

RESEARCH ARTICLE

Transcription factor-like 5 is a potential DNA- and RNA-binding protein essential for maintaining male fertility in mice

Weiya Xu¹, Yiyun Zhang¹, Dongdong Qin¹, Yiqian Gui², Shu Wang¹, Guihua Du¹, Fan Yang¹, Lufan Li¹, Shuiqiao Yuan², Mei Wang^{1,3,*} and Xin Wu^{1,*}

ABSTRACT

Transcription factor-like 5 (TCFL5) is a testis-specific protein that contains the basic helix-loop-helix domain, but the *in vivo* functions of TCFL5 remain unknown. Herein, we generated CRISPR/Cas9-mediated knockout mice to dissect the function of TCFL5 in mouse testes. Surprisingly, we found that it was difficult to generate homozygous mice with the *Tcfl5* deletion as the heterozygous males (*Tcfl5*^{+/-}) were infertile. However, we did observe markedly abnormal phenotypes of spermatids and spermatozoa in the testes and epididymides of *Tcfl5*^{+/-} mice. Mechanistically, we demonstrated that TCFL5 transcriptionally and post-transcriptionally regulated a set of genes participating in male germ cell development via TCFL5 ChIP-DNA and eCLIP-RNA high-throughput sequencing. We also identified a known RNA-binding protein, FXR1, as an interacting partner of TCFL5 that may coordinate the transition and localization of TCFL5 in the nucleus. Collectively, we herein report for the first time that *Tcfl5* is haploinsufficient *in vivo* and acts as a dual-function protein that mediates DNA and RNA to regulate spermatogenesis.

This article has an associated First Person interview with the first author of the paper.

KEY WORDS: DNA-binding protein, RNA-binding protein, Male fertility, Oligoasthenoteratozoospermia, Spermatogenesis, Spermiogenesis

INTRODUCTION

Infertility disorders affect ~15% of couples, and ~50% are caused by male factors (Schlegel, 2009), with abnormal spermatogenesis the principal cause of male infertility (Silber, 1994). Spermatogenesis in mammals includes three primary stages: mitosis of spermatogonia, meiosis of spermatocytes and spermiogenesis. Spermiogenesis is the last step of spermatogenesis in which the haploid round spermatids differentiate into mature spermatozoa amid a series of dramatic cellular reorganizations and morphologic changes, including acrosomal formation, nuclear elongation and condensation, flagellar development and cytoplasmic elimination (Eddy, 2002). Precise genetic regulation during spermatogenesis is required for


sperm production to take place in concert with functional integrity. Novel regulatory elements are therefore worthy of exploring as spermatogenesis represents one of the most complicated biologic processes *in vivo*.

Over the past several decades, key regulators – in particular the germline-specific factors in the process of spermatogenesis – have been reported to act via transcriptional and post-transcriptional regulation (Bettegowda and Wilkinson, 2010; Legrand and Hobbs, 2018). For example, testis-specific transcription factor CREM τ is a well-studied key transcription activator that controls postmeiotic germ cell differentiation. CREM τ is expressed in postmeiotic spermatids and directly targets a list of postmeiotic genes that encode structural proteins required for spermatid differentiation (Kosir et al., 2012; Martianov et al., 2010; Nantel et al., 1996). *Cremt*-deficient mice exhibit complete arrest of spermiogenesis at step 5 during the 16-step formation process from round spermatids to elongated spermatids (Nantel et al., 1996). SOX family transcription factor SOX30 is a transcription factor revealed to control the gene expression transition from late meiotic to postmeiotic stages, as well as the subsequent development of round spermatids (Bai et al., 2018), and SOX30 mutant males are sterile owing to spermiogenic arrest at the early round-spermatid stage (Bai et al., 2018). In addition, post-transcriptional regulation of gene expression also plays critical roles in spermatogenesis. RANBP9 (Ran-binding protein 9) has been identified as a regulator of alternative splicing in spermatocytes and spermatids (Bao et al., 2014). In *Ranbp9* knockout testes, expression of key spermiogenic genes is dysregulated and unique mRNA transcript isoforms are increased as a result of aberrant alternative splicing, leading to germ cell loss and infertility (Bao et al., 2014). TPAP [testis-specific poly(A) polymerase] has been demonstrated to maintain mRNA transcript stability (Kashiwabara et al., 2000) and is essential for the progression of haploid spermatid differentiation (Kashiwabara et al., 2002), and *Tpap* null mice were infertile as a result of failed spermatid morphogenesis (Kashiwabara et al., 2002). SAM68 (Src-associated substrate in mitosis of 68 kDa) is important for alternative splicing regulation and translational control in male meiotic germ cells, with loss of *Sam68* resulting in aberrant differentiation of round spermatids, increased apoptosis and consequent infertility (Paronetto et al., 2009). These post-transcriptionally regulated RNA-processing events, such as pre-mRNA splicing, mRNA export, maintenance of transcript stability and translation in the male germline, are required to maintain coordinated gene expression (Legrand and Hobbs, 2018).

The basic helix-loop-helix (bHLH) family of transcription factors is receiving increased attention due to its critical roles during the development and differentiation of a wide variety of cell types (Davis et al., 1987; Massari and Murre, 2000; Murre et al., 1989). The best understood bHLH proteins are the myogenic regulatory factors (MRFs), which are capable of converting the mesodermal cell line C3H10T1/2 into muscle progenitor cells known as

¹State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing, Jiangsu 210029, China. ²Institute of Reproductive Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China. ³Centre for Reproductive Medicine, Lianyungang Maternal and Child Health Hospital, Lianyungang, Jiangsu 222000, China.

*Authors for correspondence (xinwu@njmu.edu.cn; wangmei@njmu.edu.cn)

 S.Y., 0000-0003-1460-7682; M.W., 0000-0002-1571-2970; X.W., 0000-0001-7938-9407

Handling Editor: Caroline Hill

Received 17 June 2021; Accepted 14 December 2021

myoblasts (Olson, 1990). During spermatogenesis in the testes, studies have also revealed that multiple bHLH family members occupy essential functional roles in the development of germ cells. SPZ1 (a bHLH-Zip transcriptional factor) is expressed in the testes and epididymides of adult male mice, and overexpression induces apoptosis of germ cells at an early stage of spermatogenesis, reduces the populations of normal spermatozoa, creates smaller litter sizes and eventuates in infertility at 6 months of age (Hsu et al., 2004). *Sohlh1* and *Sohlh2* (spermatogenesis- and oogenesis-specific bHLH 1 and 2) are expressed in KIT^- and KIT^+ spermatogonia, and are both required for spermatogonial differentiation through the regulation of the expression of *Gfra1*, *Sox3* and *Kit* genes (Ballow et al., 2006a,b; Hao et al., 2008; Suzuki et al., 2012; Toyoda et al., 2009). SOHLH1 and SOHLH2 heterodimerize and homodimerize with each other *in vivo*, and double mutants of *Sohlh1/Sohlh2* phenocopy the single mutant of *Sohlh1* or *Sohlh2*, inducing improper spermatogonial differentiation and prematurely activating meiosis (Suzuki et al., 2012).

Transcription factor-like 5 (TCFL5) was previously considered to be a testis-specific protein, and the encoded domains and localization in the spermatocyte subsets of the testes are conserved between humans and mice (Maruyama et al., 1998; Siep et al., 2004); however, the *in vivo* function and mechanism(s) underlying TCFL5 action have not yet been elucidated. In the present study, we characterized the expression of TCFL5 in mouse testis, demonstrated that the loss of TCFL5 led to mouse infertility and showed that TCFL5 potentially acts as a DNA- and RNA-binding protein to regulate germ cell development at the transcriptional and post-transcriptional levels.

RESULTS

Localization of TCFL5 in mouse testes

As there are contradictory reports of differences between the expression and localization of TCFL5 RNA and protein (Maruyama et al., 1998; Shi et al., 2013; Siep et al., 2004), we analyzed the spatiotemporal expression of TCFL5 in mouse testes and showed that TCFL5 protein was expressed exclusively in the mouse testes, with two bands at 50 kDa and 55–70 kDa. In both bands, the presence of amino acid sequences of TCFL5 were verified through protein mass spectrometry (Fig. 1A; Fig. S1A,B, Tables S1,S2). The expression of TCFL5 in mouse testes was also dramatically elevated, starting from postnatal day (P)21 and continuing into adulthood (Fig. 1B). Consistent with this developmental pattern, we also uncovered a preferential location for TCFL5 protein in pachytene spermatocytes and round spermatids – subpopulations of germ cells that we purified through the STA-PUT velocity sedimentation method we published previously (Wang et al., 2019) (Fig. 1C). Next, we immunostained for TCFL5 and determined that the TCFL5 signals were located in middle- and late-pachytene spermatocytes at stages V and VI in the seminiferous tubules of mouse testes; however, TCFL5 was absent in early pachytene spermatocytes at stages I–IV (Fig. 1D; IgG control staining shown in Fig. S1C). In general, the localization of TCFL5 in mouse testes was consistent with other previous studies (Maruyama et al., 1998; Siep et al., 2004).

Notably, TCFL5 signals were present in the nucleus and cytoplasm of spermatocytes at stage V–XII and round spermatids at stage I in the mouse seminiferous tubules (verified by the immunofluorescence staining and isolation experiments of nuclear and cytoplasmic fractions using mouse testis cells; Fig. 1D; Fig. S1D). The signals of TCFL5 were higher in the nucleus of pachytene spermatocytes at stages IX and X, and diplotene spermatocytes at stage XI, whereas they showed the reversed intensity in nuclear and cytoplasmic regions

of secondary spermatocytes in meiosis II and early round spermatids at stage I (Fig. 1D). This distinct expression pattern suggested that the actions of TCFL5 may be related to *in vivo* function during the transition from late meiotic to postmeiotic stages in mouse spermatogenesis (Fig. 1E).

Infertility in *Tcfl5*-heterozygous mice

To elucidate the *in vivo* function of TCFL5, we generated *Tcfl5*-knockout mice using CRISPR/Cas9 technology-mediated gene targeting (Fig. S2A). We designed small guide RNAs (gRNAs) to target exons 1–4 and deleted these four exons (and the introns between them, Fig. S2A,B). The first generation (F1) of *Tcfl5* mutants was successfully sired from four founder lines with either indels or deletions of *Tcfl5* alleles (Fig. S2A–C). Testicular sizes and weight were identical in *Tcfl5*^{+/+} and *Tcfl5*^{-/-} allelic mice (Fig. 2A; Fig. S2D), and western blotting and immunohistochemical staining revealed a diminution in TCFL5 protein levels in *Tcfl5*^{+/-} mouse testes compared to *Tcfl5*^{+/+} mice (Fig. 2B,C). However, we were unable to generate mice with *Tcfl5*^{-/-} alleles from any F1 heterozygous fathers derived from the four founders according to the over-1-year breeding experiments, as all male *Tcfl5* mice with heterozygous alleles were likely to be infertile (Fig. 2D). Next, our histologic analysis revealed extensive defects in testicular seminiferous tubules and epididymal tubules in *Tcfl5*^{+/-} mice compared with *Tcfl5*^{+/+} mice; i.e. the numbers of elongating spermatids were significantly decreased in *Tcfl5*^{+/-} seminiferous tubules and epididymal tubules (Fig. S2E; Fig. 3G), multinucleated cells appeared around the lumen of these seminiferous tubules and we observed many degenerating cells with round-shaped nuclei in epididymal tubules, as indicated by the black arrowheads (Fig. S2E).

