123.1898. *RFX2* gene was located on chromosome 2 of the pig genome Sscrofa11.1 with a full-length of 1011,26 bp. The gene annotation using Gviz 1.40.1 revealed that ENSSSCT00000043271.2 had 19 exons and 18 introns, and three BMI samples had consistently high positive expression (Figure 1A). The fragments of 1,491 bp and 1,251 bp of *RFX2* were obtained using primers F1/R1 and F2/R2, respectively (Figure 1B). Further analysis revealed that the 2,187 bp ORF1 encoding 728 amino acids was a full-length ORF (Figure 1C).

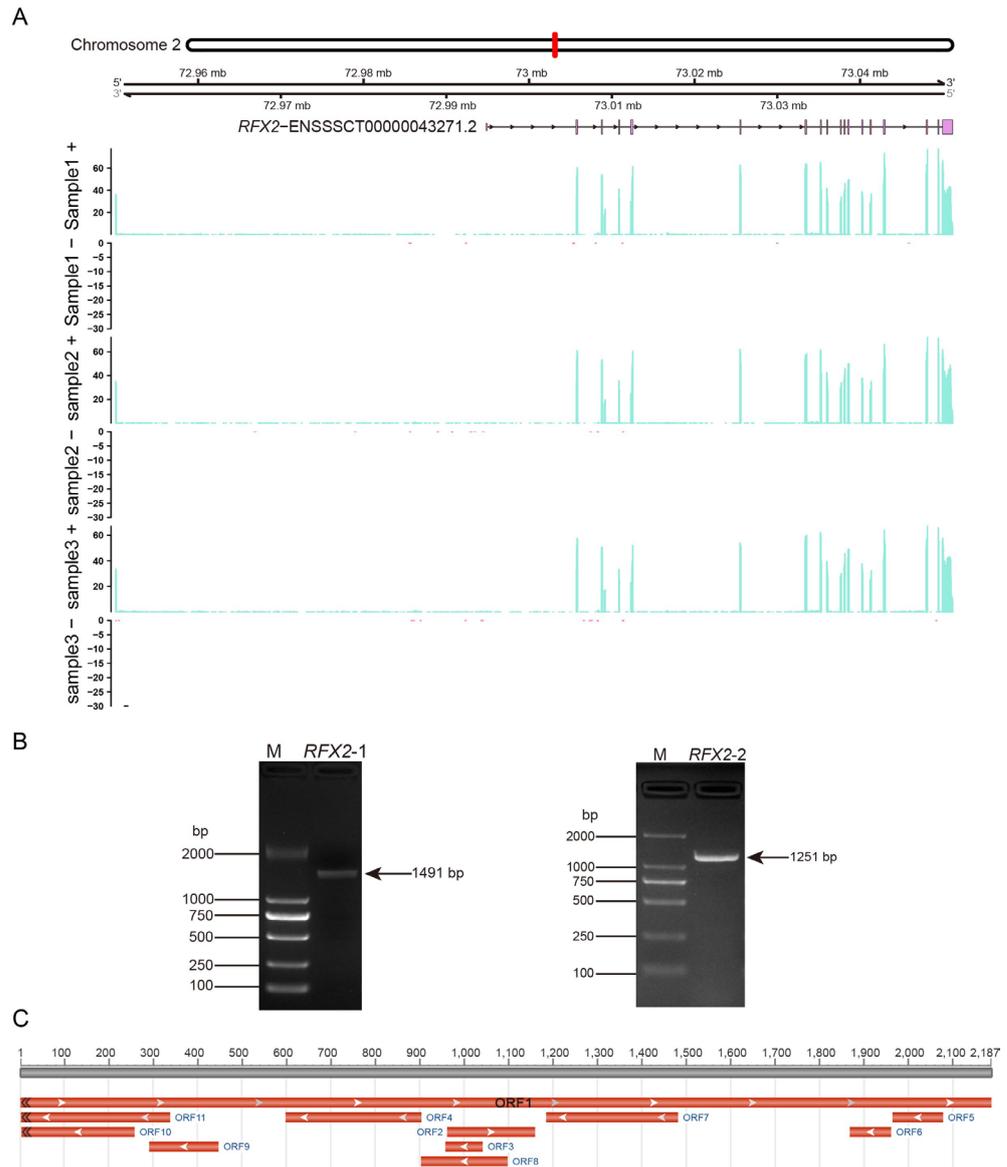


Figure 1. Expression analysis of *RFX2* gene and its CDS amplification and ORF acquisition. (A) chromosome location and exon, intron abundance of *RFX2* based on RNA-seq; (B) RT-PCR product of amplifying the CDS of *RFX2* gene. M. DL2000 DNA Marker. *RFX2-1* and *RFX2-1*, PCR products; (C) ORF analysis of *RFX2* cDNA.

Structural and functional analysis of RFX2 protein

The analysis of protein structure indicated the molecular weight (Mw), molecular formula and isoelectric point (PI), negative charge residues, and positive charge residues of BMI RFX2 were 80.297 kD, $C_{354}H_{553}N_{1007}O_{1077}S_{26}$, 6.42, 69, 62, respectively. RFX2 protein had the largest hydrophobic value of 2.056 at position 684 and the least hydrophobic value of -3.222 at position 719, and the N-terminus and C-terminus were hydrophobic and RFX2 contained active sites such as enzymatic phosphorylation and amidation, without signal peptides or transmembrane structures.

In the secondary structure of RFX2 protein, α -helix accounted for the highest proportion with 336 amino acids for 46.15%; random coil with 302 amino acids for 41.48%; extended strand with 67 amino acids for 9.20%; β turn of the least with 23 amino acids for 3.16% (Figure 2A). Structurally, the composition of the tertiary structure of RFX2 protein was similar to that of the secondary structure, including random coil, α -helix, β sheet, β turn, and extended chain (Figure 2B). RFX2 contained two conserved domains, RFX1_trans_act and RFX_DNA_binding (Figure 2C).

