Biology of Reproduction, 2019, 101(4), 842–853
doi:10.1093/biolre/ioz115
Research Article
Advance Access Publication Date: 10 July 2019

OXFORD

Research Article

A 50-bp enhancer of the mouse acrosomal vesicle protein 1 gene activates round spermatid-specific transcription in vivo†

Craig Urekar[1], Kshitish K. Acharya[1], Preeti Chhabra[1] and Prabhakara P. Reddi[2,*]

[1]Department of Pathology, University of Virginia School of Medicine, Charlottesville, Virginia, USA [2]Department of Comparative Biosciences, College of Veterinary Medicine, University of Illinois Urbana Champaign, Champaign, Illinois, USA

***Correspondence:** Department of Comparative Biosciences, College of Veterinary Medicine, University of Illinois Urbana Champaign, 2001 S Lincoln Ave, Urbana, Illinois 61802, USA. E-mail: preddi@illinois.edu

†**Grant Support:** This work was supported by National Institutes of Health HD 36239 (PPR). The UVA Biomedical Research Facility is funded by a grant from the University of Virginia Pratt Committee.

Received 6 February 2019; Revised 4 June 2019; Accepted 3 July 2019

Abstract

Enhancers are cis-elements that activate transcription and play critical roles in tissue- and cell type-specific gene expression. During spermatogenesis, genes coding for specialized sperm structures are expressed in a developmental stage- and cell type-specific manner, but the enhancers responsible for their expression have not been identified. Using the mouse acrosomal vesicle protein (*Acrv1*) gene that codes for the acrosomal protein SP-10 as a model, our previous studies have shown that *Acrv1* proximal promoter activates transcription in spermatids; and the goal of the present study was to separate the enhancer responsible. Transgenic mice showed that three copies of the −186/−135 fragment (50 bp enhancer) placed upstream of the *Acrv1* core promoter (−91/+28) activated reporter expression in testis but not somatic tissues ($n = 4$). Immunohistochemistry showed that enhancer activity was restricted to the round spermatids. The *Acrv1* enhancer failed to activate transcription in the context of a heterologous core promoter ($n = 4$), indicating a likely requirement for enhancer-core promoter compatibility. Chromatin accessibility assays showed that the *Acrv1* enhancer assumes a nucleosome-free state in male germ cells (but not liver), indicating occupancy by transcription factors. Southwestern assays (SWA) identified specific binding of the enhancer to a testis nuclear protein of 47 kDa (TNP47). TNP47 was predominantly nuclear and becomes abundant during the haploid phase of spermatogenesis. Two-dimensional SWA revealed the isoelectric point of TNP47 to be 5.2. Taken together, this study delineated a 50-bp enhancer of the *Acrv1* gene for round spermatid-specific transcription and identified a putative cognate factor. The 50-bp enhancer could become useful for delivery of proteins into spermatids.

Summary Sentence

Functional assays and chromatin characteristics show that the −186/−135 region of the mouse *Acrv1* gene acts as a transcriptional enhancer and activates round spermatid-specific gene expression in vivo.

Key words: enhancer, chromatin, gene expression, acrosome, spermatogenesis, testis, transcriptional regulation, transgenic mice

Introduction

Genes coding for protein constituents of specialized compartments of the spermatozoa such as the acrosome and flagellum are expressed in a precise developmental stage- and germ cell type-specific manner [1]. Deviation from this programmed gene expression could be detrimental for the progression of spermatogenesis and may lead to male infertility [2]. Therefore, it is important to understand the transcriptional regulatory mechanisms that ensure proper gene expression during spermatogenesis. Progress in this area has been slow owing to the lack of a convenient cell culture system that can recapitulate male germ cell differentiation. By performing reporter assays in transgenic mice, several laboratories have identified testis-specific gene promoters. One theme that emerged is that short proximal promoters (100–300 bp) are enough for male germ cell-specific transcription in vivo in mice [3–10]. In the case of the gene for testis angiotensin-converting enzyme, the −91/+17 promoter was shown to be sufficient for spermatid-specific expression of a reporter gene in transgenic mice [4]. While these studies serve as a good starting point, further delineation of the promoter into enhancer and core promoter elements and identification of the cognate transcription factors (TF) will be important to piece together the requirements for testis- and germ cell-specific gene expression.

Enhancers are defined as deoxyribonucleic acid (DNA) sequences that activate transcription from distal core promoters independent of orientation and distance [11]. When placed in a foreign chromatin environment, enhancers can still activate heterologous reporter genes. Functional assays serve as the gold standard to validate the above characteristics of an enhancer. Enhancers work in collaboration with core promoters [12, 13]. The core promoter provides a platform for the assembly of general TF and ribonucleic acid polymerase II (RNA Pol II) to direct transcription initiation [14]. Depending on whether it contains a TATA-box, initiator, or both, the composition of the pre-initiation complex varies. The enhancer-binding factors interact with the core machinery and activate transcription. Recent work has shown that enhancers are selective in terms of core promoter usage suggesting biochemical compatibility among the proteins recruited [15]. Thus, core promoters are, as much a part of the transcriptional specificity scheme as, the enhancers.

In a typical mammalian gene, enhancers are located several kilobases away from the core promoter invoking models of looping that bring the enhancer and core promoter together [16]. The male germ cells present a different scenario because all the promoters identified thus far are small and the regulatory elements are clustered within a few hundred base pairs of the transcriptional start site (TSS), but no previous work has attempted to separate the enhancer elements from the core promoter elements. This is important in order to understand how enhancers work in the male germ line. Identification of male germ cell-specific enhancers and their modular parts is critical to understand the unique mechanistic aspects of transcriptional regulation during spermatogenesis. Further, this information will be helpful to anticipate functional consequences of mutations and single nucleotide polymorphisms.