Impaired spermiogenesis in testes of *Tcfl5*^{+/-} mice

To investigate the spermatogenic defects in the seminiferous tubules of *Tcfl5*^{+/-} mice, we conducted periodic acid-Schiff (PAS) and Hematoxylin staining (Fig. 2E), and double immunostaining with peanut agglutinin (PNA, a sperm acrosome marker) and DAPI (a nuclear indicator, Fig. 3A), to reveal changes in spermatogenesis and defects in spermiogenesis. In contradistinction to the seminiferous epithelium of *Tcfl5*^{+/+} mice, in which the arrangements of spermatogonia, spermatocytes, round spermatids and elongating spermatid layers were normal from the base to the lumen (Fig. 3A), *Tcfl5*^{+/-} mouse testes exhibited extensive defects during the 12 stages encompassing spermatogenic epithelium. For example, we observed that the deformation of round spermatids into elongating spermatids was apparently impaired at stages IX and X (blue arrowhead, Fig. 2E); that the nuclei of steps 9 and 10 elongating spermatids were irregular in size and shape; that the subsequent elongating spermatids at steps 11 and 12 presented with morphologic abnormalities (Fig. 3A); and that elongating spermatids at step 16, which would normally be released into the lumen, still remained in the epithelium of stages VII and VIII, and stages IX and X (black arrows, Fig. 2E). Compared with *Tcfl5*^{+/+} mice, elongating spermatids at steps 9 and 10 of *Tcfl5*^{+/-} mice manifested malformed acrosomes. For example, the nucleus was overextended and its structural integrity (such as hooked head and dorsal angle) was lost; the acrosome did not extend normally upon losing its dorsal and ventral fins, and therefore all elongating spermatids in the following steps (11–16) were globally malformed (as illustrated in Fig. 3B); and elongating spermatids that should not be present at steps 9 and 10 were still found at stage I of the seminiferous epithelium (Fig. 3A). By contrast, all types of germ cells prior to step-9 elongating spermatids in *Tcfl5*^{+/-} mouse testes

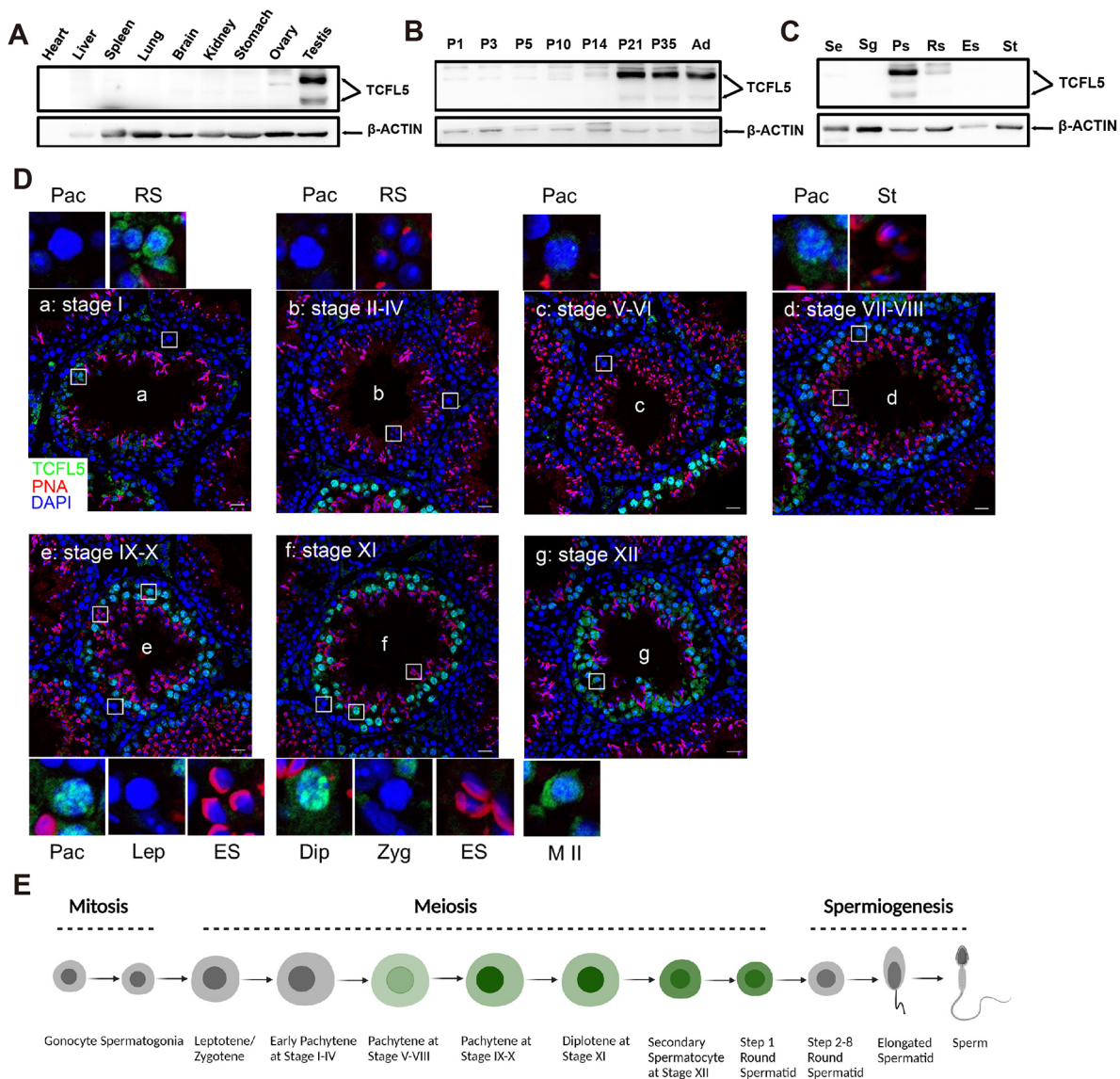


Fig. 1. Spatiotemporal expression of TCFL5 in mouse tissue. (A–C) Western blotting analysis of TCFL5 protein (50 kDa and 55–70 kDa) in lysates from adult mouse tissues (A), Testicular tissues collected at different timepoints during postnatal development (B), and isolated spermatogenic cell populations, including Sertoli cells (Se), spermatogonia (Sg), pachytene spermatocytes (Ps), round spermatids (Rs), elongated spermatids (Es) and spermatozoa (St) (C). β -ACTIN (42 kDa) served as an internal loading control. (D) Immunostaining of testicular sections from adult wild-type mice for TCFL5 (green) and PNA (red); PNA is an acrosomal marker and DNA was stained with DAPI. Boxes indicate the areas featured in the magnified images. (E) Schematic diagram demonstrating the spatiotemporal expression of TCFL5 in germ cells. Dip, diplotene spermatocytes; Es, elongated spermatids; Lep, leptotene spermatocytes; M II, second meiosis; Pac, pachytene spermatocytes; Rs, round spermatids; St, spermatozoa; Zyg, zygotene spermatocytes. Scale bars: 20 μ m.

were extant and appeared normal in the epithelium (Fig. 2E, Fig. 3A).

We then evaluated the sperm parameters and found that the proportion of head malformations of mature sperm in *Tcfl5*^{+/-} mouse epididymides was higher (79.5 \pm 4.2%) compared to the proportion in *Tcfl5*^{+/+} mice (10.9 \pm 7.9%) (mean \pm s.d.; P <0.001, unpaired t -test, Fig. 3C,E). To further examine the defects in *Tcfl5*^{+/-} mouse sperm, we performed transmission electron microscopy (TEM) to visualize sperm ultrastructure (Fig. 3D). In contrast to the *Tcfl5*^{+/+} mouse sperm that exhibited long oval shapes with their acrosomes attached to the nuclei in the integrated '9+2' arrangement of microtubules within their tails (Fig. 3Da–Dc), the nuclei of *Tcfl5*^{+/-} sperm were severely deformed and the acrosomes were detached from the nuclei (Fig. 3Dd,De), and some outer doublet microtubules were missing (indicated by the red arrows in Fig. 3Df). The percentage of

microtubules missing in *Tcfl5*^{+/-} mouse sperm tails was significantly higher relative to *Tcfl5*^{+/+} mice (P <0.01, unpaired t -test, Fig. 3F). In addition, the percentages of motile and progressively motile sperm from *Tcfl5*^{+/-} mice were significantly lower than in *Tcfl5*^{+/+} mice (P <0.001, unpaired t -test, Fig. 3H,I). Overall, we concluded that there was a gene haploinsufficiency in *Tcfl5*^{+/-} mice that prevented spermiogenesis at stages IX and X during the deformation of round spermatids to elongating spermatids.

TCFL5 activates the expression of genes functionally related to mouse spermatogenesis

To investigate the molecular events underlying the spermatogenic defects in *Tcfl5*^{+/-} mouse sperm, we performed high-throughput RNA-seq on purified pachytene spermatocytes in which TCFL5 protein appeared to be preferentially localized. Using adult

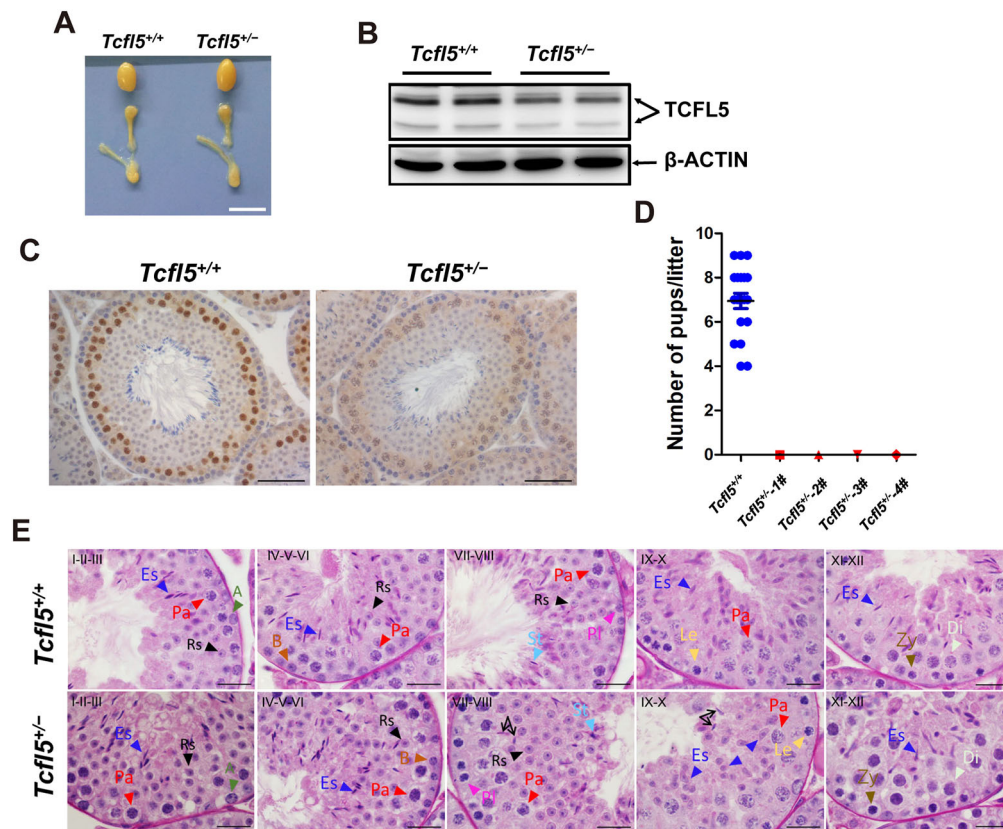


Fig. 2. Generation of *Tcf15* mutant mice by CRISPR/Cas9. (A) General morphology of testes and epididymides in adult *Tcf15*^{+/+} and *Tcf15*^{-/-} mice. (B,C) Western blotting analysis and immunohistochemical staining confirmed the decrease in TCFL5 protein (50 kDa and 55–70 kDa) in adult *Tcf15*^{-/-} mouse testes. β -ACTIN (42 kDa) served as the loading control. (D) Breeding experiments between *Tcf15*^{+/+} and *Tcf15*^{-/-} males. A mouse from each genotype was coupled with mates of normal fertility. The results were determined using nine cages of *Tcf15*^{+/+} male mice from all four lines (see Fig. S1A) [data are mean \pm s.d.; dots represent 20 offspring litters derived from *Tcf15*^{+/+} (none for *Tcf15*^{-/-} mice)]. (E) Morphology within the 12 stages governing the growth and development of spermatogenic epithelium. The 12 stages of seminiferous epithelium (top left) were identified according to PAS staining and cell arrangement patterns. A, type A spermatogonia; B, type B spermatogonia; Di, diplotene spermatocytes; Es, elongated spermatids; Le, leptotene spermatocytes; Pa, pachytene spermatocytes; PI, preleptotene spermatocytes; Rs, round spermatids; St, spermatozoa; Zy, zygotene spermatocytes. Scale bars: 50 μ m.

wild-type and *Tcf15*^{-/-} mouse testes, we collected pachytene spermatocytes purified to over 80% (Fig. S3A,B). The comparison of mRNA expression profiles revealed that 260 and 458 genes were significantly upregulated and downregulated, respectively (Fig. 4A) in *Tcf15*^{-/-} versus wild-type pachytene spermatocytes, as shown in the scatterplot (Fig. 4B; $P < 0.05$, fold-change > 2 ; see detailed data in Table S3 and in the GEO database, accession number GSE176240). Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) protein network analysis of downregulated genes revealed that these transcripts were associated with cell development, cell differentiation, intercellular adhesion and spermiogenesis (Fig. S3G). Our heatmap and qPCR depicts the downregulated genes associated with spermiogenesis and spermatogenesis (Fig. 4C; Fig. S3H; $P < 0.05$, fold-change > 2). The attenuated expression of the genes associated with intercellular adhesion may partially explain the inability of sperm at stages VII–X to be released from the epithelium in a timely manner.