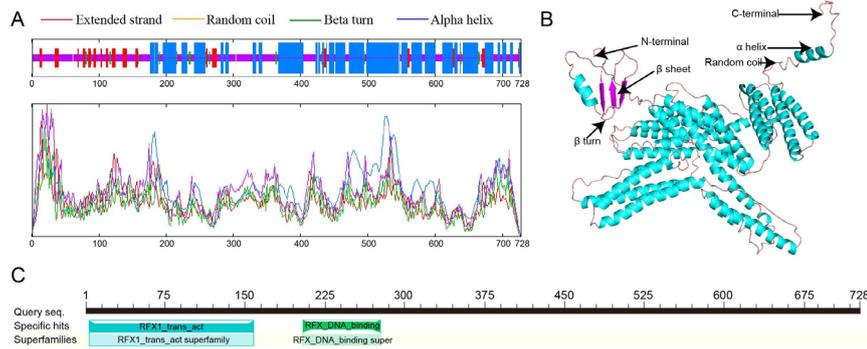


Figure 2. The analysis of spatial structure of RFX2 protein. (A) Secondary structure. The red, pink, green and blue vertical lines represents extended strand, random coil, beta turn and alpha helix, respectively; (B) Tertiary structure. The N-terminal, C-terminal, α helix, β sheet, β turn and random coil are indicated with black arrows; (C) The conserved domains, including RFX1_trans_act and RFX_DNA_binding.

Multi-species amino acid sequence homology of RFX2

Amino acid sequence homology analysis revealed that more than 85% of RFX2 sequences in 40 mammalian species were identical. Phylogenetic analysis indicated that porcine RFX2 was grouped with *Balaenoptera-musculus*, *Lipotes-vexillifer*, etc., and then grouped with Bovidae (Figure 3A). The analysis of conserved domains of RFX2 proteins revealed that two conserved domains (RFX1_trans_act at amino acids 4-159 and RFX_DNA_binding at amino acids 202-279) were shared among 40 mammalian species. Moreover, these two conserved domains showed rather small amino acid differences among 40 species, indicating their high conservation in evolution (Figure 3B).

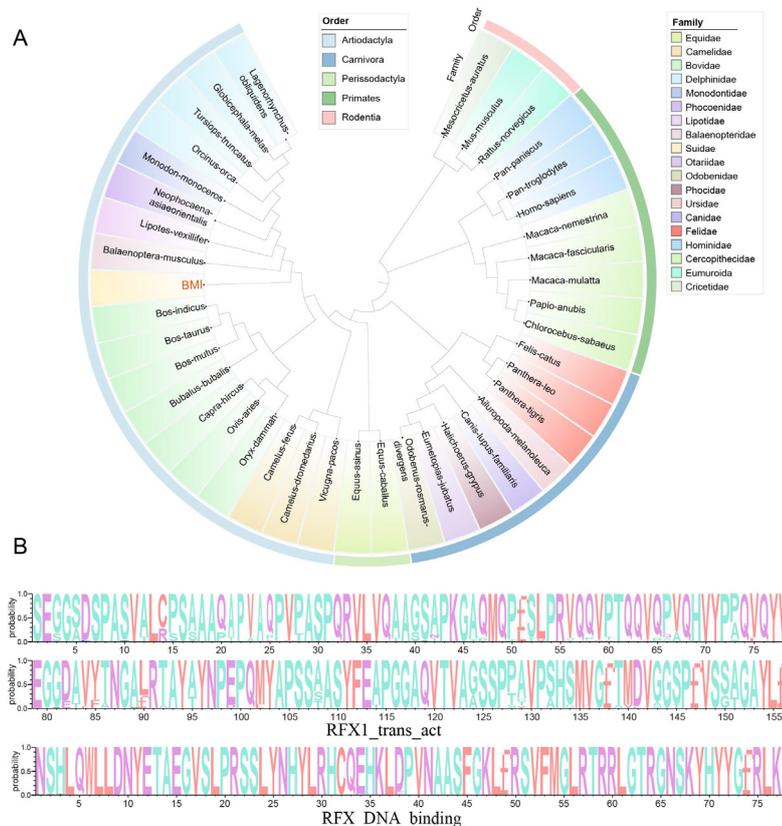


Figure 3. Amino acid sequences analysis of RFX2 from different mammals. (A) Phylogenetic tree for RFX2 from 40 mammalian species; (B) The analysis of conserved domains (RFX1_trans_act and RFX_DNA_binding) among 40 mammalian species using WebLogo3.

Protein interaction network and gene correlation analysis

Protein-Protein Interaction Networks Functional Enrichment Analysis revealed potential interactions between *RFX2* and 43 proteins (Figure 4A). To further investigate the functions of *RFX2*, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment were implemented. KEGG enrichment analysis revealed that these proteins were mainly involved in cell cycle, TGF-beta signaling pathway, cellular senescence and other pathways (Figure 4B). GO enrichment revealed that these proteins were mainly involved in spermatid development, spermatid differentiation, cilium assembly and cilium organization, etc. (Figure 4C). Subsequently, we analyzed the correlation between these proteins obtained above and our RNA-seq data and constructed the expression correlation network (Figure 4D), and we found that *RFX2* was significantly correlated with *FOXJ1*, *DNAH9*, *TMEM138*, *E2F7* and *ATR* (Table 1).

Table 1. Genes that significantly correlating with the *RFX2*.