Our studies have used the mouse acrosomal vesicle protein (*Acrv1*) gene, which codes for the evolutionarily conserved acrosomal protein SP-10, as a model gene to study the regulation of round spermatid-specific gene transcription. Our initial work showed that the −408/+28 *Acrv1* promoter drives round spermatid-specific transcription in transgenic mice [10]. Subsequent works using truncated versions showed that the −186/+28 proximal promoter retains the same transcriptional specificity in vivo [17]. The −91/+28 promoter, however, lacked reporter gene activity in vivo [10]. This suggested that cis-elements within the −186/−92 recruit the TF necessary for gene activation and that the −91/+28, which contains a canonical initiator element at the TSS, acts as a core promoter. Interestingly, negative regulatory elements, i.e., those required for gene repression in somatic tissues, are also located within the *Acrv1* proximal promoter. Our previous work showed that the −408/−92 *Acrv1* proximal promoter possesses insulator activity, which silences the *Acrv1* gene in somatic tissues [18]. Further work showed that the −186/−135 region acts as a minimal insulator in cell culture as well as in transgenic mice capable of preventing ectopic expression in somatic tissues [19].

In the present study, we asked if the same 50-bp fragment (−186/−135) could act as an enhancer for activating gene expression in round spermatids. We have addressed this using functional assays to assess the ability of the 50 bp to drive correct spatiotemporal gene expression in vivo. Functional characterization of the *Acrv1* enhancer will provide a focus to probe the mechanism of the tight regulation of male germ cell-specific gene transcription. The present study also examined the chromatin structure of the endogenous mouse *Acrv1* gene promoter in somatic tissue vs. germ cells to gain clues as to the regulation at the level of chromatin. Finally, we present data on attempts to identify the cognate testis nuclear factors binding to the −186/−135 *Acrv1* enhancer.

Materials and methods

Generation of luciferase reporter gene constructs

The −91/+28 *Acrv1* core promoter was polymerase chain reaction (PCR) amplified from the −408GFP plasmid [10] with Bgl II and Hind III ends and cloned into pGL3 basic. Next, the −186/−135 *Acrv1* promoter fragment was PCR amplified with Bgl II ends and cloned into the Bgl II site of the above construct. Clones with one −186/−135 insertion or three tandem insertions of the −186/−135 fragment (all in forward orientation) were selected. To generate the construct with the thymidine kinase (TK) core promoter, PCR amplified the −61/−31 region of the herpes simplex virus thymidine kinase (HSV TK) gene promoter from pRL-TK vector (Promega, WI) with Bgl II and Hind III sites and cloned into pGL3 basic. Next, the −186/−135 fragment was cloned into the Bgl II site as above; clones with one copy of −186/−135 were selected. All constructs were sequenced to ensure that the promoter fragments are correctly oriented.

Generation of transgenic mice

All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of Virginia. Mice were euthanized following the CO_2 asphyxiation method approved by the institution. Preparation of transgene fragments and generation of transgenic mice were performed as reported previously [19]. The above reporter constructs were cleaved with Bam HI + Xho I and the transgenes consisting of the promoter elements and luciferase reporter gene were excised from the above vectors, verified by sequencing, and used for microinjections to generate transgenic mice

by the University of Virginia Transgenic Core Facility. Genotyping was performed by PCR amplification of the luciferase gene as described previously [19].

Detection of reporter gene activity

Testis and various other tissues were collected from 10- to 12-week old transgenic mice and littermate controls and processed for measuring luciferase activity as reported before [19]. Briefly, tissues were homogenized in passive lysis buffer (Promega, WI), centrifuged at 16 000 × *g* for 2 min, supernatant was collected, and protein quantified using the Bio-Rad reagent (cat # 500-0006). Equal amounts of protein (50 µg) from each tissue used in luciferase assays were performed using the Promega Luciferase Assay reagent and measured in Turner Designs TD20/20 Luminometer. Luciferase activities were normalized based on the amount of protein used. Background value from non-transgenic mouse tissues (highest 0.15) was subtracted from the luciferase activities of transgenic mouse tissues. Next, formalin fixed testes from the above mice were processed for immunofluorescence using anti-firefly luciferase antibody (Abcam ab21176) following manufacturer's instructions. Images were taken on fluorescent microscope.

Micrococcal nuclease digestion of nuclei

Seminiferous tubules and liver from mice of specified ages were cut into 2 mm pieces in pre-chilled DMEM-F12 (Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12) medium and washed twice in the same medium. Collagenase (1 mg/mL) digestion was carried out at 34 °C for 12 min and stopped by adding ice cold medium. Next, digestion was carried out in collagenase, hyaluronidase, and trypsin (all at 1 mg/mL) at 34 °C with shaking. After the cells have separated as determined by light microscopic observation, the supernatant was spun at 150 × *g* at 4 °C and resuspended in DMEM-F12 media. The testis/liver cells were dounce homogenized in 30 mL of NB-1.0 buffer (15 mM Tris-Cl pH 7.4, 15 mM NaCl, 60 mM KCl, 0.1 mM EGTA (ethylene glycol-bis(β-aminoethyl ether), and 5 mM $MgCl_2$ + 1 M sucrose + 1 mM DTT (Dithiothreitol) + 1 mM PMSF (phenylmethylsulfonyl fluoride)). The homogenate was filtered through a single layer of Miracloth and centrifuged for 10 min at 7000 rpm using the HB6 rotor to collect nuclei. The nuclear pellet was resuspended in 10 mL of NB1.0 buffer and the suspension of nuclei layered over 20 mL of NB-1.7 (15 mM Tris-Cl pH 7.4, 15 mM NaCl, 60 mM KCl, 0.1 mM EGTA, and 5 mM $MgCl_2$ + 1.7 M sucrose + 1 mM DTT + 1 mM PMSF) and centrifuged for 20 min at 11 000 rpm in the HB6 rotor. The nuclei (pellet) were resuspended in NB-0.25 (15 mM Tris-Cl pH 7.4, 15 mM NaCl, 60 mM KCl, 0.1 mM EGTA, and 5 mM $MgCl_2$ + 0.25 M sucrose + 1 mM DTT) containing 0.4 mM $CaCl_2$. Micrococcal nuclease (MNase) was serially diluted in 25 µl NB0.25/2 mM $CaCl_2$ to obtain 0–4 units of MNase. Control did not receive MNase. 475 µL aliquots of nuclei (1 mg DNA/mL) were dispensed into the MNase containing tubes as well as the no MNase control labeled "0" tube. Micrococcal nuclease treatment was for 4 min at room temperature. Digestion was stopped by the addition of 12.5 µL of 0.5 M EDTA (Ethylenediaminetetraacetic acid)/25 µL of 10% SDS. Samples were deproteinized with Proteinase K (50 µg/mL) and digested overnight at 55 °C. DNA was extracted with Phenol:$CHCl_3$:Isoamylalcohol (25:24:1), $CHCl_3$ and digested with 25 µg of RNase A for 2 h at 37 °C. Following extraction first with Phenol:$CHCl_3$:Isoamylalcohol (25:24:1) and then $CHCl_3$, ethanol precipitation and cold 70% ethanol wash, the DNA was resuspended in water.