To identify genes directly regulated by TCFL5, we next performed chromatin immunoprecipitation with massively parallel DNA sequencing (ChIP-seq) to identify the binding of TCFL5 by using 21-day testes (in which TCFL5 was abundantly expressed). The enrichment of TCFL5 protein was verified by western immunoblotting (Fig. S3C). Model-based analysis of ChIP-seq (Zang et al., 2009) identified 12,623 peak-associated genes with significant enrichment by TCFL5 (Table S4). We then analyzed the

ChIP-seq data by mapping DNA sequences to genomic regions and found that TCFL5 bound to the promoter regions of genes as expected (Fig. 4D), and that a majority of TCFL5-binding sites were close to the center of the transcription start sites (TSS, Fig. 4E). The significant enrichment of downregulated genes compared to upregulated genes suggested broader transcriptional regulation functions by TCFL5 in the mouse testes (58% versus 35%; Fig. 4F; Fig. S3D–F).

Next, we analyzed RNA-seq data in combination with the ChIP-seq data and identified 268 downregulated genes, including spermatogenesis-associated genes (as listed in Fig. 4F). A genome browser view revealed that TCFL5 bound to the promoter regions of *Tnfrsf2*, *Aurkb*, *Nme8* and *Scml2*, and that it activated their expression (Fig. 4G). We noted that *Scml2* was one of the genes that showed a pronounced reduction in transcription (Table S3; Fig. 4F; and qPCR data in Fig. S3H). To evaluate the regulation of *Scml2* by TCFL5, we constructed a TCFL5 expression vector by inserting TCFL5 CDS into a pcDNA3.1+ vector, and determined the peak sequence in which TCFL5 bound to the *Scml2* promoter region using a dual-luciferase assay by co-transfecting 293T cells with a TCFL5 expression plasmid, a pGL3 basic plasmid-containing promoter peak sequence and a *Renilla* luciferase vector. Our results established that relative luciferase activity increased significantly in the experimental group compared to the control groups (Fig. S3I). In addition, western blotting analysis indicated an attenuated

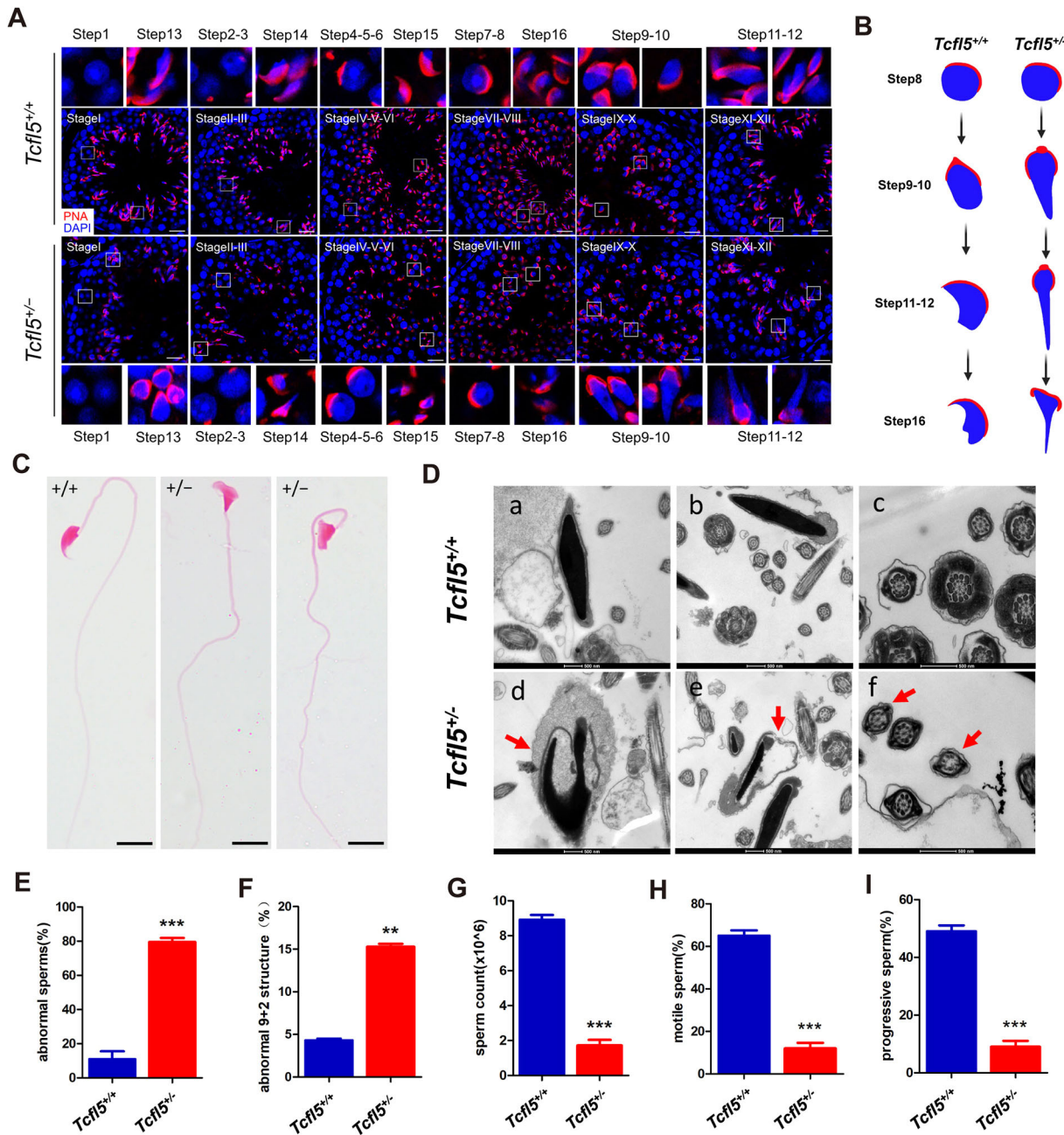


Fig. 3. Defects in spermiogenesis between *Tcf15^{+/+}* and *Tcf15^{+/-}* mice. (A) PNA (red)- and DAPI (blue)-stained testis sections from adult *Tcf15^{+/+}* and *Tcf15^{+/-}* mice. The seminiferous tubule stage was identified based on the acrosomal morphology marked by PNA and arrangement of spermatogenic cells, and the 16-step spermiogenesis process of round and elongated spermatids was identified by the PNA signal pattern. (B) Schematic diagram showing the key steps of the differences in sperm deformation between *Tcf15^{+/+}* and *Tcf15^{+/-}* mice. (C) Sperm from *Tcf15^{+/+}* and *Tcf15^{+/-}* mouse epididymides stained by eosin showing abnormal morphologies. (D) Electron microscopic analysis of epididymal sperm from *Tcf15^{+/+}* and *Tcf15^{+/-}* mice, with red arrows indicating the head and tail of the sperm of the *Tcf15^{+/-}* mice. (E) Sperm with abnormal morphologies. $n=3$. (F) Percentages of sperm with an abnormal '9+2' structure were also determined. $n=2$. (G-I) Comparison of sperm count (G), motile sperm (H) and progressive sperm (I) in *Tcf15^{+/+}* and *Tcf15^{+/-}* mouse epididymides by CASA. $n=3$ per group. Data are mean \pm s.d. ** $P<0.01$, *** $P<0.001$ (unpaired two-tailed Student's *t*-test). Scale bars: 20 μ m (A); 20 μ m (C), 500 nm (D).

expression of SCML2 protein in *Tcf15^{+/-}* mouse testes (Fig. S3J). These results indicated that TCFL5 activated the expression of *Scml2* and other genes that are related to spermatogenesis by binding to their promoters (Fig. S5).

TCFL5 interacts with the RNA-binding proteins FXR1 and DHX9

To further scrutinize the regulatory mechanism(s) subserving TCFL5 function in spermatogenesis, we profiled TCFL5-interacting proteins

with TCFL5 antibody immunoprecipitation using lysates of mouse testes at P21. In two independent immunoprecipitation-mass spectrometric experiments, we identified 57 candidate proteins that may interact with TCFL5 (Fig. S4A, Tables S5, S6). Intriguingly, we found that a majority of the interacting proteins of TCFL5 were functionally related to RNA and post-transcriptional processing, including heterogeneous nuclear ribonucleoproteins (HnRNPs), splicing factor-related proteins, ribosomal proteins and some classical well-studied RNA-binding proteins (i.e. FXR1 and

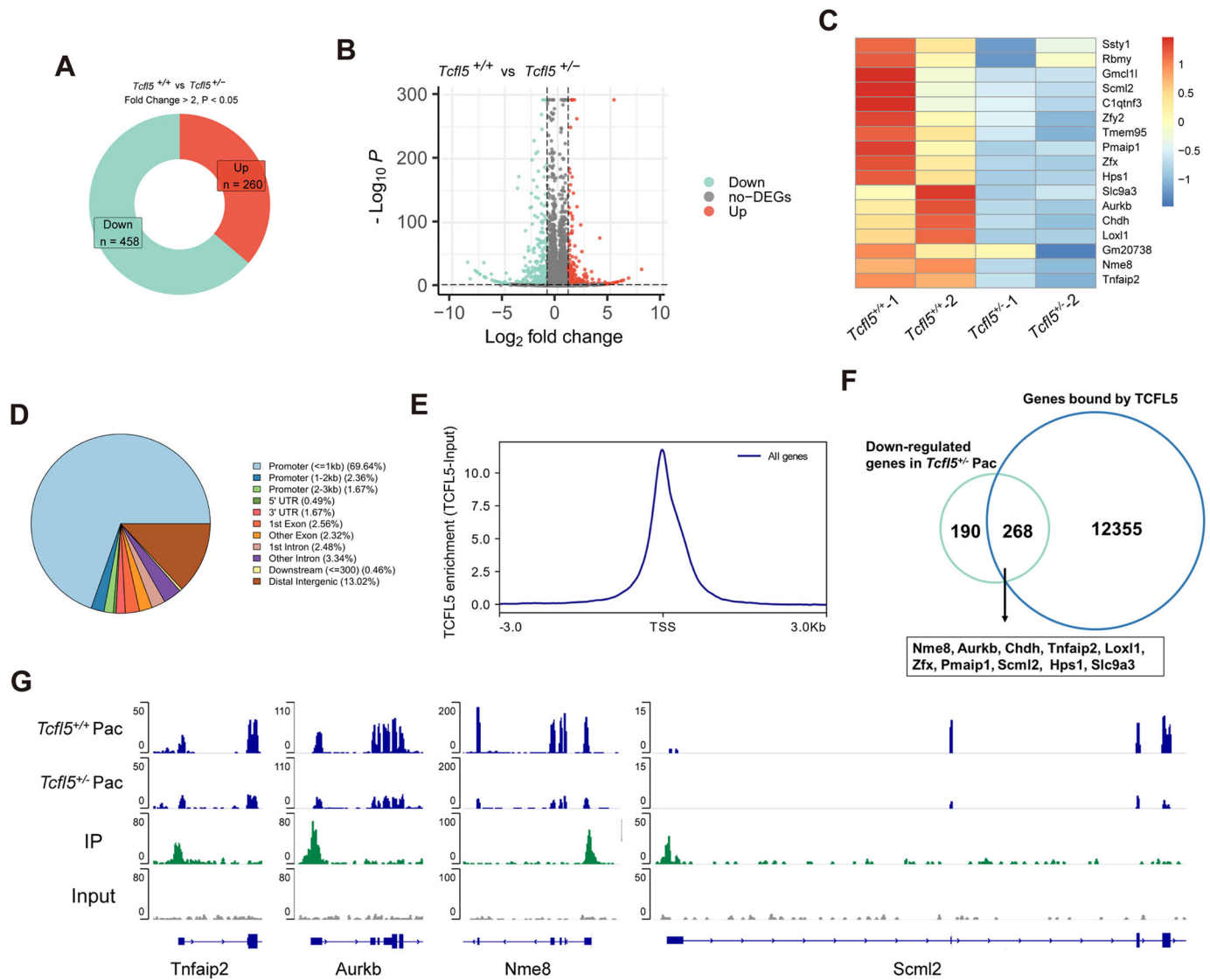


Fig. 4. Transcriptional activation by TCFL5 of gene expression in mouse spermatogenesis. (A) Genes upregulated or downregulated using high-throughput RNA-seq between *Tcfl5*^{+/+} and *Tcfl5*^{+/-} mouse testes. Three independent samples were collected from each genotype and genes manifesting at least a twofold change and $P < 0.05$ were selected for comparison. (B) Scatterplot of differentially expressed transcripts between *Tcfl5*^{+/-} and *Tcfl5*^{+/+} pachytene spermatocytes isolated from adult *Tcfl5*^{+/+} and *Tcfl5*^{+/-} mice with the STA-PUT method (*Tcfl5*^{+/+}, $n=2$; *Tcfl5*^{+/-}, $n=2$). Each dot represents a gene that was significantly changed (fold change > 2, $P < 0.05$). (C) Heatmap of genes that are differentially upregulated and downregulated by TCFL5, and that are closely related to spermatogenesis. DEGs, differentially expressed genes. (D) Distribution map of the proportions of potential binding positions of genes found in the TCFL5 ChIP experiments. UTR, untranslated region. (E) Prediction of the distribution area of TCFL5-binding sites at or near the targeting gene TSS. (F) ChIP and RNA-seq experiments suggested that 268 genes (58.5%) of the 458 downregulated genes were directly regulated by the transcriptional regulation of TCFL5. (G) Genome browser view of TCFL5 ChIP-seq and RNA-seq reads on *Tnfaip2*, *Aurkb*, *Nme8* and *Scml2* gene loci. IP, immunoprecipitation.