Node1	Node2	Correlation	Pvalue
<i>RFX2</i>	<i>FOXJ1</i>	0.997465	0.045336
<i>RFX2</i>	<i>DNAH9</i>	0.998620	0.033444
<i>RFX2</i>	<i>TMEM138</i>	0.998728	0.032110
<i>RFX2</i>	<i>E2F7</i>	0.999118	0.026740
<i>RFX2</i>	<i>ATR</i>	0.999916	0.008247

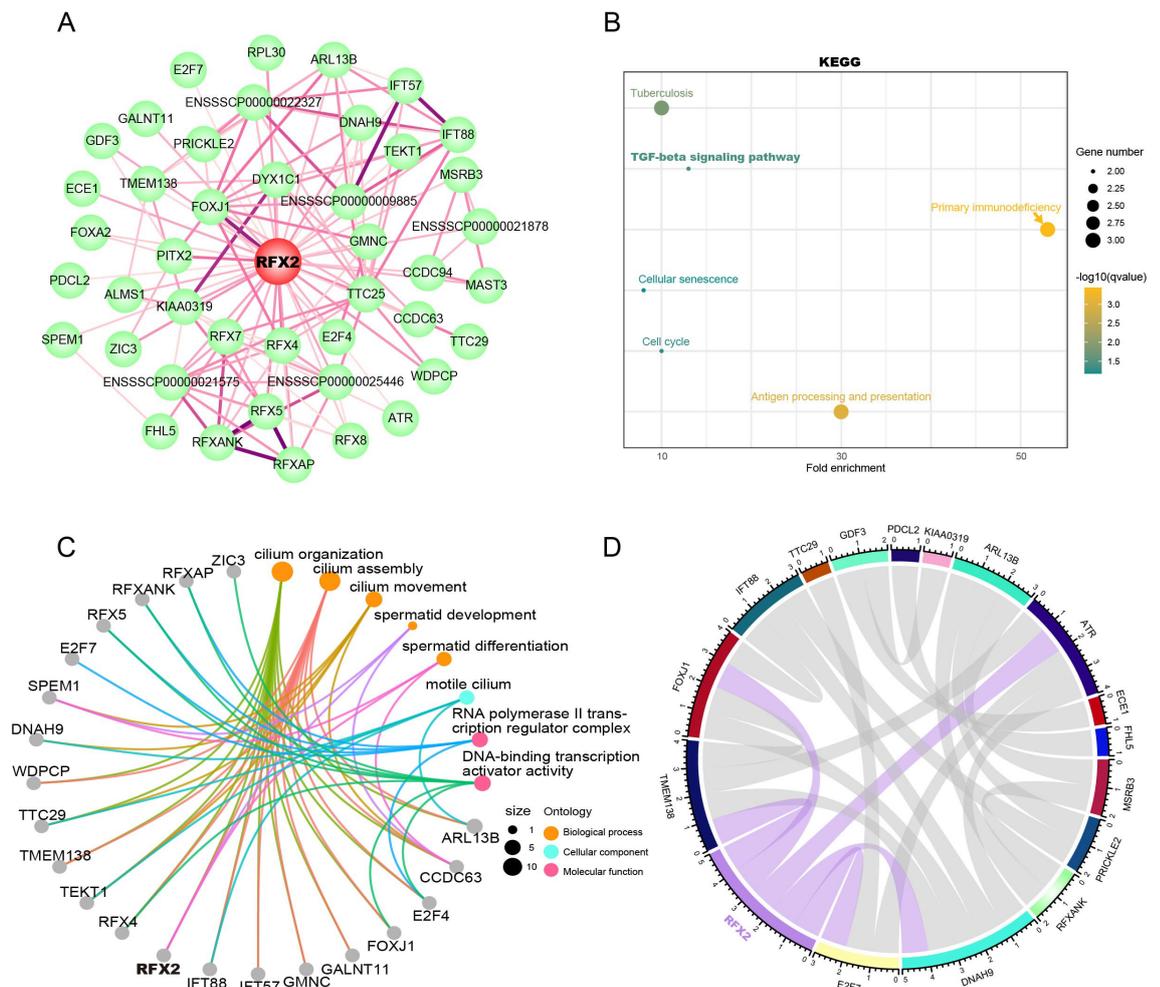


Figure 4. Interaction network of *RFX2*. (A) Protein-protein Interaction networks. The lines indicate the interactions among proteins, and more lines indicate higher confidence; (B) Interaction network of KEGG

enrichment of *RFX2*; (C) Interaction network of GO enrichment of *RFX2*, including proteins and the pathways related to biological process, cellular component and molecular function; (D) Chord plot of genes correlating with the *RFX2*. The pink lines in the circle indicate a significant correlation with *RFX2*.

ceRNA regulatory network of *RFX2*

Functional annotation indicated that *RFX2* mRNA mainly involved three GO terms, including cytoplasm, nucleus, and chromatin in the Cellular Component (CC), and six GO terms including sequence-specific double-stranded DNA binding, RNA polymerase II cis-regulatory region sequence-specific DNA binding, DNA-binding transcription factor activity, DNA-binding transcription factor activity RNA polymerase II-specific, DNA binding and protein binding in the Molecular Function (MF), and eight GO terms including spermatogenesis, cell projection, cell differentiation, acrosome assembly, cilium assembly, spermatid development, regulation of transcription DNA-templated, regulation of transcription by RNA polymerase II in the Biological Process (BP) (Figure 5). *RFX2* was mainly regulated by two miRNAs, including *ssc-miR-365-5p* and *ssc-miR-744*. Of note, we found that four (ENSSSCG00000042985.1, ENSSSCG00000047419.1, ENSSSCG00000041557.1, ENSSSCG00000047216.1) and five lncRNAs (ENSSSCG00000045377.1, ENSSSCG00000040582.2, ENSSSCG00000041587.1, ENSSSCG00000036820.2, ENSSSCG00000042991.1) competitively bound *ssc-miR-365-5p* and *ssc-miR-744* with *RFX2*, respectively (Figure 5).