Indirect end labeling

Twenty micrograms, each of MNase-treated DNA (as well as the no MNase control), were digested to completion with Eco RI, ethanol precipitated, and washed with cold 70% ethanol. Air-dried pellet was resuspended in 20 µL of water. The MNase alone and MNase + Eco RI digested samples were separated on 2% agarose gels and subjected to standard Southern blotting procedure. The MNase hypersensitive cleavage sites were then identified by Southern blot hybridization analysis using indirect end labeling method. The indirect end labeling detection method employs a short hybridization probe that abuts one end of the restriction enzyme fragment under investigation (Eco RI) and reveals specific MNase cleavage sites that occur at discrete distances from that end. The Eco RI site is located at −47 within the *Acrv1* promoter. We have used the −186/−47 and the +1/+106 (Exon 1) regions as probes. The DNA fragments were generated by PCR, purified, and radiolabeled with [^{32}P] dCTP using the Random Primer kit (Promega). Southern blotting and hybridization were performed as per standard procedures [20].

Nuclear protein extraction

Nuclear proteins were extracted from brain, heart, kidney lung, liver, spleen, and testes of adult (3–4 months old) SW (South western) male mice. TNPs were also extracted from pre-pubertal Swiss strain mice of two additional age groups: postnatal days 17 and 24. Round spermatids are absent in day 17 testes and abundant on day 24 [21].

Nuclear proteins were isolated following the previously published protocols with minor modifications [22]. Collagenase digestion step, usually performed after the decapsulation of the testes, was eliminated. For lysing the cells, minced and washed fragments of testes were homogenized in buffer A (10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), pH 7.9, 1.5 mM KCl, 0.5 mM DTT, 0.05% NP40, and protease inhibitors) using Kontes glass dounce homogenizer with 20 strokes of B type pestle. The volume of buffer C (20 mM HEPES, pH 7.9, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 6.25% glycerol, and protease inhibitors) used for lysing nuclei was equivalent to the volume of the nuclear pellet. The lysate was centrifuged to separate nuclear proteins in the supernatant. The supernatant was mixed with equal volumes of buffer D (20 mM HEPES, pH 7.9, 50 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, 20% glycerol, and protease inhibitors) and snap frozen in small aliquots before storing at –70 °C. The extracts were usually used within two months of the preparations. A protease inhibitor cocktail solution (Sigma) was used in all steps. The entire procedure was carried out at 4 °C.

Radiolabeling of oligonucleotide DNA probes

For radioactive labeling, various 40-bp *Acrv1* promoter fragments and unrelated (UR) control DNA of similar length, equimolar amounts of sense and anti-sense oligonucleotides corresponding to each fragment were annealed by first heating to 100 °C in annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, and 1 mM EDTA) for 5 min and then allowing to cool slowly at room temperature for 1 h. Radiolabeling was carried out by mixing 1 µL of fresh γ-^{32}P—dATP (Amersham-Pharmacia) was mixed with 20 pm of DNA in the presence of 10 units of T4 polynucleotide kinase in kinase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, and 5 mM DTT) for 30 min at 37 °C. The labeled probe(s) were diluted to 0.02 pm/µL. The probes were always used within the first half-life of the γ-^{32}P.

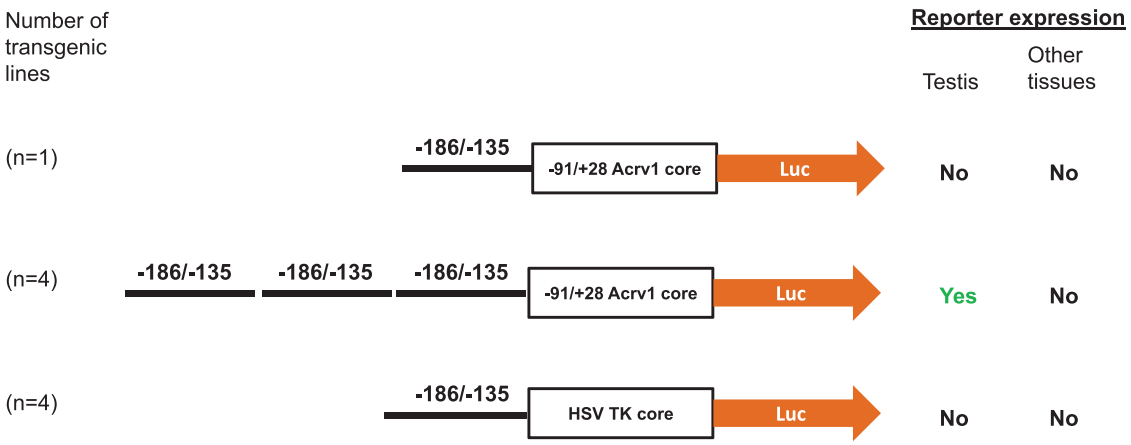

Figure 1. Schematic summary of the reporter gene constructs used to generate transgenic mouse lines in this study and the resulting reporter gene activity. The nucleotide coordinates of the *Acrv1* gene proximal promoter are with reference to the TSS as +1.