DHX9), as shown by Coomassie Blue staining prior to sequencing by mass spectrometry (Fig. 5A). Gene ontology (GO) analysis displayed terms significantly related to mRNA processing, RNA splicing and/or mRNA stability (Fig. 5B). These results suggested that different regions of TCFL5 may be also responsible for binding to DNA and RNA with or without other proteins, and that TCFL5 may possess dual DNA- and RNA-binding capacities. Of 57 candidate proteins that interacted with TCFL5, we further confirmed at least two known dual DNA- and RNA-binding proteins – FXR1 (fragile-X mental retardation syndrome-related protein 1) and DHX9 (DEXH-box helicase 9) – by western blotting after TCFL5 immunoprecipitation, indicating that the interactions between/among FXR1, DHX9 and TCFL5 may coordinate transcriptional regulation in mouse testes (Fig. 5C). Immunofluorescence staining showed that TCFL5

colocalized with FXR1 and DHX9 in mouse testes (Fig. 5D). We examined the co-expression of TCFL5-FLAG and FXR1-EGFP proteins (Fig. 5E, lower panel) in 293T cells to compare with TCFL5-FLAG and EGFP proteins (Fig. 5E, upper panel), and showed that FXR1 likely supported the transport of TCFL5 into the nucleus (Fig. 5E). Collectively, the aforementioned data indicated that TCFL5 may not only be involved in transcriptional regulation but also in RNA processing during spermatogenesis.

TCFL5 was an RNA player in regulating mouse spermatogenesis

Next, we asked whether TCFL5 binding to DNA could be stabilized by its binding to RNA; if so, then RNase treatment of chromatin should reduce TCFL5 occupancy. To test this idea, we treated

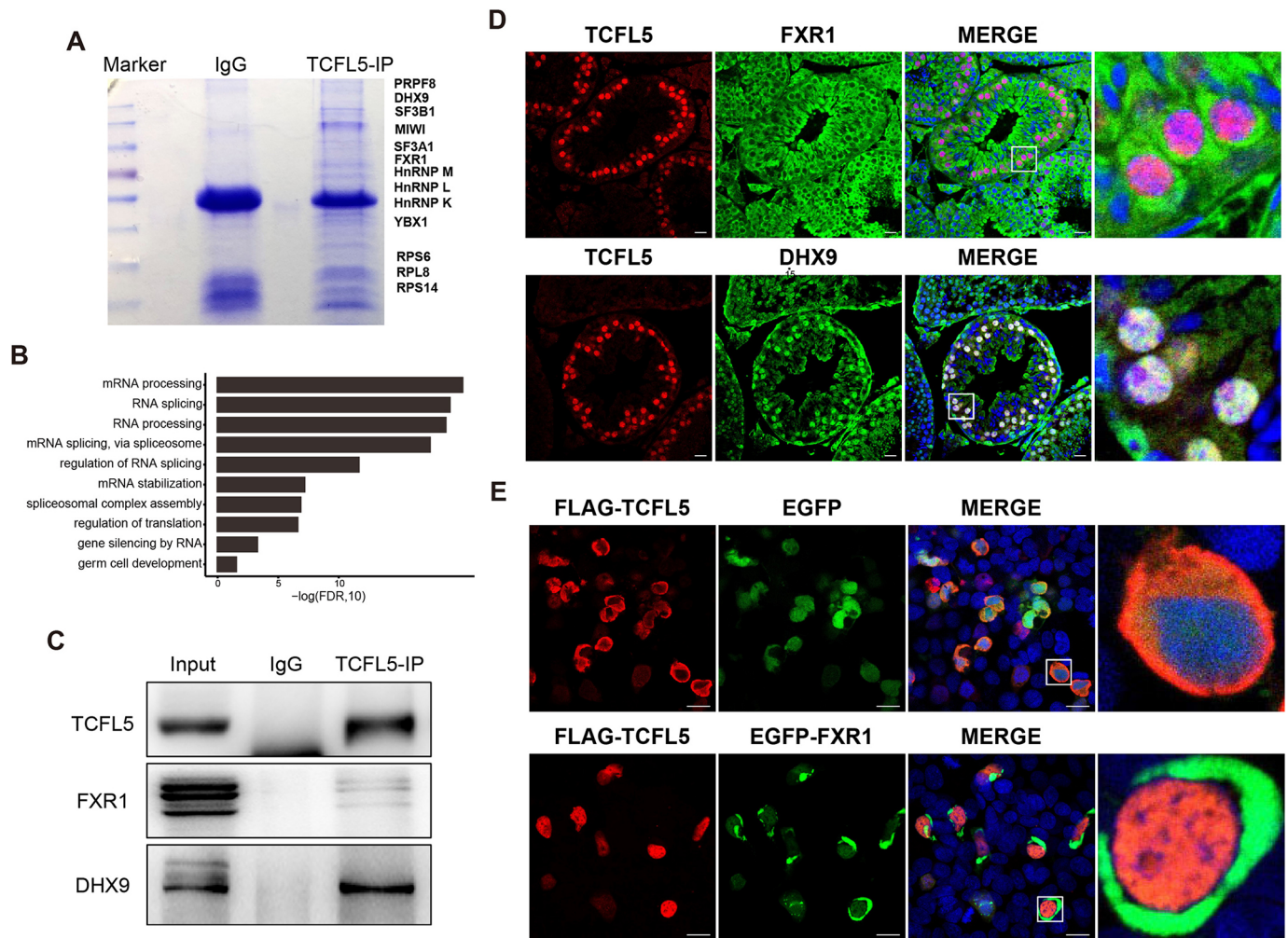


Fig. 5. Interaction of TCFL5 with FXR1 and DHX9. (A) Coomassie Blue staining after TCFL5 immunoprecipitation (IP) in 21-day mouse testes indicating the differentially enriched protein positions as revealed by protein mass spectrometry. (B) GO term-enrichment analysis of 57 candidates (see Fig. S4) obtained from two independent TCFL5 immunoprecipitation and mass-spectrometric experiments. (C) Immunoprecipitation followed by western blotting in 21-day mouse testes confirming the interaction of TCFL5 (55–70 kDa) with FXR1 (70–80 kDa) and DHX9 (140 kDa). (D) Colocalization of TCFL5 with FXR1 and DHX9 in the seminiferous tubules of mouse testes. (E) Comparison of the co-expression of TCFL5-FLAG protein with EGFP protein (upper panel) and TCFL5-FLAG protein with FXR1-EGFP protein (lower panel) in 293T cells after *in vitro* transfection. The boxed areas in D and E indicate the magnified areas in the images to the right. Scale bars: 20 μm .

nuclear chromatin with RNase inhibitor or RNase A, and, indeed, the levels of TCFL5 bound to chromatin were significantly diminished when the chromatin preparation was treated with RNase A, consistent with the concept that RNA may contribute to the stability of TCFL5 in chromatin (Fig. 6A).

In order to explore the possibility of TCFL5 as an RNA player during mouse germ cell development and verify the direct RNA-binding function and post-transcriptional regulation of TCFL5, we performed the high-throughput sequencing of RNAs isolated by enhanced crosslinking immunoprecipitation (eCLIP) experiments. TCFL5 antibody was used to pulldown RNAs bound by TCFL5 from 21-day-old mouse testes, and the enrichment of TCFL5 protein was confirmed by western blotting (Fig. S4B). The RNAs obtained through two independent pulldowns were used to generate two cDNA libraries (Fig. S4C). The binding peaks with at least eightfold enrichment when compared with the background input were considered as significant. TCFL5 bound to protein-coding RNA and a small amount of long non-coding (lnc)RNA, such as antisense RNA and lincRNA (intergenic lincRNA), as shown in Fig. S4D. The 4614 overlapping mRNA targets were identified by two

independent experiments (Fig. 6B). We then mapped RNA sequences to genomic regions and found that the overwhelming majority of TCFL5-binding sites were in intron regions (accounting for 96.09%, Fig. 6C,D). GO processing analysis revealed that these mRNA targets were significantly correlated with cilium assembly, flagellated sperm motility, spermatogenesis, cell division and differentiation, as well as regulation of translation, gene silencing by RNA, piwi-interacting (pi)RNA metabolic process and meiotic cell cycle (Fig. 6E). Of these, TCFL5 bound to multiple intron regions of spermiogenesis-related genes, e.g. *Rfx2*, *Rnf17*, *Crem* and *Boll* (Fig. 6F). Together, these results indicate that TCFL5 regulates spermatogenesis at the post-transcriptional level by binding to the intron regions of spermatogenesis genes, and may be involved in the alternative splicing for mRNA precursors.

DISCUSSION

Owing to the scarcity of available clinical cases and detailed genetic background data – and because clinical characterization of information regarding male infertility is often lacking – animal models have been applied to examine the roles of genes and to