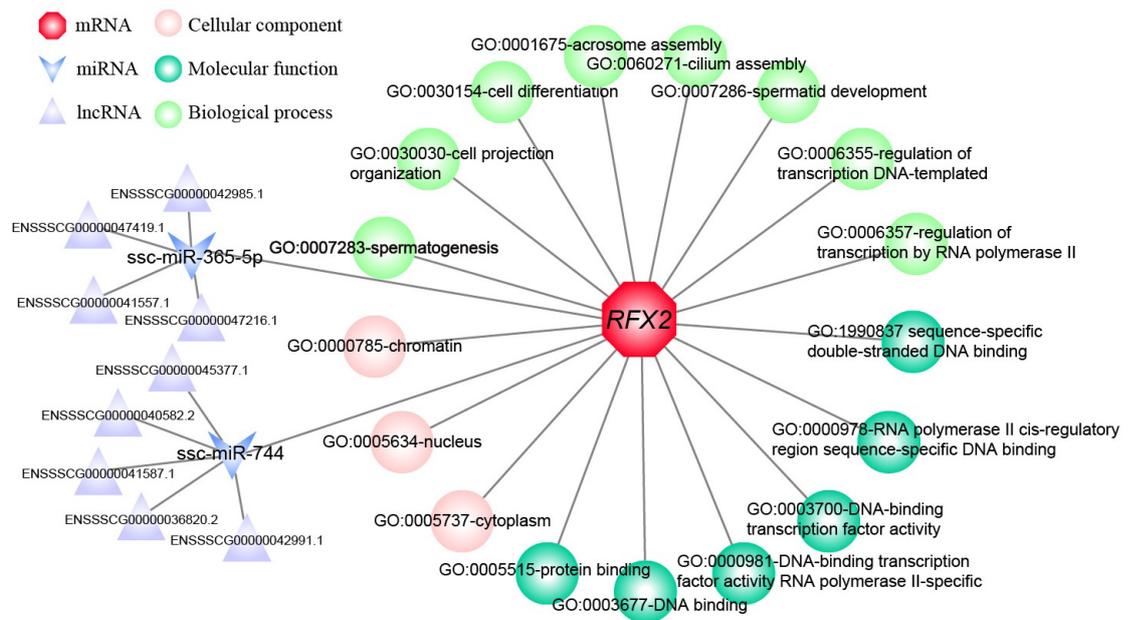


Figure 5. The functional annotation of BMI *RFX2* and potential ceRNA regulatory network. The hexagon, “V”, triangle, sphere, respectively represent *RFX2* mRNA, miRNA, lncRNA and GO pathways including cellular component and molecular function, biological process.

Validation of *RFX2*-miRNAs targeting relationship

To identify the targeting effect of miRNA on *RFX2*, we verified the function of *ssc-miR-365-5p* and *ssc-miR-744* based on the dual-luciferase reporter assay. Compared with the luciferase assay performed on the *RFX2* 3'UTR mutant reporter, the luciferase activity of 3'UTR wild-type reporter vector was significantly decreased by *ssc-miR-365-5p* ($P < 0.001$) in 293T cells (Figure 6). Similarly, *ssc-miR-744* also significantly reduced the luciferase activity of *RFX2* 3'UTR wild-type reporter vector in 293T cells ($P < 0.001$), but the mutant reporter vector was not affected ($P > 0.05$), suggesting that there may be other unknown binding sites of *ssc-miR-744* in the *RFX2* 3'UTR. These results indicated that *ssc-miR-365-5p* and *ssc-miR-744* did regulate the expression of *RFX2*.