The sense strand of the −186/−148 *Acrv1* promoter fragment used as probe in Southwestern assays (SWA) in the present study is shown below (−186) 5′-GAAGCTACCCCTAACACACTATTCTA-CACACAGAAAAT-3′ (−148).

Southwestern analysis

This method [23, 24] was used for detecting proteins interacting with DNA. Nuclear proteins (50 µg/lane) were separated under reducing conditions on 8–12% SDS-polyacrylamide gels and electro-transferred to a nitrocellulose membrane (0.22 µ, Micron separation Inc., 135 Flanders Road, Westborough, MA 01581). Pre-stained broad range proteins (Amersham Pharmacia or Biorad and/or Life sciences technologies) were used as standard molecular weight markers. The proteins on the membranes were then blocked using 3% non-fat milk powder in a solution of 25 mM HEPES (pH 8.0), 1 mM DTT, 10% glycerol, 50 mM NaCl, and 1 mM EDTA for 2–3 h at room temperature. The membranes were then probed with labeled oligonucleotides in TNE-50 (Tris-Sodium Chloride-EDTA) (10 mM Tris of pH 7.5, 50 mM NaCl, 1 mM EDTA, and 1 mM DTT) in the presence of 10 µg/µL of non-specific plasmid DNA. Additional, size-matched but unlabeled oligonucleotides were added in different fold molar excess as competitors to this probe solution for sequence specificity studies. After 2 h of incubation with label at room temperature, the blots were washed in excess of TNE-50 for 1 h with repeated changes, until no radioactivity was detected in the wash buffer. The washed membranes were dried and subjected to autoradiography.

Two-dimensional electrophoresis

TNPs were separated by two-dimensional (2-D) electrophoresis following the procedure described previously [25]. Isoelectric focusing was performed in 15 × 0.15 cm acrylamide tube gels. Following the first-dimensional electrophoresis, the proteins in the tube gels were separated on 10% SDS-polyacrylamide slab gel using the BIO-RAD Mini-PROTEAN II electrophoresis cell. The slab gels with proteins were either silver-stained or processed for SWA as mentioned above. To identify the molecular weight and the iso-electric point (pI) of the protein spots on the silver-stained gels, 2-D SDS-PAGE standards (BIO-RAD) were used. The membranes with 2-D separated proteins were temporarily stained with ponceau (before probing with labeled oligonucleotide). Several gels and membranes with 2-D standards or the TNPs were stained and compared to identify and confirm the location of the protein spots of interest.

Proteomics

After determining the location of testis nuclear protein of 47 kDa (TNP47) as stated above, eight spots were cored (named KA1 through KA8) corresponding to 47 kDa molecular weight and pI of 5.2 (Supplemental Figure 1A) and submitted to the Biomolecular Research Facility of the University of Virginia. The core performed sequencing by liquid chromatography–mass spectrometry (LC–MS) using an LCQ (brand name) ion trap machine (Finnigan Austin, TX). The peptide data were analyzed by searching the NCBI non-redundant database using the Sequest algorithm. No data were obtained for spots KA1, 6, and 7. All other spots identified as *Mus musculus* tubulin beta-4B chain (NCBI Accession number NP_666228).

Results

The −186/−135 fragment activates spermatid-specific gene expression in vivo

The goal of this experiment was to test if the −186/−135 *Acrv1* proximal promoter fragment possessed enhancer activity and whether it would be necessary and sufficient for activating round spermatid-specific transcription. To address this, we generated mice with transgenes in which either one or three copies of the −186/−135 fragment were placed upstream of the −91/+28 *Acrv1* core promoter and luciferase reporter gene (Figure 1). Thus, the −186/−135 fragment was used in combination with the native (*Acrv1*) core promoter to test if the 50-bp fragment was necessary for round spermatid-specific transcription. We obtained only one transgenic mouse line using the transgene with one copy of the −186/−135 fragment, and it did not express the reporter gene in testis or any other tissue tested. Reporter gene expression was determined by measuring the luciferase activity. Reverse transcriptase-PCR was used to confirm that lack of luciferase expression was not due to post-transcriptional regulation (data not shown). Based on the data obtained from the lone transgenic line, one cannot conclude whether the 50-bp *Acrv1* promoter fragment was necessary for spermatid-specific gene expression. On the other hand, we obtained four separate transgenic lines with the transgene that included three tandem copies of the −186/−135