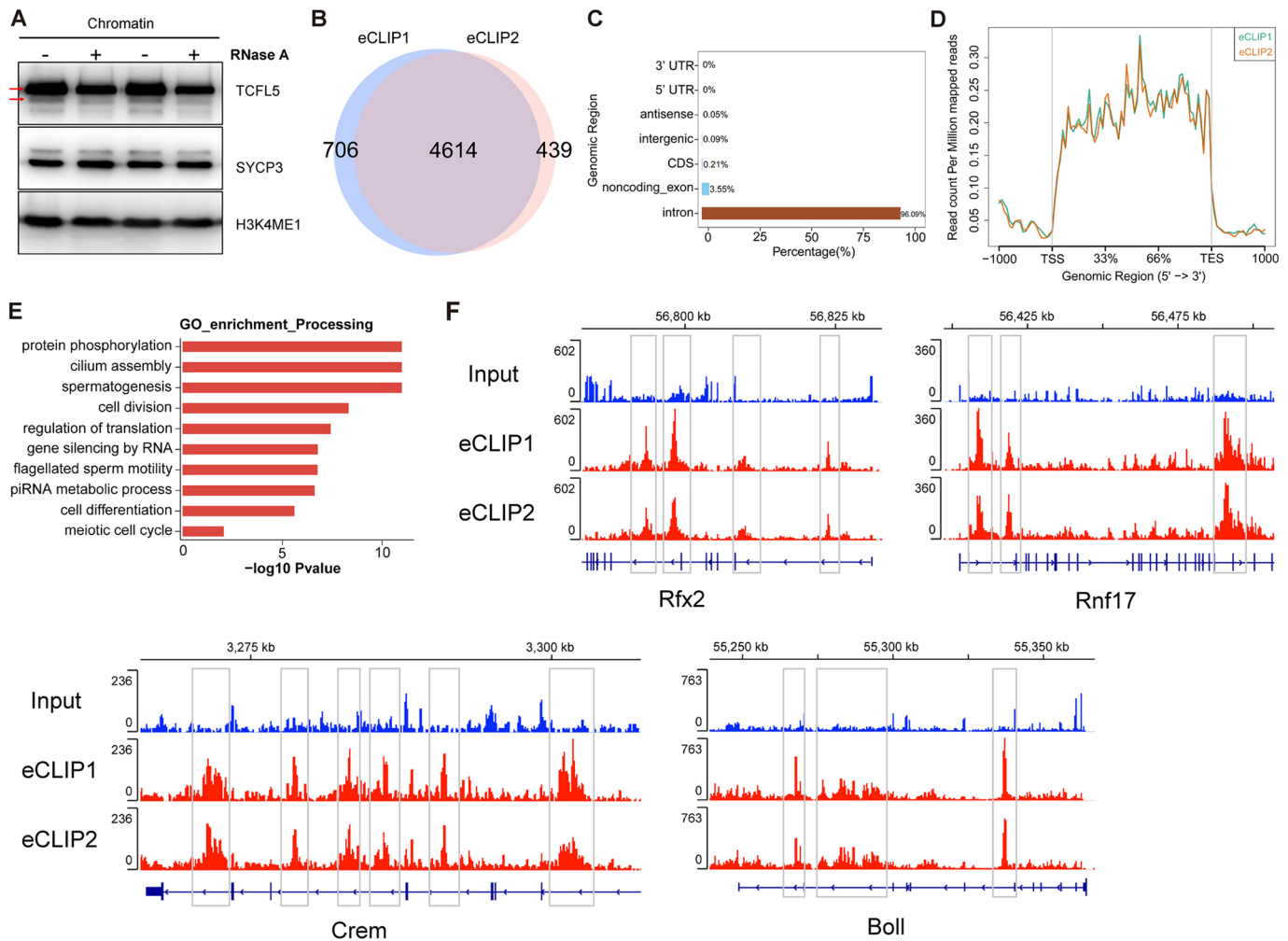


Fig. 6. TCFL5 regulates spermatogenesis at the post-transcriptional level. (A) Chromatin-binding assay as assessed by western blotting analysis showing the levels of TCFL5 (red arrows represent 50 kDa and 55–70 kDa), SYCP3 (28 kDa) and H3K4ME1 (15 kDa) in a nuclear chromatin preparation after RNase A inhibitor or RNase A treatment. (B) Venn diagram analysis of genes identified by eCLIP. Data were from two independent experiments. (C) Distribution of overlapping peaks in two experiments based on their genomic locations. UTR, untranslated region. (D) Plot showing the distribution of TCFL5 eCLIP reads on gene bodies. TCFL5 was enriched at the regions from TSS to transcription end sites (TES). (E) GO term-enrichment analysis of 4614 genes obtained from the two independent TCFL5 eCLIP experiments. (F) Genome browser view of TCFL5 eCLIP-seq on *Rfx2*, *Rnf17*, *Crem* and *Boll* gene loci.

clarify their underlying mechanism(s) of action in the processes of germ cell production (Matzuk and Lamb, 2008; Nalam and Matzuk, 2010). In the present study, we identified a new testis-specific gene, TCFL5, that belongs to the bHLH family, and is required for spermatogenesis in mouse testes and for male fertility. TCFL5 is localized in middle and late spermatocytes starting from stage V to the round spermatid at stage I, and demonstrates both nuclear and cytoplasmic expression. We observed the signal intensity of TCFL5 to be weak in stage V–VIII spermatocytes, but augmented from stage IX onward, with signal intensity in the nucleus stronger than in the cytoplasm. By contrast, intensity in these compartments was reversed in secondary spermatocytes and round spermatids. These distinct expression patterns support the concept that TCFL5 functions from late meiotic to postmeiotic stages.

From our RNA-seq and ChIP-seq results, it was obvious that TCFL5 bound to at least 258 genes (58.5% of the downregulated genes by RNA-Seq), validating the transcriptional regulation by TCFL5 of genes associated with germ cell development and serving as a critical factor in the mouse testis. *Scml2* in the testis has been reported to deubiquitylate RNF8-dependent monoubiquitylation of histone H2A at lysine 119 (H2AK119ub), thus inhibiting

the establishment of H3K27 acetylation (Adams et al., 2018; Hasegawa et al., 2015). H3K27 acetylation is a marker of active enhancers and is followed by H3K4 dimethylation, a modification associated with active transcription (Creighton et al., 2010; Talasz et al., 2005). We showed that TCFL5 bound to the promoter region of *Scml2* and regulated the transcription of *Scml2*. Therefore, the decrease in *Scml2* expression in *Tcfl5* mutants may have affected the reactivation of male reproductive genes after meiosis, and ultimately led to the production of abnormal sperm. *Tnfrsf2* and *Nme8* are components of the acrosome and fibrous sheath, respectively (Miranda-Vizuete et al., 2003; Smith et al., 2013; Wolf et al., 1994). *Nme8* mutation is associated with accelerated loss of sperm motility and impaired protamination of sperm chromatin (Smith et al., 2013). *Aurkb* mutant mice exhibited reduced sperm counts, and in some cases complete absence of spermatozoa, manifesting multinucleated cells in the seminiferous tubules and subfertility defects (Kimmins et al., 2007). AURKB is a chromosomal passenger protein that regulates kinetochore-microtubular attachments and cytokinesis, preventing tetraploidization (Adams et al., 2001; Steigemann et al., 2009), and this may explain the phenotype of multinucleated cells present in *Tcfl5*^{+/-} mouse testes. TCFL5 may also regulate

spermatogenesis by activating *Tnfrsf2*, *Nme8* and *Aurkb* expression. A previous study (Siep et al., 2004) showed that the *calmegin* gene might be the target of TCFL5, indicated by gel-shift and yeast one-hybrid experiments; however, TCFL5 as a candidate *calmegin* promoter-binding protein was not experimentally validated in mouse germ cell or *in vivo*. According to our ChIP experiment and high-throughput sequencing data, we did not find the binding relationship between TCFL5 protein and the *calmegin* gene.

Proteins that bind DNA or RNA are often considered and studied independently of one another. However, recent studies have indicated that ~8% of nuclear proteins possess dual DNA- and RNA-binding capacities, and many transcriptional factors are capable of binding diverse types of RNA (Hudson and Ortlund, 2014). For example, although the DNA-binding protein CTCF was initially described as a conserved transcription factor (Filippova et al., 1996), it was subsequently also reported to be a specific RNA-binding protein with a previously unrecognized RNA-binding region (Saldaña-Meyer et al., 2014) and shown to be involved in alternative splicing (Kung et al., 2015; Shukla et al., 2011). The classical transcription factor YY1 also exhibits RNA-binding function (Jeon and Lee, 2011; Sigova et al., 2015). The binding of YY1 to RNA transcripts associated with regulatory elements can enhance its occupation at these elements (Jeon and Lee, 2011; Sigova et al., 2015), and the chromatin-associated protein PARP1 associates with nucleosomes that harbor exonic H3K4me3 marks, binds nascent RNA and regulates splicing co-transcriptionally by recruiting splicing factors (Matveeva et al., 2016; Melikishvili et al., 2017). Transcription factor Sox2 can associate with RNA while binding to its cognate DNA sequence to form ternary RNA–Sox2–DNA complexes that co-transcriptionally regulate RNA metabolism during somatic cell reprogramming (Hou et al., 2020). In addition, some RNA-binding proteins localize to active chromatin regions and directly participate in transcriptional control, with active gene promoters constituting the major binding regions (Xiao et al., 2019). SR proteins comprise a family of RNA-binding proteins involved in both constitutive and regulated splicing (Lin and Fu, 2007), and SR protein SRSF2 has been reported to be a transcriptional regulator associating with gene promoters as part of the 7SK complex and activating transcription via promoter-proximal nascent RNA (Ji et al., 2013). We discovered that TCFL5 interacts with many RNA-binding proteins according to our mass spectrometric results, and therefore we speculate that TCFL5 exerts RNA-binding actions and is involved in post-transcriptional regulation.

Of the plethora of interacting proteins, the first group was composed of hnRNPs, i.e. HnRNP M, HnRNP L and HnRNP K. HnRNPs represent a large family of RNA-binding proteins that contribute to multiple aspects of nucleic acid metabolism, including alternative splicing, mRNA stabilization and transcriptional and translational regulation (Geuens et al., 2016). The second group comprised splicing factors, i.e. SF3B1, SF3A1 and PRPF8, and the third consisted of ribosomal proteins (e.g. RPS6, RPL8 and RPS14). In addition, there were also some classical RNA-binding proteins, such as DHX9 and FXR1. DHX9 is a member of the DEX/H-box family of helicases and was initially portrayed as a transcriptional co-activator acting as a bridging factor between the transcriptional co-factor CREB-binding protein (CBP)/p300 and Pol-II, recruiting the latter to the CREB/CBP/p300 complexes at the promoter (Nakajima et al., 1997). In addition to transcriptional regulation, DHX9 is necessary for post-transcriptional regulation, such as pre-mRNA splicing (Hartmuth et al., 2002; Zhang et al., 1999). FXR1 was also reported to regulate transcription by recruiting transcription

factor STAT1 or STAT3 to gene promoters at the chromatin interface, and at least partially mediating cellular proliferation (Fan et al., 2017). Notably, FXR1 has both nuclear localization and nuclear export functions (Ceman et al., 1999). For example, leptomycin B was applied to inhibit the export receptor exportin 1, resulting in FXR proteins accumulating in the nucleus (Tamanini et al., 1999). Therefore, it is possible that FXR1 helps TCFL5 enter the nucleus but brings itself out of the nucleus.

More meaningful are that the eCLIP-seq results showing TCFL5 binds to at least 4616 mRNAs. These mRNAs are largely associated to a series of spermatogenesis-associated genes. GO processing analysis showed that a large portion of genes are involved in spermiogenesis and flagellum assembly, as well as piRNA metabolism/meiosis (Fig. S5). We assume that TCFL5 may play a role in sperm development and differentiation by regulating the splicing of these spermiogenesis genes and/or flagellum assembly genes, as TCFL5 mainly targets the intron regions of these mRNAs. For example, RFX2 transcription factor is a master regulator of genes required for the haploid phase, and *Rfx2* mutant mice display sterility and complete arrest before spermatid elongation in the testes (Kistler et al., 2015). RNF17 forms distinct nuages in germ cells, and *Rnf17* mutant male mice were sterile, with complete arrest in round spermatids (Pan et al., 2005). Transcription activator Crem controls a series of postmeiotic cell-specific gene expression, and the mutant mice exhibit arrest at the first step of spermiogenesis (Nantel et al., 1996). Boll is one of the oldest members of the DAZ family of genes, and the mutant mice displayed arrest at step 6 round spermatids (VanGompel and Xu, 2010). Additionally, as suggested by our eCLIP-seq data, TCFL5 may be also involved in piRNA metabolism, and this downstream mechanism may rely on TCFL5 co-participating with other RNA players, such as DHX9 (Fig. S5), as shown by immunoprecipitation-mass spectrometry data. Our chromatin-binding experiment also showed that after the removal of RNA by RNase A, the chromatin binding by TCFL5 protein was reduced, indicating that TCFL5 binds DNA in an RNA-dependent manner, and confirming the overall RNA-binding function of TCFL5.

Collectively, we postulate that TCFL5 plays an important role in spermiogenesis via two pathways: one in the regulation of *Scml2*, *Tnfrsf2*, *Aurkb* and *Nme8* gene transcription, and the other in the regulation of *Rfx2*, *Rnf17*, *Crem* and *Boll* mRNA maturation, by interacting with RNA-binding proteins, such as DHX9 and FXR1 (Fig. S5). Because of the important functions ascribed to TCFL5, we posit that *Tcfl5* haploinsufficiency leads to dysfunction in spermatogenesis. These data thus provide novel clues toward understanding spermatogenesis and the treatment of male infertility.