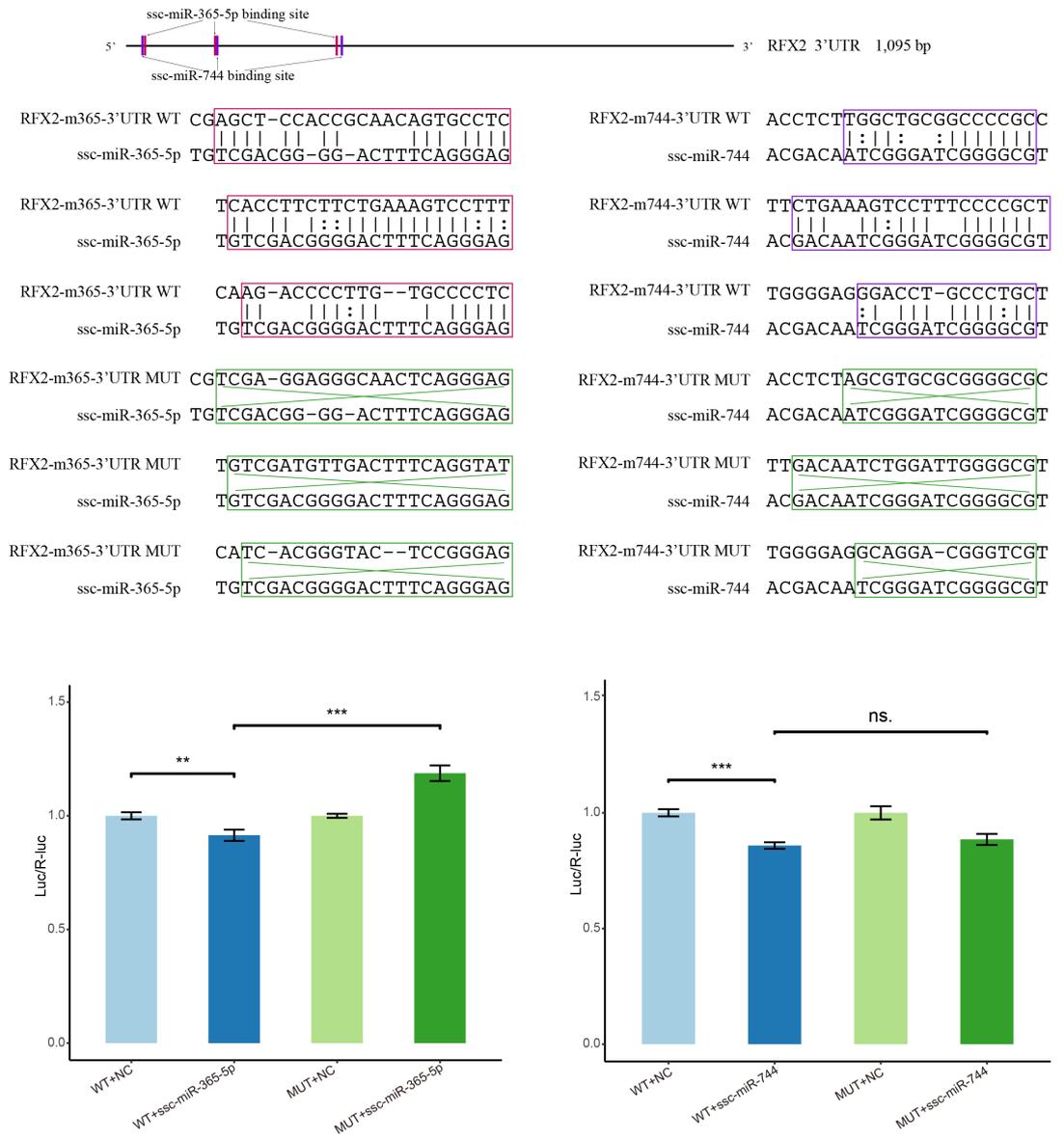


Figure 6. The validated results of *RFX2* regulation by miRNA ssc-miR-365-5p and ssc-miR-744. The left and right indicate ssc-miR-365-5p and ssc-miR-744, respectively. ns.: no significant difference. **, $P < 0.01$. ***, $P < 0.001$.

Discussion

RFX factors typically share a winged-helix DNA-binding domain (DBD), which can recognize the minor groove of promoter DNA by monomer or dimers form (Gajiwala et al., 2000). DBD is a highly conserved motif of dimeric binding protein, which has various functions in regulating unicellular organisms and eukaryotic organisms (Emery et al., 1996b; Cornille et al., 1998). RFX containing DBD motif is a regulator that can bind to MHC class II genes and regulate their expression (Nekrep et al., 2002). Domain RFX1_trans_act locates the transcriptional activation region of the N-terminal of RFX2, binds to DNA and activates transcription (Katan et al., 1997). The amino acid sequences of the two domains were slightly different among the 40 species, indicating that these two domains are highly conserved among mammals.

Protein-protein interaction analysis revealed that RFX2 interacted with 43 proteins, and KEGG enrichment analysis indicated that these proteins were involved in cell cycle, TGF-beta signaling pathway, cellular senescence, etc. GO enrichment analysis revealed that these proteins were involved in spermatid development, spermatid differentiation, cilium assembly,

DNA-binding transcription activator activity, RNA polymerase II transcription regulator complex, motile cilium and cilium movement, etc. Together, these signaling pathways formed a complex network to regulate spermatogenesis and ciliogenesis. After querying the gene expression levels of these proteins in our RNA-seq data, we calculated the correlation of expression levels. Our results showed that *RFX2* was significantly correlated with *ATR*, *E2F7*, *TMEM138*, *DNAH9* and *FOXJ1*. *ATR* plays important roles in the cell-cycle and early mammalian development. Targeted disruption of *ATR* leads to the death of early mice embryos (de Klein et al., 2000). In humans, when DNA is damaged, the kinase encoded by *ATR* gene is activated as the checkpoint transducer to regulate cell cycle and DNA replication through phosphorylating a variety of substrates, enabling cell survival (Karnitz and Zou, 2015). *E2F7* is a transcriptional repressor encoding cell cycle-related proteins, which inhibits cell proliferation by regulating the transcription of miRNAs (Mitxelena et al., 2016). *E2F7* is essential to embryonic development and cell survival, and its dysfunction leads to apoptosis and dilation of blood vessels, resulting in embryonic death in mice (Li et al., 2008). *TMEM138* is a connecting protein for cilium transport and biogenesis, and mice with mutated *TMEM138* suffer from impaired primary cilia, learning and memory, resulting in Joubert syndrome (Guo et al., 2022). *DNAH9* expressed on sperm tail is an important heavy chain component of the dynein arm of cilia and flagella, and mutation of *DNAH9* leads to primary ciliary dyskinesia (PCD) and nonsyndromic asthenozoospermia (Fassad et al., 2018; Tang et al., 2021). *FOXJ1* plays an important role in the cilia gene express by synergistically cooperating with *RFX2* (Koay et al., 2021). *RFX2* stabilizes the proximal promoter and distal enhancer and facilitates the core domains bound by *FOXJ1* enabling the functional roles of cilia genes (Quigley and Kintner, 2017). Taken together, the findings of these proteins interacting with *RFX2* will provide information and research directions for shedding further light on the function and molecular mechanism underlying the *RFX2* in BMI.

Spermatogenesis is a complex biological process, which is regulated not only by mRNAs but also by ncRNAs (Gao et al., 2020). microRNAs (miRNAs) are highly conserved endogenous non-coding small RNAs with ~22 nt length that regulate gene expression post-transcriptionally (Calin and Croce, 2006; Gao et al., 2020). A growing number of reports have revealed that miRNAs are required for primordial germ cell development and spermatogenesis and play vital roles in the development of testis and spermatozoa as potential biomarkers (Hayashi et al., 2008). The functional annotation of *RFX2* and analysis of the regulatory network on competing endogenous RNA (ceRNA) indicated that *RFX2* was targeted by two miRNAs, *ssc-miR-365-5p* and *ssc-miR-744*.

miR-365-5p regulated some genes involved in modifying miRNA biogenesis (Singh and Storey, 2021), and *miR-365-5p* in porcine milk prevented the damage of intestinal epithelial cells via regulating the expression of p53 (Xie et al., 2020). *miR-365-5p* in cervical cancer was also a benign regulator that inhibited the production of Interleukin 6 (Arenas-Padilla and Mata-Haro, 2018). *miR-744* was a potential prognostic marker for patients with liver cancer and inhibited the growth of Hepatocellular carcinoma (HCC) cells by targeting c-MYC (Tan et al., 2015). As tumor suppressor microRNA, *miR-744* inhibited the proliferation of colorectal cancer cells by targeting RFC2 (Hu et al., 2020). *miR-744-5p* inhibited multiple myeloma proliferation by targeting the *SOX12/Wnt/β-catenin* pathway (Guo et al., 2021), and inhibited the proliferation and metastasis of HCC by targeting TGF-β1 (Huang et al., 2021). Conversely, *miR-744* increased the risk of pancreatic cancer, prostate cancer and nasopharyngeal carcinoma by activating the *Wnt/β-catenin* pathway, AMPK signaling, and transcriptional regulation of ARHGAP5, respectively (Zhou et al., 2015; Zhang et al., 2019). The results identified by the dual-luciferase reporter assay indicated that *ssc-miR-365-5p* and *ssc-miR-744* regulated the expression of *RFX2*, which will pave the way for further study on the mechanism of *ssc-miR-365-5p* and *ssc-miR-744* targeting *RFX2*.