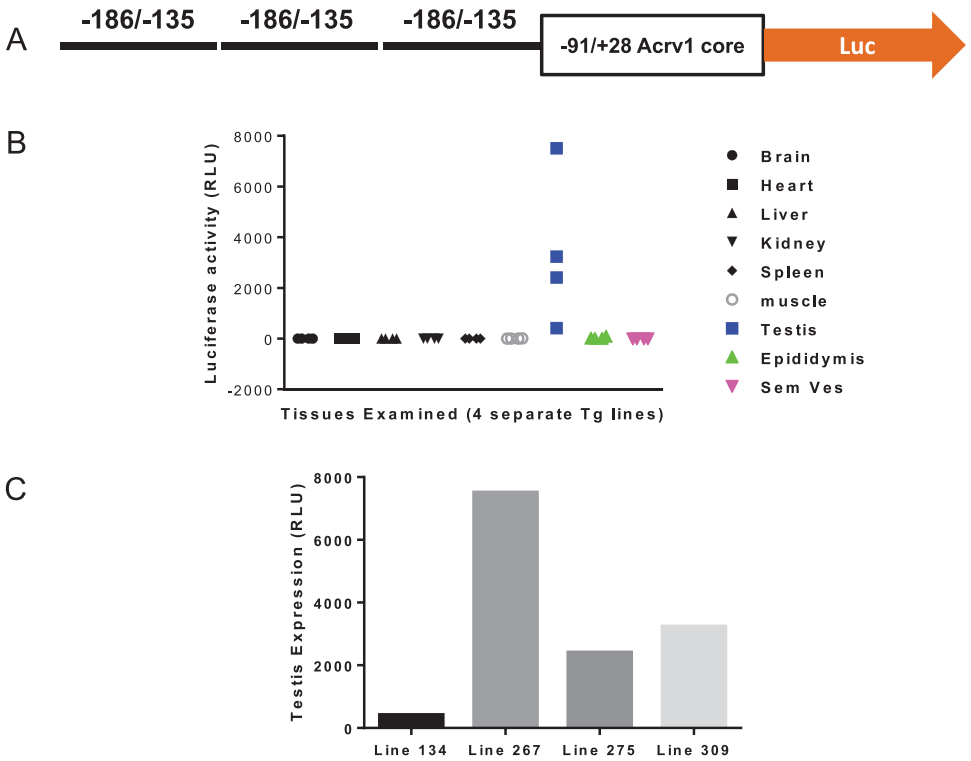

Figure 2. Three copies of the −186/−135 *Acrv1* promoter drove testis-specific reporter activity in transgenic mice in the context of *Acrv1* core promoter. (A) Schematic showing the components of the transgene; (B) luciferase reporter gene activity in various tissues of four independent mouse lines (four data points for each tissue) showing testis-specificity; and (C) levels of reporter expression varied but all four lines consistently expressed the transgene in the testis.

fragment placed upstream of the *Acrv1*–91/+28 core promoter and luciferase reporter (Figure 1). All four lines expressed luciferase in the testis but not in any of the other tissues tested including the brain, liver, lung, kidney, intestine, muscle, epididymis, and seminal vesicle (Figure 2B). Luciferase expression level within the testis varied for four independent mouse lines, which is consistent with the random integration of the transgene (Figure 2C). This suggested that the −186/−135 *Acrv1* promoter fragment was necessary and it acts as an enhancer for testis-specific transcription. To address sufficiency, the 50-bp enhancer was placed in conjunction with the heterologous HSV TK core promoter (Figure 1, lowest panel). The HSV TK promoter is ubiquitously an active promoter. Further, it was previously shown to be active in the testis and in the male germ cells, and therefore suitable for testing sufficiency [26]. Four separate transgenic lines were obtained, but none of the lines expressed the reporter gene in testis or any other tissue (Figure 1). Because the data were obtained from four independent transgenic lines ($n = 4$), we conclude that the 50-bp *Acrv1* enhancer is not sufficient to drive spermatid-specific transcription.

Next, using mice that expressed luciferase in the testis, we asked if the 50-bp enhancer retained the endogenous *Acrv1*-like developmental stage- and cell-type specificity within the seminiferous epithelium. We performed immunohistochemistry on adult mouse testis cross-sections from each transgenic mouse line using anti-luciferase antibody and found that reporter gene expression was restricted to the postmeiotic round spermatids (Figure 3B). There was no expression in spermatogonia, spermatocytes, sertoli, or in the interstitial cells of the testis (Figure 3A and B). Testis cross-sections from non-transgenic littermates were negative, indicating the specificity of the anti-luciferase antibody (Figure 3C and D). The above data suggested that the −186/−135 enhancer was capable of activating round spermatid-specific transcription in vivo.

The *Acrv1* proximal promoter assumes nucleosome-free state in germ cells

We probed the structure of the chromatin surrounding the mouse *Acrv1* gene promoter using the MNase method. Nucleosomes, the fundamental units of chromatin, assume a beads-on-a-string configuration when the gene is actively transcribed; they are organized into higher order chromatin in tissues, where the gene is not expressed. To determine the nucleosome organization of the endogenous *Acrv1* promoter region, we took advantage of the characteristic property of MNase, which causes double strand breaks in nucleosomal linker regions as well as in regions that are free of nucleosomes. Nuclei were prepared from testicular germ cells and liver cells (negative control). Germ cells from postnatal day 18 (PND18) and PND21 mice testes represent premeiotic and postmeiotic timepoints, respectively [27]. They were chosen for the experiment because PND18 germ cells do not express *Acrv1* mRNA, whereas PND21 germ cells do [10]. Nuclei isolated from germ cells and liver were subjected to limited MNase digestion and genomic DNA was isolated; a portion of this was subjected to Eco RI digestion (reason explained below). Micrococcal nuclease digested as well as MNase + Eco RI digested DNA samples were separated by agarose gel electrophoresis, transferred to membrane, and used for Southern hybridization (Figure 4). Micrococcal nuclease treatment resulted in genomic DNA fragments that corresponded to increments of the nucleosome core particle (147pb). To investigate the nucleosome structure at the *Acrv1* gene, Southern hybridization was performed