MATERIALS AND METHODS

Generation of *Tcfl5* knockout mice by CRISPR/Cas9

Exons 1–4 of the *Tcfl5* gene were targeted with four single guide (sg)RNAs, and the synthesized paired oligonucleotides for sgRNAs were annealed and cloned into the pUC57-sgRNA expression vector (51132, Addgene). Oligo sequences were as follows: sgRNA1 up, 5'-taggTGGCCGGCGCGT-CGTCTG-3'; sgRNA1 down, 5'-aacCAGACGACGCGCCGCCA-3'; sgRNA2 up, 5'-taggCGGGCGCGACGGGCGCG-3'; sgRNA2 down, 5'-aacCGCCGCCGTCGCGCCCG-3'; sgRNA3 up, 5'-taggCCTT-CAGGTTAGTTGGCA-3'; sgRNA3 down, 5'-aacTGCCAACCTAACCT-GAAGG-3'; sgRNA4 up, 5'-taggTAGCTTCGTATTGTTTTG-3'; and sgRNA4 down, 5'-aacCAAACAATACGAAGCTA-3'. The Cas9 expression plasmid (44758, Addgene) was linearized with AgeI and used as the template for *in vitro* transcription using a T7 Ultra Kit (AM1345, Ambion). sgRNA expression plasmids were linearized with DraI and used as templates for *in vitro* transcription using a MEGAshortscript Kit (AM1345, Ambion). Transcribed Cas9 mRNA and sgRNA were both purified using a

MEGAclear Kit (AM1908, Ambion). C57BL/6 mice were purchased from Beijing Vital River Laboratories animal center and housed in standard cages, and maintained on a 12-h light/dark cycle with food and water *ad libitum*. The microinjection of fertilized mouse eggs was performed in the Zhang Laboratory, Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences. In brief, fertilized eggs were injected with a mixture of Cas9 mRNA (12 ng/ μ l) and four sgRNAs (6 ng/ μ l each). Founder mice were backcrossed to C57BL/6J. The PCR primers for genotyping were as follows: forward primer (P1), 5'-GCTTCATCAGACGCTTGGTC-3'; reverse primer (P2), 5'-TGCATGATGTGTGGAGCTGACAGC-3'; forward primer (P3), 5'-CAGTGGGCGTGTCTTTATGAG-3'; and reverse primer (P4), 5'-CTT-GCTACTTAAGACTGGC-3'. Primers P1 and P2 were used for wild type and primers P3 and P4 for the mutant. The wild-type allele generated a 407-bp band, and the knockout allele produced a 5160-bp band. Mice were maintained under specific pathogen-free conditions according to the guidelines of the Institutional Animal Care and Use Committee of Nanjing Medical University.

Isolation of pachytene spermatocytes

We isolated pachytene spermatocytes from adult mice using the STA-PUT method. Mice testes were first digested with collagenase IV (1 mg/ml), and the dispersed seminiferous tubules were washed with Dulbecco's modified Eagle medium and then centrifuged at 500 g. The pellet was further digested with 0.25% trypsin-containing DNase I (1 mg/ml), and we terminated digestion with 10% fetal bovine serum. The resulting cellular suspension was filtered with a 40- μ m filter to prepare a single-cell suspension, which was then loaded onto cell-separation apparatus (ProScience, Canada) using a 2–4% bovine serum albumin (BSA) gradient (2% and 4% BSA were loaded onto the separation-apparatus chamber). After 2–3 h of sedimentation, we harvested the cell fractions and identified pachytene spermatocytes according to their morphologic characteristics, cell diameters and Hoechst staining under light microscopy (Ti2-U, Nikon, Japan).

Western immunoblotting analysis

Testes were lysed in RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 1% NP40, 1 mM DTT, 0.5% sodium deoxycholate, 0.05% SDS and 1 mM EDTA) containing protease-inhibitor cocktail (11697498001, Roche). Protein lysates were separated on 10% SDS-PAGE gels followed by electrotransfer onto PVDF membranes. The blots were blocked for 2 h with 5% nonfat milk and then incubated overnight at 4°C with the following primary antibodies: anti-TCFL5 (1:1000; ab188075, Abcam); anti-SCML2 (1:1000; kindly provided by Dr Mengcheng Luo, Wuhan University, China; Luo et al., 2015); anti-FLAG a (1:1000; F1804, Sigma-Aldrich); anti-FXR1 (1:1000; A8697, Abclonal); anti-DHX9 (1:1000; A17955, Abclonal); anti-SYCP3 (1:500; ab97672, Abcam); anti-H3K4ME1 (1:500; ab8895, Abcam); anti-GAPDH (1:2000; GB11002, Servicebio); anti-VASA (1:500; ab13840, Abcam); and anti-H3K9ME3 (1:1000; A15624, Abclonal). Horseradish peroxidase-conjugated secondary antibody (1:2000; ZB2301, Zhongshan Jinqiao Biotechnology) and enhanced chemiluminescence reagent were used to visualize immunopositive bands.

Immunofluorescence

Testes were fixed in 4% paraformaldehyde at 4°C, dehydrated in sucrose, embedded in optimal cutting temperature compound and then sectioned (5 μ m) using a cryostat microtome (Cryotome FSE, Thermo Fisher Scientific). We executed immunostaining on testicular cryosections with primary antibody (anti-TCFL5), and Cy2-conjugated secondary antibody (1:200; 711-225-152, Jackson ImmunoResearch, USA) was used for visualization of immunostaining. For acrosomal staining, cryosections were incubated for 1 h at 37°C with rhodamine-conjugated PNA (1:2000; RL-1072, Vector Laboratories), and nuclear DNA was stained with DAPI (1:500; D9542, Sigma-Aldrich). Stained testis sections were visualized using a confocal microscope (LSM800, Zeiss).

Immunocytochemistry

The Fxr1 CDS was cloned into a pEGFP-C1 vector and co-expressed with EGFP TAG (the co-expression plasmid was constructed by Tsingke

Biological Technology, Nanjing, China). The Tcf15/pcDNA3.1+ and Fxr1/pEGFP-C1 plasmids were transfected into HEK 293T cells (SCSP-502, National Collection of Authenticated Cell Cultures, China) simultaneously using calcium phosphate/CaCl₂ transfection. After transfection for 36 h, the cells were fixed with 4% paraformaldehyde, permeabilized with 0.5% Triton X-100 and blocked with 2% BSA/10% normal donkey serum/0.2% Triton X-100 in PBS for 2 h at room temperature. The cells were then incubated overnight at 4°C with FLAG antibody (1:500; F1804, Sigma-Aldrich), and incubated with tetramethylrhodamine-conjugated secondary antibody (1:200, 715-025-150, Jackson ImmunoResearch, USA) for 2 h at room temperature. Images were taken using a confocal microscope (LSM800, Zeiss).

Histology and transmission electron microscopic analysis

Testes and epididymides were fixed in Hartman's fixative (H0290, Sigma-Aldrich) for 24 h. The tissues were dehydrated with increasing concentrations of ethanol (70%, 80%, 90% and 100%), cleared in xylene, embedded in paraffin and sectioned (5 μ m). The sections were deparaffinized, rehydrated, stained with Hematoxylin and Eosin (Sigma-Aldrich) or stained with PAS reagent. For TEM analysis, cauda epididymal sperm obtained from *Tcf15*^{+/+} and *Tcf15*^{+/-} mice were fixed in 2.5% glutaraldehyde in 0.2 M cacodylate buffer at 4°C overnight. Sperm were then washed with 0.2 M cacodylate buffer, dehydrated in a graded series of ethanols and embedded and polymerized in an automated microwave tissue processor (EMAMW, Leica). Ultrathin sections were stained with uranyl acetate and lead citrate, and were examined by TEM (JEM-1010, JEOL).

Sperm analysis

Sperm from *Tcf15*^{+/+} and *Tcf15*^{+/-} mice were extracted and incubated in Hanks' balanced salt solution medium (14175-079, Gibco) at 37°C for 20 min, and samples were diluted and analyzed using Hamilton Thorne's Ceros II system.

Co-immunoprecipitation

For immunoprecipitation, whole testes from mice at P21 were lysed in RIPA buffer containing protease-inhibitor cocktail. Protein lysates were then blocked by incubation with protein A Dynabeads (10002D, Invitrogen) for 2 h at 4°C, followed by incubation with primary antibody against TCFL5 (1:50; 188075, Abcam) or control IgG (Rabbit IgG, A7016, Beyotime) overnight at 4°C. The lysates were then incubated with protein A Dynabeads for 2 h at 4°C, and proteins were eluted in lysis buffer for 10 min at 100°C after stringent washing in lysis buffer. The co-immunoprecipitated protein complexes were detected using western blotting with the following antibodies: anti-TCFL5 (1:1000; ab188075, Abcam); anti-FXR1 (1:1000; A8697, Abclonal); and anti-DHX9 (1:1000; A17955, Abclonal).

Coomassie Blue staining

SDS-PAGE gels were dipped in R250 dye solution for 3 h at room temperature with gentle rocking and then washed in washing buffer until the gels became transparent.

Mass spectrometry

The TCFL5 protein complexes pulled down with an anti-TCFL5 antibody from adult mice testes were analyzed by the Mass Spectrometry Core Center, Nanjing Medical University, using the entire elutions.

RNA isolation and sequencing

Total RNA was extracted from pachytene cells taken from wild-type and *Tcf15*^{+/-} mice using TRIzol reagent (15596018, Life Technologies). Two biologic replicates were executed and used for sequencing. RNAs were subjected to oligo dT selection and adaptor ligation. We performed sequencing using the BGISEQ-500 platform at BGI Tech (Shenzhen, China). Low-quality reads were filtered using the internal software SOAPnuke. Afterward, clean reads were obtained and stored in FASTQ format. The clean reads were mapped to the mouse genome (GRCm38.p5) using HISAT2. Bowtie 2 was applied to align the clean reads to the reference coding gene set, and the expression level of each gene was calculated by

RSEM. Differentially expressed genes between samples were generated by DESeq2 algorithms. An expression fold change of ≤ -2 or ≥ 2 and adjusted q -value of < 0.05 were considered significantly different.

ChIP-seq and data analysis

Testis tissues were collected from wild-type and *Tcf15* mutants at P21, and crosslinked with 1% formaldehyde, followed by quenching with glycine solution. Chromatin fragmentation was performed by sonication in ChIP SDS lysis buffer [50 mM Tris-HCl (pH 8), 10 mM EDTA (pH 8) and 1% SDS] using a Covaris-S220 sonicator. Crosslinked chromatin was incubated with anti-Tcf15 antibodies in ChIP dilution buffer (16.7 mM Tris-HCl, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100 and 0.01% SDS) with protease inhibitors overnight. Crosslinking was then reversed overnight at 65°C, and DNA extracted using chloroform. Purified DNA was sequenced using the BGISEQ-500 platform. Raw reads were filtered first to remove low-quality or adaptor sequences using SOAPnuke, and cleaned reads were mapped to the mouse genome (GRCm38.p5) using SOAPaligner/SOAP2. Quality control for ChIP-seq data was performed with FastQC software, and peak calling for TCFL5 was carried out using Model-based Analysis for ChIP-Seq (MACS2). Genome-wide signal coverage tracks were visualized in the Integrative Genome Browser (IGV).