Conclusion

In this study, RNA-seq technology was used to obtain the expression level of *RFX2* gene in the adult BMI testis and the full-length coding sequence of *RFX2* was obtained by RT-PCR. The molecular characteristics of *RFX2* gene and corresponding conserved protein domains and

evolutionary relationships among multiple mammals were obtained by RNA-seq data combined with bioinformatics analysis. *RFX2* was mainly involved in cell cycle, TGF-beta signaling pathway, cellular senescence and spermatid development, spermatid differentiation, cilium assembly and cilium organization, etc. *RFX2* was significantly correlated with *FOXJ1*, *DNAH9*, *TMEM138*, *E2F7* and *ATR*, in particular, with the strongest correlation with *ATR* protein. *RFX2* was involved in 17 GO terms, including 3 cellular components, 6 molecular functions and 8 biological processes. The dual-luciferase reporter assay verified that *RFX2* was mainly regulated by two miRNAs, *ssc-miR-365-5p* and *ssc-miR-744*. These results broaden the understanding of transcriptional regulatory characteristics of spermatogenesis-related gene *RFX2*, thus building a foundation for further studying function and molecular mechanism of *RFX2* in BMI testis.

Acknowledgements

The authors would like to thank Dr. Yu H Sun, PhD (University of Rochester) for his help in the technical support, reviewing and editing this manuscript.

References

- Arenas-Padilla M, Mata-Haro V. Regulation of TLR signaling pathways by microRNAs: implications in inflammatory diseases. *Cent Eur J Immunol*. 2018;43(4):482-9. <http://dx.doi.org/10.5114/ceji.2018.81351>. PMID:30799997.
- Bisgrove BW, Makova S, Yost HJ, Brueckner M. *RFX2* is essential in the ciliated organ of asymmetry and an *RFX2* transgene identifies a population of ciliated cells sufficient for fluid flow. *Dev Biol*. 2012;363(1):166-78. <http://dx.doi.org/10.1016/j.ydbio.2011.12.030>. PMID:22233545.
- Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer*. 2006;6(11):857-66. <http://dx.doi.org/10.1038/nrc1997>. PMID:17060945.
- Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi Y-C, Hecht NB, Eddy EM. Haploinsufficiency of protamine-1 or-2 causes infertility in mice. *Nat Genet*. 2001;28(1):82-6. <http://dx.doi.org/10.1038/ng0501-82>. PMID:11326282.
- Choksi SP, Lauter G, Swoboda P, Roy S. Switching on cilia: transcriptional networks regulating ciliogenesis. *Development*. 2014;141(7):1427-41. <http://dx.doi.org/10.1242/dev.074666>. PMID:24644260.
- Chung M-I, Kwon T, Tu F, Brooks ER, Gupta R, Meyer M, Baker JC, Marcotte EM, Wallingford JB. Coordinated genomic control of ciliogenesis and cell movement by *RFX2*. *eLife*. 2014;3:e01439. <http://dx.doi.org/10.7554/eLife.01439>. PMID:24424412.
- Cornille F, Emery P, Schüler W, Lenoir C, Mach B, Roques BP, Reith W. DNA binding properties of a chemically synthesized DNA binding domain of hRFX1. *Nucleic Acids Res*. 1998;26(9):2143-9. <http://dx.doi.org/10.1093/nar/26.9.2143>. PMID:9547272.
- de Klein A, Muijtjens M, van Os R, Verhoeven Y, Smit B, Carr A, Lehmann A, Hoeijmakers J. Targeted disruption of the cell-cycle checkpoint gene *ATR* leads to early embryonic lethality in mice. *Curr Biol*. 2000;10(8):479-82. [http://dx.doi.org/10.1016/S0960-9822\(00\)00447-4](http://dx.doi.org/10.1016/S0960-9822(00)00447-4). PMID:10801416.
- Emery P, Durand B, Mach B, Reith W. RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Res*. 1996a;24(5):803-7. <http://dx.doi.org/10.1093/nar/24.5.803>. PMID:8600444.
- Emery P, Strubin M, Hofmann K, Bucher P, Mach B, Reith W. A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. *Mol Cell Biol*. 1996b;16(8):4486-94. <http://dx.doi.org/10.1128/MCB.16.8.4486>. PMID:8754849.
- Fassad MR, Shoemark A, Legendre M, Hirst RA, Koll F, le Borgne P, Louis B, Daudvohra F, Patel MP, Thomas L, Dixon M, Burgoyne T, Hayes J, Nicholson AG, Cullup T, Jenkins L, Carr SB, Aurora P, Lemullois M, Aubusson-Fleury A, Papon JF, O'Callaghan C, Amselem S, Hogg C, Escudier E, Tassin AM, Mitchison HM. Mutations in outer dynein arm heavy chain *DNAH9* cause motile cilia defects and situs inversus. *Am J Hum Genet*. 2018;103(6):984-94. <http://dx.doi.org/10.1016/j.ajhg.2018.10.016>. PMID:30471717.
- Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, Burley SK. Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature*. 2000;403(6772):916-21. <http://dx.doi.org/10.1038/35002634>. PMID:10706293.