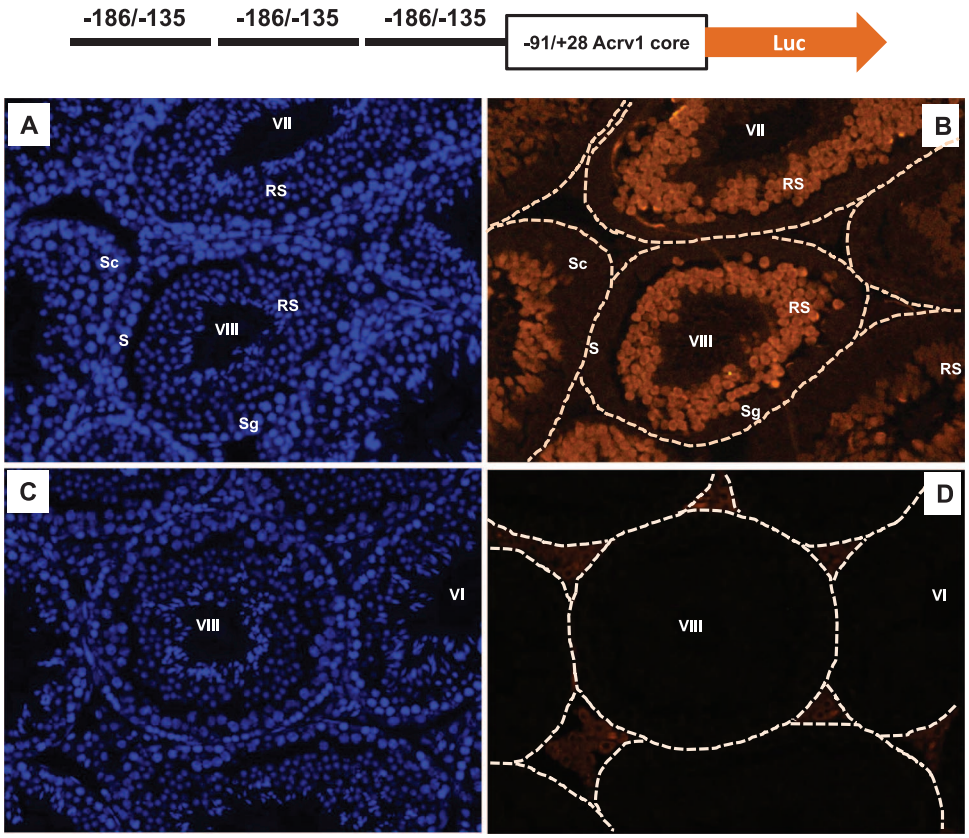

Figure 3. The −186/−135 promoter drove round spermatid-specific expression of the reporter gene. Cross sections of testis from mouse line 267 (Figure 1) were immunolabeled with anti-luciferase antibody (B). Note that the signal is restricted to round spermatids within each tubule cross section. Non-transgenic control testis showed no immunoreactivity with germ cells (D). Panels A and C show DAPI (4′,6-diamidino-2-phenylindole) counterstain for B and D, respectively. Roman numeral indicates stages; S, sertoli; Sg, spermatogonia; Sc, spermatocytes; and RS, round spermatids.

with probes derived from regions upstream and downstream of the TSS of *Acrv1* (Figure 4A and B, schematic on the right). Southern hybridization using the Exon 1 probe, which spans from +1 to +106 nucleotides of the *Acrv1* gene, showed a ladder of bands in liver that corresponded to multiples of 147-bp DNA, indicating that the *Acrv1* gene locus is assembled into a higher order structure in tissues that do not express *Acrv1* mRNA. In testicular germ cell nuclei, however, with increasing amount of MNase, the nucleosomal ladder gave way to predominantly mono- and di-nucleosomes (black arrow head and arrow, respectively in Figure 4A), indicating that the *Acrv1* gene assumes an open configuration (allowing easy access to MNase to cut between nucleosomes) in germ cells (Figure 4A). Eco RI cleavage of the MNase-digested DNA was done to gain insight into nucleosome positioning at the *Acrv1* promoter. An Eco RI site is located at the −47 position in the endogenous mouse *Acrv1* gene (Figure 4, schematic on the right). Using a short hybridization probe abutting one end of the Eco RI cleaved fragment would reveal specific MNase cleavage sites at discrete distances from that end. When the +1 to +106 (Exon 1) was used as a probe, prominent mono- and di-nucleosome bands were observed in germ cells, and there was no difference in size between MNase alone vs. MNase + EcoRI. This suggests that the +1 nucleosome encompassing the Exon 1 region is a stable nucleosome.

Next, to obtain a comparative picture of the nucleosome status at the *Acrv1* promoter, the above blot was stripped and re-hybridized with the −186/−47 probe (Figure 4B). Liver showed the typical nucleosomal ladder indicative of ordered nucleosomes, but the germ cells showed only faint hybridization signals. Although the size of the hybridized bands corresponded to mono- and di-nucleosomes, the intensity of hybridization was much weaker compared to the Exon 1 probe indicating that the *Acrv1* proximal promoter region assumes a relatively nucleosome-free state in germ cells (Figure 4B). Nucleosome-free region (NFR) would be accessible for MNase cuts at multiple places leaving less representation of nucleosomal length DNA. Further, the MNase + Eco RI digested DNA showed a smear below the mononucleosomal band. In Figure 4B, note how the smear denoted by the red arrow is lower compared to the mononucleosomal length shown by the blue arrow. The smear indicates that the −1 nucleosome is unstable in germ cells. It is noteworthy that in both PND18 and PND21samples, the −186/−47 promoter region showed a nucleosome-free state (Figure 4B).

Identification of testicular transcription factors binding to the spermatid-specific *Acrv1* enhancer

Our previous work showed that the −186/+28 promoter of the mouse *Acrv1* gene was capable of directing testis- and round spermatid-specific expression of a reporter gene in transgenic mice [17]. On the other hand, the truncated −91/+28 region did not support transcription in mice and is considered as the core promoter of the *Acrv1* gene [10]. This suggested that the −186/−92 region