Dual-luciferase reporter assay

Peak sequences from TCFL5 targets were amplified by PCR, and the *Tcf15* CDS was cloned into a pcDNA3.1+ vector and co-expressed with FLAG TAG (the co-expression plasmid was constructed by Tsingke Biological Technology, Nanjing, China). The peak sequences of target mRNAs were cloned into a pGL3 basic vector (PM0874, BioNice Co., Ningbo) using a ClonExpress MultiS One Step Cloning Kit (C113-02, Vazyme Biotech). Primers synthesized for peak sequences were as follows: *Tcf15*-peak sequence, 5'-tctcgatcactaagtaagcttGTTATATGGTGGTGCTATCTTTGCTGG-3' (forward 1) and 5'-GTTTAATCACACCCTCATAAAGACACGC-3' (reverse 1), 5'-CTTTATGAGGGTGTGATTAACCTGGGCGTG-3' (forward 2) and 5'-CTGCGTGTACTCCACTTCCGTCATCTC-3' (reverse 2), and 5'-CGGAAGTGGAGTACACGCAGC-3' (forward 3) and 5'-agtaccggaatccaagcttGACTGCCTGCCCTTTACTG-3' (reverse 3); and *Scml2*-peak sequence, 5'-tctcgatcactaagtaagcttAAAAGTCCAAAGGCTTTTAGCATC-3' (forward) and 5'-agtaccggaatccaagcttCAGAGAACA-CTAATTTAGACACAGCAGC-3' (reverse). PCR conditions were as follows: 95°C for 5 min; 95°C for 45 s, 60°C for 30 s, and 72°C for 1 min for 35 cycles; 72°C for 10 min, and then the PCR was paused. The peak/pGL3 basic plasmid *Tcf15*/pcDNA3.1+ plasmid and *Renilla* plasmid (graciously provided by Dr Yunjun Xu, Nanjing Medical University, China) were transfected into HEK 293T cells simultaneously using CaCl₂ transfection. After transfection for 36 h, the cells were collected and luciferase activities were determined according to the manufacturer's instructions for the Dual-Luciferase Reporter Assay System (E1910, Promega).

eCLIP

TCFL5 eCLIP was performed according to the published protocol (Van Nostrand et al., 2016). Testes from 21-day mice were collected, detunicated, triturated in a grinder until the seminiferous tubules were evenly dispersed and immediately UV-irradiated three times at 254 nm (400 MJ/cm²). UV light-treated tissues were lysed in 1 ml iCLIP lysis buffer with protease inhibitor (04693132001, Roche) and RNase inhibitor (N251B, Promega); lysates were treated with RNase I (AM2295, LifeTech) for 5 min at 37°C, with shaking at 1200 rpm. The 10 µg of anti-TCFL5 antibody were used for each immunoprecipitation. RNA adapter X1A and X1B (Table S7) were ligated to RNA 3' ends dephosphorylated by FastAP Alkaline Phosphatase (EF0652, Thermo Fisher Scientific). Immunoprecipitation beads were eluted for 10 min at 70°C using washing buffer. Samples were analyzed by NuPAGE (4%–12% gradient precast gels run with MOPS buffer). Crosslinked RNA-protein complexes were transferred onto nitrocellulose membrane (1060008, Amersham Protran Premium GE). Then, the membrane was cut from the protein band to 75 kDa above it and RNA was purified using acid phenol/chloroform/isoamyl alcohol (25:24:1) (P1011, Solarbio). RNA adapter RiL19 (Table S7) was ligated to input

RNA 3' ends dephosphorylated by FastAP Alkaline Phosphatase. AR17 primer (Table S7) was used to reverse transcribe all RNA, and Rand103Tr3 adapter (Table S7) was ligated to the 5' ends of cDNA. The ligated cDNA was purified with MyONE Silane beads (37002D, Thermo Fisher Scientific), quantified by qPCR and amplified with Q5 PCR master mix (M0492L, New England Biolabs). The amplified cDNA was run in 3% low melting temp agarose gel (50180, Lonza), cut from 175 bp to 350 bp and extracted using a Qiagen gel extraction kit. The cDNA libraries were sequenced using an Illumina NovaSeq.

Sequenced reads were processed as described previously (Van Nostrand et al., 2016). The clean reads were evaluated using FastQC. Clean reads were aligned to the mouse GRCm38 genome (mm10) by HISAT2 (Kim et al., 2019). After reads were aligned onto the genome, only uniquely mapped reads were used for the following analysis. Peaks were called on immunoprecipitation samples using the CLIPper tool (Lovci et al., 2013) with options -s mm10. Identified peaks were normalized to the appropriate size-matched input control with 'Peak_input_normalization_wrapper.pl' (<https://github.com/YeoLab/gscripts>), with criteria $P \leq 10^{-5}$ and fold-enrichment ≥ 8 . Genomic annotations of peaks obtained by TCFL5 eCLIP experiments are shown in Table S8 and S9.

Chromatin-binding assay

The testicular tissues from two groups (100 mg/group) were prepared using 1 ml of hypotonic buffer each. One group was treated with a 1:100 dilution of RNase A (ST576, Beyotime) for 10 min at 37°C and the other was treated with 1 U/µl RNase inhibitor. For chromatin extraction, the tissue was homogenized in 1 ml of buffer A [250 mM sucrose, 10 mM Tris HCl (pH 8.0), 10 mM MgCl₂, 1 mM EGTA, 1× protease-inhibitor cocktail (11697498001, Roche)], and incubated for 30 min at 4°C with gentle rocking. The lysates were then passed through a 40-µm cell strainer followed by centrifugation at 1200 g for 5 min at 4°C to separate nuclei (pellet) from the cytoplasmic extract (supernatant). The pellet was washed in buffer A three times and then homogenized in 400 µl of buffer B [250 mM sucrose, 10 mM Tris HCl (pH 8.0), 10 mM MgCl₂, 1 mM EGTA, 1× protease-inhibitor cocktail (11697498001, Roche), 0.1% Triton X-100 and 0.25% NP-40] for 10 min at 4°C. The sample was then centrifuged at 100 g for 5 min at 4°C to remove nuclear membrane particles (pellet). The supernatant was collected and centrifuged at 1700 g for 5 min at 4°C to separate insoluble chromatin (pellet) from soluble nuclear protein (supernatant). The pellet was then washed in buffer B three times. We used 50 µl of 0.2 M HCl to release the bound protein from chromatin, and 10 µl of 1 M Tris HCl (pH 8.0) was used to neutralize the pH. The chromatin fraction was then analyzed by western blotting analysis to determine the expression of TCFL5.

Data processing and statistical analysis

We performed a computer-aided sperm analysis (CASA) experiment with three experimental replicates and captured fluorescence images using confocal microscopy (LSM800, Zeiss). All image data were assembled according to Adobe Illustrator (2021 edition, USA) software. Data are reported as mean±s.d. unless otherwise noted in the figure legends. Significance was assigned using an unpaired Student's *t*-test ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$) with Excel 2016 (Microsoft, USA). We carried out the quantitative data presentation using GraphPad Prism 8.0 software (USA).

Acknowledgements

We thank Dr Mengcheng Luo and Yuming Cao at Wuhan University in China for their kind help with the *Tcf15* mouse phenotype data.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: W.X., D.Q., G.D., X.W.; Methodology: W.X., D.Q., G.D.; Software: W.X., D.Q., G.D., L.L.; Validation: W.X., X.W.; Formal analysis: W.X., X.W.; Investigation: W.X., Y.Z., Y.G., F.Y., S.Y., X.W.; Resources: W.X., Y.Z., Y.G., S.W., F.Y., S.Y., M.W., X.W.; Data curation: W.X., Y.Z., D.Q., Y.G., S.W., M.W., X.W.; Writing - original draft: W.X., X.W.; Writing - review & editing: W.X., S.W., L.L., X.W.;

Visualization: W.X., Y.Z., F.Y., X.W.; Supervision: W.X., X.W.; Project administration: X.W.; Funding acquisition: X.W.

Funding

The work was supported by the National Key Research and Development Program of China (2021YFC2700201 and 2018YFC1003302 to X.W.), the National Natural Science Foundation of China (32070831 and 31872844 to X.W.) and the Natural Science Foundation of Jiangsu Province for Youth (BK20210140 to M.W.).

Data availability

Sequencing data have been deposited in GEO under accession numbers GSE176240 and GSE187385.

Peer review history

The peer review history is available online at <https://journals.biologists.com/jcs/article-lookup/doi/10.1242/jcs.259036>.

References

- Adams, R. R., Carmena, M. and Earnshaw, W. C. (2001). Chromosomal passengers and the (aurora) ABCs of mitosis. *Trends Cell Biol.* **11**, 49-54. doi:10.1016/S0962-8924(00)01880-8
- Adams, S. R., Maezawa, S., Alavattam, K. G., Abe, H., Sakashita, A., Shroder, M., Broering, T. J., Sroga Rios, J., Thomas, M. A., Lin, X. et al. (2018). RNF8 and SCML2 cooperate to regulate ubiquitination and H3K27 acetylation for escape gene activation on the sex chromosomes. *PLoS Genet.* **14**, e1007233. doi:10.1371/journal.pgen.1007233
- Bai, S., Fu, K., Yin, H., Cui, Y., Yue, Q., Li, W., Cheng, L., Tan, H., Liu, X., Guo, Y. et al. (2018). Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. *Development* **145**, dev164855. doi:10.1242/dev.164855
- Ballow, D., Meistrich, M. L., Matzuk, M. and Rajkovic, A. (2006a). Sohlh1 is essential for spermatogonial differentiation. *Dev. Biol.* **294**, 161-167. doi:10.1016/j.ydbio.2006.02.027
- Ballow, D. J., Xin, Y., Choi, Y., Pangas, S. A. and Rajkovic, A. (2006b). Sohlh2 is a germ cell-specific bHLH transcription factor. *Gene Expr. Patterns* **6**, 1014-1018. doi:10.1016/j.modgep.2006.04.007
- Bao, J., Tang, C., Li, J., Zhang, Y., Bhetwal, B. P., Zheng, H. and Yan, W. (2014). RAN-binding protein 9 is involved in alternative splicing and is critical for male germ cell development and male fertility. *PLoS Genet.* **10**, e1004825. doi:10.1371/journal.pgen.1004825
- Bettgowda, A. and Wilkinson, M. F. (2010). Transcription and post-transcriptional regulation of spermatogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 1637-1651. doi:10.1098/rstb.2009.0196
- Ceman, S., Brown, V. and Warren, S. T. (1999). Isolation of an FMRP-associated messenger ribonucleoprotein particle and identification of nucleolin and the fragile X-related proteins as components of the complex. *Mol. Cell. Biol.* **19**, 7925-7932. doi:10.1128/MCB.19.12.7925
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931-21936. doi:10.1073/pnas.1016071107
- Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000. doi:10.1016/0092-8674(87)90585-X
- Eddy, E. M. (2002). Male germ cell gene expression. *Recent Prog. Horm. Res.* **57**, 103-128. doi:10.1210/rp.57.1.103
- Fan, Y., Yue, J., Xiao, M., Han-Zhang, H., Wang, Y. V., Ma, C., Deng, Z., Li, Y., Yu, Y., Wang, X. et al. (2017). FXR1 regulates transcription and is required for growth of human cancer cells with TP53/FXR2 homozygous deletion. *eLife* **6**, e26129. doi:10.7554/eLife.26129
- Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J. and Lobanov, V. V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell. Biol.* **16**, 2802-2813. doi:10.1128/mcb.16.6.2802
- Geuens, T., Bouhy, D. and Timmerman, V. (2016). The hnRNP family: insights into their role in health and disease. *Hum. Genet.* **135**, 851-867. doi:10.1007/s00439-016-1683-5
- Hao, J., Yamamoto, M., Richardson, T. E., Chapman, K. M., Denard, B. S., Hammer, R. E., Zhao, G. Q. and Hamra, F. K. (2008). Sohlh2 knockout mice are male-sterile because of degeneration of differentiating type A spermatogonia. *Stem Cells* **26**, 1587-1597. doi:10.1634/stemcells.2007-0502
- Hartmuth, K., Urlaub, H., Vornlocher, H.-P., Will, C. L., Gentzel, M., Wilm, M. and Lührmann, R. (2002). Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc. Natl. Acad. Sci. USA* **99**, 16719-16724. doi:10.1073/pnas.262483899
- Hasegawa, K., Sin, H.-S., Maezawa, S., Broering, T. J., Kartashov, A. V., Alavattam, K. G., Ichijima, Y., Zhang, F., Bacon, W. C., Greis, K. D. et al. (2015). SCML2 establishes the male germline epigenome through regulation of histone H2A ubiquitination. *Dev. Cell* **32**, 574-588. doi:10.1016/j.devcel.2015.01.014
- Hou, L., Wei, Y., Lin, Y., Wang, X., Lai, Y., Yin, M., Chen, Y., Guo, X., Wu, S., Zhu, Y. et al. (2020). Concurrent binding to DNA and RNA facilitates the pluripotency reprogramming activity of Sox2. *Nucleic Acids Res.* **48**, 3869-3887. doi:10.1093/nar/gkaa067
- Hsu, S.-H., Hsieh-Li, H.-M. and Li, H. (2004). Dysfunctional spermatogenesis in transgenic mice overexpressing bHLH-Zip transcription factor, Spz1. *Exp. Cell Res.* **294**, 185-198. doi:10.1016/j.yexcr.2003.10.029
- Hudson, W. H. and Ortlund, E. A. (2014). The structure, function and evolution of proteins that bind DNA and RNA. *Nat. Rev. Mol. Cell Biol.* **15**, 749-760. doi:10.1038/nrm3884
- Jeon, Y. and Lee, J. T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **146**, 119-133. doi:10.1016/j.cell.2011.06.026
- Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C. Y., Xiao, R., Burge, C. B. and Fu, X.-D. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855-868. doi:10.1016/j.cell.2013.04.028
- Kashiwabara, S.-I., Zhuang, T., Yamagata, K., Noguchi, J., Fukamizu, A. and Baba, T. (2000). Identification of a novel isoform of poly(A) polymerase, TPAP, specifically present in the cytoplasm of spermatogenic cells. *Dev. Biol.* **228**, 106-115. doi:10.1006/dbio.2000.9894
- Kashiwabara, S.-I., Noguchi, J., Zhuang, T., Ohmura, K., Honda, A., Sugiura, S., Miyamoto, K., Takahashi, S., Inoue, K., Ogura, A. et al. (2002). Regulation of spermatogenesis by testis-specific, cytoplasmic poly(A) polymerase TPAP. *Science (New York, N.Y.)* **298**, 1999-2002. doi:10.1126/science.1074632
- Kim, D., Paggi, J. M., Park, C., Bennett, C. and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907-915. doi:10.1038/s41587-019-0201-4
- Kimmins, S., Crosio, C., Kotaja, N., Hirayama, J., Monaco, L., Höög, C., van Duin, M., Gossen, J. A. and Sassone-Corsi, P. (2007). Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. *Mol. Endocrinol.* **21**, 726-739. doi:10.1210/me.2006-0332
- Kistler, W. S., Baas, D., Lemeille, S., Paschaki, M., Seguin-Estevez, Q., Barras, E., Ma, W., Duteyrat, J.-L., Morlé, L., Durand, B. et al. (2015). RFX2 is a major transcriptional regulator of spermiogenesis. *PLoS Genet.* **11**, e1005368. doi:10.1371/journal.pgen.1005368
- Kosir, R., Juvan, P., Perse, M., Budefeld, T., Majdic, G., Fink, M., Sassone-Corsi, P. and Rozman, D. (2012). Novel insights into the downstream pathways and targets controlled by transcription factors CREM in the testis. *PLoS ONE* **7**, e31798. doi:10.1371/journal.pone.0031798
- Kung, J. T., Kesner, B., An, J. Y., Ahn, J. Y., Cifuentes-Rojas, C., Colognori, D., Jeon, Y., Szanto, A., del Rosario, B. C., Pinter, S. F. et al. (2015). Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* **57**, 361-375. doi:10.1016/j.molcel.2014.12.006
- Legrand, J. M. D. and Hobbs, R. M. (2018). RNA processing in the male germline: Mechanisms and implications for fertility. *Semin. Cell Dev. Biol.* **79**, 80-91. doi:10.1016/j.semcdb.2017.10.006
- Lin, S. and Fu, X.-D. (2007). SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.* **623**, 107-122. doi:10.1007/978-0-387-77374-2_7
- Lovci, M. T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T. Y., Stark, T. J., Gehman, L. T., Hoon, S. et al. (2013). Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* **20**, 1434-1442. doi:10.1038/nsmb.2699
- Luo, M., Zhou, J., Leu, N. A., Abreu, C. M., Wang, J., Anguera, M. C., de Rooij, D. G., Jasin, M. and Wang, P. J. (2015). Polycomb protein SCML2 associates with USP7 and counteracts histone H2A ubiquitination in the XY chromatin during male meiosis. *PLoS Genet.* **11**, e1004954. doi:10.1371/journal.pgen.1004954
- Martianov, I., Choukallah, M.-A., Krebs, A., Ye, T., Legras, S., Rijkers, E., Van Ijcken, W., Jost, B., Sassone-Corsi, P. and Davidson, I. (2010). Cell-specific occupancy of an extended repertoire of CREM and CREB binding loci in male germ cells. *BMC Genomics* **11**, 530. doi:10.1186/1471-2164-11-530
- Maruyama, O., Nishimori, H., Katagiri, T., Miki, Y., Ueno, A. and Nakamura, Y. (1998). Cloning of TCFL5 encoding a novel human basic helix-loop-helix motif protein that is specifically expressed in primary spermatocytes at the pachytene stage. *Cytogenet. Cell Genet.* **82**, 41-45. doi:10.1159/000015061
- Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429-440. doi:10.1128/MCB.20.2.429-440.2000
- Matveeva, E., Maiorano, J., Zhang, Q., Eteleeb, A. M., Convertini, P., Chen, J., Infantino, V., Stamm, S., Wang, J., Rouchka, E. C. et al. (2016). Involvement of PARP1 in the regulation of alternative splicing. *Cell Discovery* **2**, 15046. doi:10.1038/celldisc.2015.46
- Matzuk, M. M. and Lamb, D. J. (2008). The biology of infertility: research advances and clinical challenges. *Nat. Med.* **14**, 1197-1213. doi:10.1038/nm.f.1895