- Gao Y, Wu F, Ren Y, Zhou Z, Chen N, Huang Y, Lei C, Chen H, Dang R. MiRNAs expression profiling of bovine (*Bos taurus*) testes and effect of bta-miR-146b on proliferation and apoptosis in bovine male germline stem cells. *Int J Mol Sci*. 2020;21(11):3846. <http://dx.doi.org/10.3390/ijms21113846>. PMID:32481702.
- Girardet L, Augière C, Asselin MP, Belleannée C. Primary cilia: biosensors of the male reproductive tract. *Andrology*. 2019;7(5):588-602. <http://dx.doi.org/10.1111/andr.12650>. PMID:31131532.
- Guo B, Xiao C, Liu Y, Zhang N, Bai H, Yang T, Xiang Y, Nan Y, Li Q, Zhang W, Huang D. MiR-744-5p inhibits multiple myeloma proliferation, epithelial mesenchymal transformation and glycolysis by targeting SOX12/Wnt/ β -catenin signaling. *OncoTargets Ther*. 2021;14:1161-72. <http://dx.doi.org/10.2147/OTT.S270636>. PMID:33654408.
- Guo D, Ru J, Xie L, Wu M, Su Y, Zhu S, Xu S, Zou B, Wei Y, Liu X, Liu Y, Liu C. Tmem138 is localized to the connecting cilium essential for rhodopsin localization and outer segment biogenesis. *Proc Natl Acad Sci USA*. 2022;119(15):e2109934119. <http://dx.doi.org/10.1073/pnas.2109934119>. PMID:35394880.
- Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, Das PP, Tarakhovskiy A, Miska EA, Surani MA. MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. *PLoS One*. 2008;3(3):e1738. <http://dx.doi.org/10.1371/journal.pone.0001738>. PMID:18320056.
- Horvath GC, Kistler WS, Kistler MK. RFX2 is a potential transcriptional regulatory factor for histone H1t and other genes expressed during the meiotic phase of spermatogenesis. *Biol Reprod*. 2004;71(5):1551-9. <http://dx.doi.org/10.1095/biolreprod.104.032268>. PMID:15229132.
- Hu T, Shen H, Li J, Yang P, Gu Q, Fu Z. RFC2, a direct target of miR-744, modulates the cell cycle and promotes the proliferation of CRC cells. *J Cell Physiol*. 2020;235(11):8319-33. <http://dx.doi.org/10.1002/jcp.29676>. PMID:32239691.
- Huang W, Chen Q, Dai J, Zhang Y, Yi Y, Wei X, Wu Z. miR-744-5p suppresses tumor proliferation and metastasis by targeting transforming growth factor-beta 1 (TGF- β 1) in hepatocellular carcinoma (HCC). *J Gastrointest Oncol*. 2021;12(4):1811-22. <http://dx.doi.org/10.21037/jgo-21-319>. PMID:34532130.
- Huo JL, Wang P, Miao YW, Huo HL, Liu H, Zeng YZ, Xiao H. Isolation, sequence identification and tissue expression profile of a novel ribokinase gene (RBKS) from Chinese Banna mini-pig inbred line (BMI). *Afr J Biotechnol*. 2012;11(1):46-53. <http://dx.doi.org/10.5897/AJB11.2885>.
- Huo JL, Zhang LQ, Zhang X, Wu XW, Ye XH, Sun YH, Cheng WM, Yang K, Pan WR, Zeng YZ. Genome-wide single nucleotide polymorphism array and whole-genome sequencing reveal the inbreeding progression of Banna minipig inbred line. *Anim Genet*. 2022;53(1):146-51. <http://dx.doi.org/10.1111/age.13149>. PMID:34658041.
- Karnitz LM, Zou L. Molecular pathways: targeting ATR in cancer therapy. *Clin Cancer Res*. 2015;21(21):4780-5. <http://dx.doi.org/10.1158/1078-0432.CCR-15-0479>. PMID:26362996.
- Katan Y, Agami R, Shaul Y. The transcriptional activation and repression domains of RFX1, a context-dependent regulator, can mutually neutralize their activities. *Nucleic Acids Res*. 1997;25(18):3621-8. <http://dx.doi.org/10.1093/nar/25.18.3621>. PMID:9278482.
- Kistler WS, Baas D, Lemeille S, Paschaki M, Seguin-Estevez Q, Barras E, Ma W, Duteyrat J-L, Morlé L, Durand B, Reith W. RFX2 is a major transcriptional regulator of spermiogenesis. *PLoS Genet*. 2015;11(7):e1005368. <http://dx.doi.org/10.1371/journal.pgen.1005368>. PMID:26162102.
- Koay TW, Osterhof C, Orlando IMC, Keppner A, Andre D, Yousefian S, Suárez Alonso M, Correia M, Markworth R, Schödel J, Hankeln T, Hoogewijs D. Androglobin gene expression patterns and FOXJ1-dependent regulation indicate its functional association with ciliogenesis. *J Biol Chem*. 2021;296:100291. <http://dx.doi.org/10.1016/j.jbc.2021.100291>. PMID:33453283.
- Li J, Ran C, Li E, Gordon F, Comstock G, Siddiqui H, Cleghorn W, Chen H-Z, Kornacker K, Liu C-G, Pandit SK, Khanizadeh M, Weinstein M, Leone G, de Bruin A. Synergistic function of E2F7 and E2F8 is essential for cell survival and embryonic development. *Dev Cell*. 2008;14(1):62-75. <http://dx.doi.org/10.1016/j.devcel.2007.10.017>. PMID:18194653.
- Mitxelena J, Apraiz A, Vallejo-Rodríguez J, Malumbres M, Zubiaga AM. E2F7 regulates transcription and maturation of multiple microRNAs to restrain cell proliferation. *Nucleic Acids Res*. 2016;44(12):5557-70. <http://dx.doi.org/10.1093/nar/gkw146>. PMID:26961310.
- National Center for Biotechnology Information – NCBI [Internet]. 2022 [cited 2022 Sept 27]. Available from: https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/003/025/GCF_000003025.6_Sscrofa11.1/GCF_000003025.6_Sscrofa11.1_rna_from_genomic.fna.gz