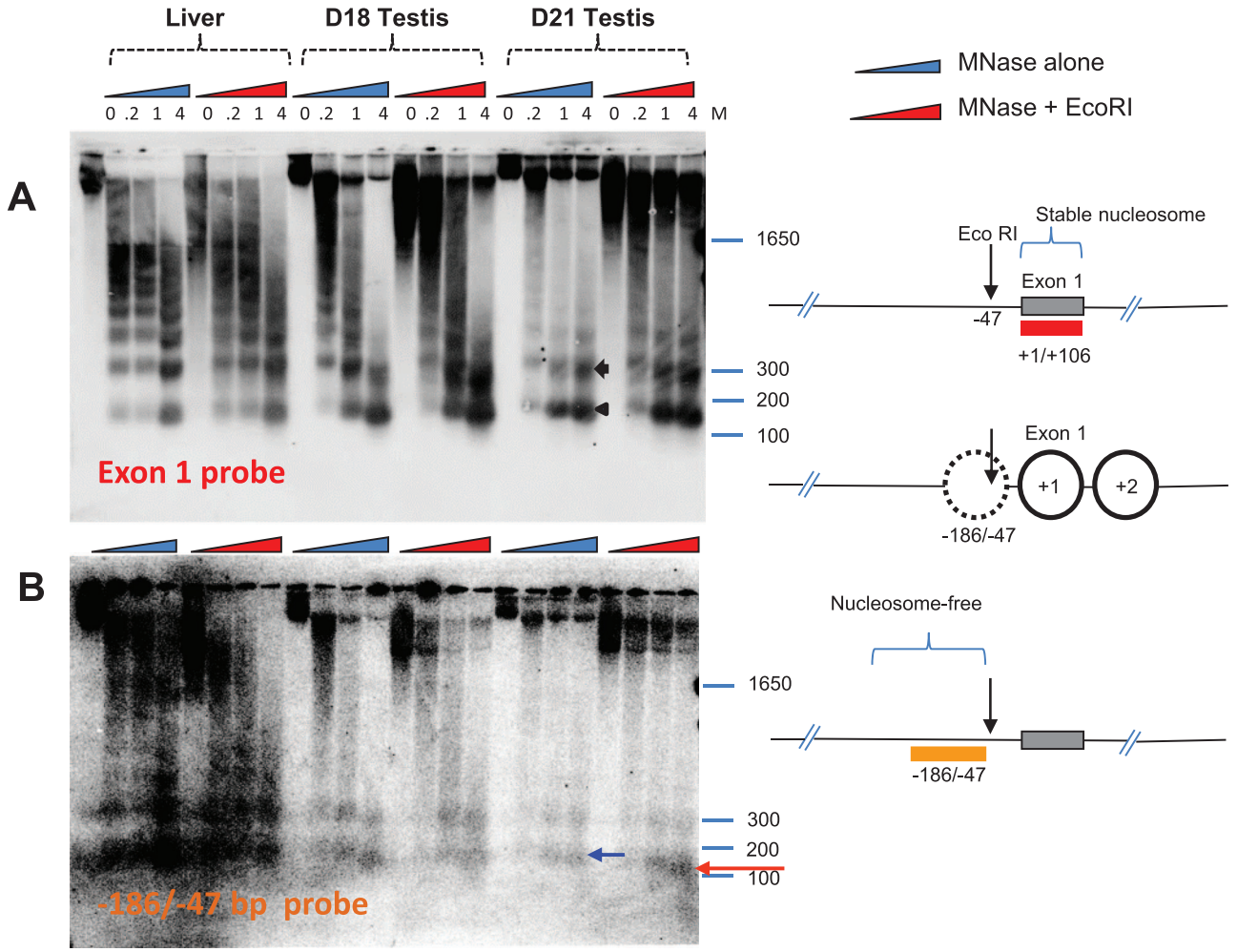

Figure 4. Chromatin structure at the *Acrv1* proximal promoter. Southern hybridization of MNase or MNase + Eco RI digested nuclei from liver, and testis of 18 (D18) or 21 (D21) days old mice probed with *Acrv1* Exon 1 probe (A) or the −186/−47 *Acrv1* promoter (B). Upper and lower schematics on the right depict the *Acrv1* gene region being analyzed, location of the Eco RI site, and the respective probes used in the experiment. The middle schematic shows the interpretation of the data.

must contain binding sites for TF responsible for spermatid-specific gene expression in vivo. To identify the cognate DNA binding proteins in testis extracts, we have employed the SWA. Nuclear extracts isolated from mouse testicular germ cells as well as several somatic tissues including the brain, heart, lung, liver, kidney, and spleen were separated by SDS-PAGE and transferred to nylon membrane. Several 40-bp double-stranded promoter fragments spanning the −186/−92 region of the *Acrv1* gene were radiolabeled to probe the membranes containing nuclear proteins (data not shown). Among the probes tested, the −186/−148 region (Figure 5, upper panel) showed specific binding to a 47 kDa protein present in testis nuclear extracts but not in heart, liver, lung, spleen, or kidney (Figure 5A). Two fainter bands at 47 and 38 kDa were observed in the brain nuclear extract. Thus, the −186/−148 region interacted with a 47 kDa band present predominantly in testis nuclear extracts. We named this band as testis nuclear protein of 47 kDa or TNP47. Binding of the −186/−148 dsDNA (double stranded DNA) to TNP47 was sequence-specific because UR DNA fragment could not compete for binding (Figure 5C). Interestingly, the −186/−148 probe falls within the −186/−135 50-bp *Acrv1* enhancer, capable of activating round spermatid-specific gene expression. To learn more about the significance of TNP47 to male germ cell differentiation, we used nuclear extracts from testes of mice aged 17 days, 24 days, or 3 months (Figure 5B). In postnatal testis, on day 17 (PND17), mouse testis will contain spermatogonia and spermatocytes but not round spermatids, whereas PND24 testis will contain round spermatids also. We observed that TNP47 binding activity was considerably higher in PND24 compared to PND17 indicating that TNP47 production begins in meiotic cells, but is upregulated within the haploid round spermatids. The source of TNP47 is not likely to be the somatic cells of the testis because the enzymatic digestion method used for germ cell preparation eliminates somatic cell contamination. Finally, TNP47 was abundant in the nuclear but not cytoplasmic extracts (Figure 5B), which would be expected of a protein involved in transcriptional activation. Thus, TNP47 is considered a putative cognate factor for the 50-bp enhancer.