- Melikishvili, M., Chariker, J. H., Rouchka, E. C. and Fondufe-Mittendorf, Y. N.** (2017). Transcriptome-wide identification of the RNA-binding landscape of the chromatin-associated protein PARP1 reveals functions in RNA biogenesis. *Cell Discov.* **3**, 17043. doi:10.1038/celldisc.2017.43
- Miranda-Vizuete, A., Tsang, K., Yu, Y., Jiménez, A., Pelto-Huikko, M., Flickinger, C. J., Sutovsky, P. and Oko, R.** (2003). Cloning and developmental analysis of murid spermatid-specific thioredoxin-2 (SPTRX-2), a novel sperm fibrous sheath protein and autoantigen. *J. Biol. Chem.* **278**, 44874-44885. doi:10.1074/jbc.M305475200
- Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B. et al.** (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-544. doi:10.1016/0092-8674(89)90434-0
- Nakajima, T., Uchida, C., Anderson, S. F., Lee, C.-G., Hurwitz, J., Parvin, J. D. and Montminy, M.** (1997). RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* **90**, 1107-1112. doi:10.1016/S0092-8674(00)80376-1
- Nalam, R. L. and Matzuk, M. M.** (2010). Local signalling environments and human male infertility: what we can learn from mouse models. *Expert Rev. Mol. Med.* **12**, e15. doi:10.1017/S1462399410001468
- Nantel, F., Monaco, L., Foulkes, N. S., Masquillier, D., LeMeur, M., Henriksen, K., Dierich, A., Parvinen, M. and Sassone-Corsi, P.** (1996). Spermiogenesis deficiency and germ-cell apoptosis in CREM-mutant mice. *Nature* **380**, 159-162. doi:10.1038/380159a0
- Olson, E. N.** (1990). MyoD family: a paradigm for development? *Genes Dev.* **4**, 1454-1461. doi:10.1101/gad.4.9.1454
- Pan, J., Goodheart, M., Chuma, S., Nakatsuji, N., Page, D. C. and Wang, P. J.** (2005). RNF17, a component of the mammalian germ cell nuage, is essential for spermiogenesis. *Development* **132**, 4029-4039. doi:10.1242/dev.02003
- Paronetto, M. P., Messina, V., Bianchi, E., Barchi, M., Vogel, G., Moretti, C., Palombi, F., Stefanini, M., Geremia, R., Richard, S. et al.** (2009). Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis. *J. Cell Biol.* **185**, 235-249. doi:10.1083/jcb.200811138
- Saldaña-Meyer, R., González-Buendía, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F. and Reinberg, D.** (2014). CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* **28**, 723-734. doi:10.1101/gad.236869.113
- Schlegel, P. N.** (2009). Evaluation of male infertility. *Minerva Ginecol.* **61**, 261-283.
- Shi, Y., Zhang, L., Song, S., Teves, M. E., Li, H., Wang, Z., Hess, R. A., Jiang, G. and Zhang, Z.** (2013). The mouse transcription factor-like 5 gene encodes a protein localized in the manchette and centriole of the elongating spermatid. *Andrology* **1**, 431-439. doi:10.1111/j.2047-2927.2013.00069.x
- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R. and Oberdoerffer, S.** (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**, 74-79. doi:10.1038/nature10442
- Siep, M., Sleddens-Linkels, E., Mulders, S., van Eenennaam, H., Wassenaar, E., Van Cappellen, W. A., Hoogerbrugge, J., Grootegoed, J. A. and Baarends, W. M.** (2004). Basic helix-loop-helix transcription factor Tcf15 interacts with the Calmegin gene promoter in mouse spermatogenesis. *Nucleic Acids Res.* **32**, 6425-6436. doi:10.1093/nar/gkh979
- Sigova, A. A., Abraham, B. J., Ji, X., Molinie, B., Hannett, N. M., Guo, Y. E., Jangi, M., Giallourakis, C. C., Sharp, P. A. and Young, R. A.** (2015). Transcription factor trapping by RNA in gene regulatory elements. *Science (New York, N.Y.)* **350**, 978-981. doi:10.1126/science.aad3346
- Silber, S. J.** (1994). A modern view of male infertility. *Reprod. Fertil. Dev.* **6**, 93-103; discussion 103-104. doi:10.1071/rd9940093
- Smith, T. B., Baker, M. A., Connaughton, H. S., Habenicht, U. and Aitken, R. J.** (2013). Functional deletion of Txndc2 and Txndc3 increases the susceptibility of spermatozoa to age-related oxidative stress. *Free Radic. Biol. Med.* **65**, 872-881. doi:10.1016/j.freeradbiomed.2013.05.021
- Steigemann, P., Wurzenberger, C., Schmitz, M. H. A., Held, M., Guizetti, J., Maar, S. and Gerlich, D. W.** (2009). Aurora B-mediated abscission checkpoint protects against tetraploidization. *Cell* **136**, 473-484. doi:10.1016/j.cell.2008.12.020
- Suzuki, H., Ahn, H. W., Chu, T., Bowden, W., Gassei, K., Orwig, K. and Rajkovic, A.** (2012). SOHLH1 and SOHLH2 coordinate spermatogonial differentiation. *Dev. Biol.* **361**, 301-312. doi:10.1016/j.ydbio.2011.10.027
- Talasz, H., Lindner, H. H., Sarg, B. and Helliger, W.** (2005). Histone H4-lysine 20 monomethylation is increased in promoter and coding regions of active genes and correlates with hyperacetylation. *J. Biol. Chem.* **280**, 38814-38822. doi:10.1074/jbc.M505563200
- Tamanini, F., Bontekoe, C., Bakker, C. E., van Unen, L., Anar, B., Willemsen, R., Yoshida, M., Galjaard, H., Oostra, B. A. and Hoogveen, A. T.** (1999). Different targets for the fragile X-related proteins revealed by their distinct nuclear localizations. *Hum. Mol. Genet.* **8**, 863-869. doi:10.1093/hmg/8.5.863
- Toyoda, S., Miyazaki, T., Miyazaki, S., Yoshimura, T., Yamamoto, M., Tashiro, F., Yamato, E. and Miyazaki, J.-I.** (2009). Sohlh2 affects differentiation of KIT positive oocytes and spermatogonia. *Dev. Biol.* **325**, 238-248. doi:10.1016/j.ydbio.2008.10.019
- Van Nostrand, E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., Blue, S. M., Nguyen, T. B., Surka, C., Elkins, K. et al.** (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508-514. doi:10.1038/nmeth.3810
- VanGompel, M. J. W. and Xu, E. Y.** (2010). A novel requirement in mammalian spermatid differentiation for the DAZ-family protein Boule. *Hum. Mol. Genet.* **19**, 2360-2369. doi:10.1093/hmg/ddq109
- Wang, Y., Wang, H., Zhang, Y., Du, Z., Si, W., Fan, S., Qin, D., Wang, M., Duan, Y., Li, L. et al.** (2019). Reprogramming of meiotic chromatin architecture during spermatogenesis. *Mol. Cell* **73**, 547-561.e46. doi:10.1016/j.molcel.2018.11.019
- Wolf, F. W., Sarma, V., Seldin, M., Drake, S., Suchard, S. J., Shao, H., O'Shea, K. S. and Dixit, V. M.** (1994). B94, a primary response gene inducible by tumor necrosis factor-alpha, is expressed in developing hematopoietic tissues and the sperm acrosome. *J. Biol. Chem.* **269**, 3633-3640. doi:10.1016/S0021-9258(17)41909-0
- Xiao, R., Chen, J. Y., Liang, Z., Luo, D., Chen, G., Lu, Z. J., Chen, Y., Zhou, B., Li, H., Du, X. et al.** (2019). Pervasive chromatin-RNA binding protein interactions enable RNA-based regulation of transcription. *Cell* **178**, 107-121.e118. doi:10.1016/j.cell.2019.06.001
- Zang, C., Schones, D. E., Zeng, C., Cui, K., Zhao, K. and Peng, W.** (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics (Oxford, England)* **25**, 1952-1958. doi:10.1093/bioinformatics/btp340
- Zhang, S., Herrmann, C. and Grosse, F.** (1999). Pre-mRNA and mRNA binding of human nuclear DNA helicase II (RNA helicase A). *J. Cell Sci.* **112**, 1055-1064. doi:10.1242/jcs.112.7.1055