- Nekrep N, Jabrane-Ferrat N, Wolf HM, Eibl MM, Geyer M, Peterlin BM. Mutation in a winged-helix DNA-binding motif causes atypical bare lymphocyte syndrome. *Nat Immunol.* 2002;3(11):1075-81. <http://dx.doi.org/10.1038/ni840>. PMID:12368908.
- Pan J, Goodheart M, Chuma S, Nakatsuji N, Page DC, Wang PJ. RNF17, a component of the mammalian germ cell nuage, is essential for spermiogenesis. *Development.* 2005;132(18):4029-39. <http://dx.doi.org/10.1242/dev.02003>. PMID:16093322.
- Quigley IK, Kintner C. Rfx2 stabilizes Foxj1 binding at chromatin loops to enable multiciliated cell gene expression. *PLoS Genet.* 2017;13(1):e1006538. <http://dx.doi.org/10.1371/journal.pgen.1006538>. PMID:28103240.
- Satir P, Christensen ST. Overview of structure and function of mammalian cilia. *Annu Rev Physiol.* 2007;69(1):377-400. <http://dx.doi.org/10.1146/annurev.physiol.69.040705.141236>. PMID:17009929.
- Shawlot W, Vazquez-Chantada M, Wallingford JB, Finnell RH. Rfx2 is required for spermatogenesis in the mouse. *Genesis.* 2015;53(9):604-11. <http://dx.doi.org/10.1002/dvg.22880>. PMID:26248850.
- Shima JE, McLean DJ, McCarrey JR, Griswold MD. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod.* 2004;71(1):319-30. <http://dx.doi.org/10.1095/biolreprod.103.026880>. PMID:15028632.
- Singh G, Storey KB. MicroRNA cues from nature: a roadmap to decipher and combat challenges in human health and disease? *Cells.* 2021;10(12):3374. <http://dx.doi.org/10.3390/cells10123374>. PMID:34943882.
- Tan YL, Bai ZG, Zou WL, Ma XM, Wang TT, Guo W, Liu J, Li JS, Jie-Yin, Zang YJ, Zhang ZT. miR-744 is a potential prognostic marker in patients with hepatocellular carcinoma. *Clin Res Hepatol Gastroenterol.* 2015;39(3):359-65. <http://dx.doi.org/10.1016/j.clinre.2014.09.010>. PMID:25543521.
- Tang D, Sha Y, Gao Y, Zhang J, Cheng H, Zhang J, Ni X, Wang C, Xu C, Geng H, He X, Cao Y. Novel variants in DNAH9 lead to nonsyndromic severe asthenozoospermia. *Reprod Biol Endocrinol.* 2021;19(1):27. <http://dx.doi.org/10.1186/s12958-021-00709-0>. PMID:33610189.
- Wang P, Zhang X, Huo H, Wang S, Song X, Huo J. Septin14, a gene specifically expressed in the testis and seminal vesicle of the Banna mini-pig inbred line (BMI). *Anim Reprod.* 2020;17(4):e20200521. <http://dx.doi.org/10.1590/1984-3143-ar2020-0521>. PMID:33791027.
- Wolfe SA, van Wert J, Grimes SR. Transcription factor RFX2 is abundant in rat testis and enriched in nuclei of primary spermatocytes where it appears to be required for transcription of the testis-specific histone H1t gene. *J Cell Biochem.* 2006;99(3):735-46. <http://dx.doi.org/10.1002/jcb.20959>. PMID:16676351.
- Wolfe SA, Vanwert JM, Grimes SR. Transcription factor RFX4 binding to the testis-specific histone H1t promoter in spermatocytes may be important for regulation of H1t gene transcription during spermatogenesis. *J Cell Biochem.* 2008;105(1):61-9. <http://dx.doi.org/10.1002/jcb.21793>. PMID:18459115.
- Wu Y, Hu X, Li Z, Wang M, Li S, Wang X, Lin X, Liao S, Zhang Z, Feng X, Wang S, Cui X, Wang Y, Gao F, Hess RA, Han C. Transcription factor RFX2 is a key regulator of mouse spermiogenesis. *Sci Rep.* 2016;6(1):20435. <http://dx.doi.org/10.1038/srep20435>. PMID:26853561.
- Xie M-Y, Chen T, Xi Q-Y, Hou L-J, Luo J-Y, Zeng B, Li M, Sun J-J, Zhang Y-L. Porcine milk exosome miRNAs protect intestinal epithelial cells against deoxynivalenol-induced damage. *Biochem Pharmacol.* 2020;175:113898. <http://dx.doi.org/10.1016/j.bcp.2020.113898>. PMID:32145262.
- Yu CH, Xie T, Zhang RP, A ZC. Association of the common SNPs in RNF212, STAG3 and RFX2 gene with male infertility with azoospermia in Chinese population. *Eur J Obstet Gynecol Reprod Biol.* 2018;221:109-12. <http://dx.doi.org/10.1016/j.ejogrb.2017.12.030>. PMID:29277047.
- Zhang M, Li H, Zhang Y, Li H. Oncogenic miR-744 promotes prostate cancer growth through direct targeting of LKB1. *Oncol Lett.* 2019;17(2):2257-65. PMID:30675291.
- Zhou W, Li Y, Gou S, Xiong J, Wu H, Wang C, Yan H, Liu T. MiR-744 increases tumorigenicity of pancreatic cancer by activating Wnt/ β -catenin pathway. *Oncotarget.* 2015;6(35):37557-69. <http://dx.doi.org/10.18632/oncotarget.5317>. PMID:26485754.

Author contributions

ZL: Formal analysis, Writing – original draft; HD: Writing – original draft; HH: Methodology, Data curation; WL: Data curation; YJ: Review & editing; XZ: Designing the article, Writing – review & editing; JH: Funding acquisition, Writing – review & editing.