Proteomics approach to identify TNP47

In order to determine the identity of TNP47, we followed 2D-SWA approach, in which testis nuclear proteins (TNP) were separated

Radiolabeled probe 5′-GAAGCTACCCCTAACACACTATTCTACACACAGAAAAT-3′
3′-CTTCGATGGGGATTGTGTGATAAGATGTGTGTCTTTTA-5′

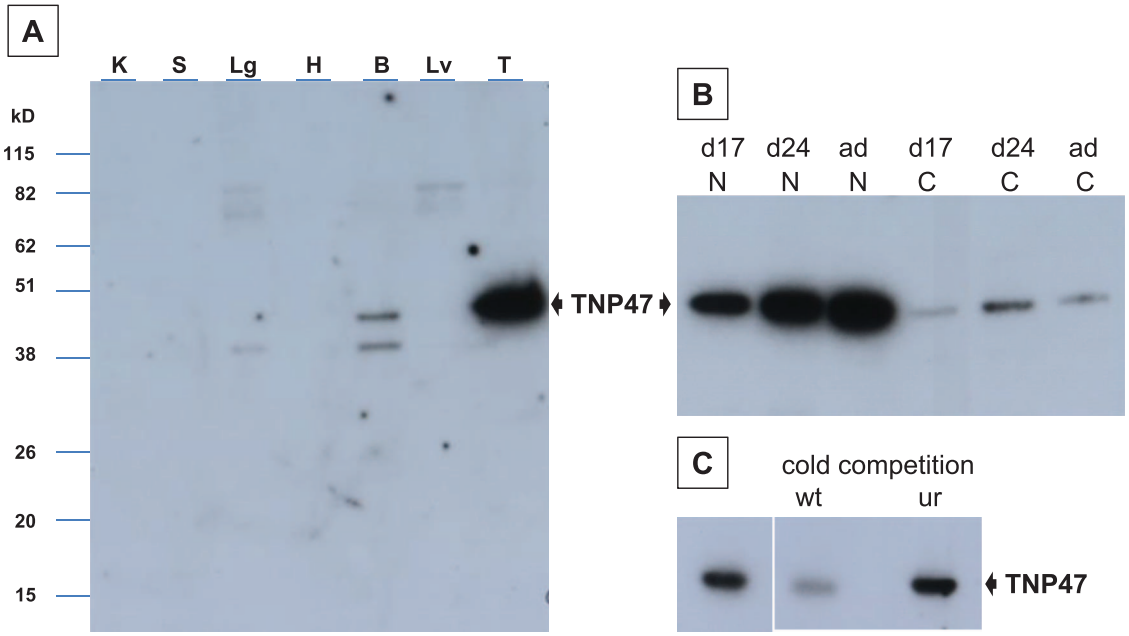

Figure 5. TNP47 identified as the cognate factor for the *Acrv1* promoter fragment. (A) SWA shows that a 47 kDa band in testis nuclear extract reacted with radiolabeled −186/−148 dsDNA. Brain extracts showed mild signal at a similar molecular weight. (B) TNP47 is a predominantly nuclear protein and it increases in abundance following haploid cell formation in the testis. (C) Specificity of TNP47. Cold (unlabeled) wild type −186/−148 DNA (wt) competes for TNP47 binding of the radiolabeled −186/−148 probe but not an UR DNA sequence of identical length.

by 2-D electrophoresis followed by SWA. In the first-dimensional testis, nuclear proteins were separated based on their isoelectric point (pI). The gel strip containing proteins focused according to their pI were then separated in the second dimension based on their molecular weight. Silver-stained image of the resulting 2-D gel is shown in Figure 6A. The gel was transferred to nitrocellulose membrane and used in SWA assays as above. The radiolabeled −186/−148 dsDNA interacted with a 2-D spot with pI of 5.2 and molecular weight of 47 kDa (Figure 6B.I). No hybridization signals were observed elsewhere in the blot. The area corresponding to TNP47 2-D hybridization signal is circled in Figure 6A. Competitive inhibition experiments showed that the wild type unlabeled −186/−148 fragment (Figure 6B.II). But not an UR fragment of similar size (Figure 6B.III), could compete for probe binding indicating the sequence-specificity of binding to the 2-D spot (pI 5.2; molecular weight [Mw] 47 kDa). With this information, we then performed proteomics. Parallel 2D gels were run; one was stained with silver stain to reveal the protein spot profile. The other gel was transferred to nitrocellulose membrane, stained with Ponceau S briefly and then used for SWA to reveal the TNP47 spot. After careful comparison of the resulting autoradiograph and the silver stained gel, TNP47 2D spot was located on the latter for coring (Supplemental Figure 1A). A total of eight spots (KA1 through KA8) were cored to make sure that the TNP47 2D spot at pI 5.2 and molecular weight of 47 kDa was recovered (Supplemental Figure 1A). Cored samples (KA1–KA8) were subjected to capillary LC–MS.

The peptide data were analyzed by searching the NCBI non-redundant database using the sequest algorithm. No data were obtained from spots KA1, KA6, and KA7. The remaining five (KA2–KA5, and KA8) spots were identified as tubulin-beta 4b chain (*M. musculus*). The peptides identified are shown in Supplemental Table 1. The protein coverage by amino acid count for each spot was as follows: KA 2 (28.6%), KA3 (32.9%), KA4 (6.5%), KA5 (4.5%), and KA8 (19.1%). Data showed that the cored spots were homogeneous and belonged to mouse tubulin-beta 4b chain (Supplemental Figure 1B). The calculated molecular weight for tubulin beta-4B (*M. musculus*) is 49.8 kDa and theoretical pI is 4.79 (Expasy Compute pI/Mw tool). This isoform of tubulin beta chain has been shown to be highly expressed in the testis compared to other tissues. We do not believe that TNP47 is beta tubulin for several reasons. First, the primary amino acid sequence of beta tubulin does not predict nucleic acid binding regions and there are no reports of beta tubulin being a DNA binding protein. Second, TNP47 showed testis-specific tissue distribution, whereas beta tubulin, the cytoskeletal protein, is ubiquitously expressed. Thus, one interpretation is that beta tubulin dominates in abundance at the same spot where TNP47 (putative TF) would be present in 2D gels. Thus, the identity of TNP47 has yet to be determined.

In silico approach to identify transcription factor binding sites in the 50-bp enhancer

To find potential TF binding sites within the 50-bp Acrv1 enhancer, we have performed TF database search using the TFBIND site (http://tfbind.hgc.jp/, Supplemental Table 2). This tool uses weight matrix in TF database TRANSFAC R.3.4. Although perfect matches

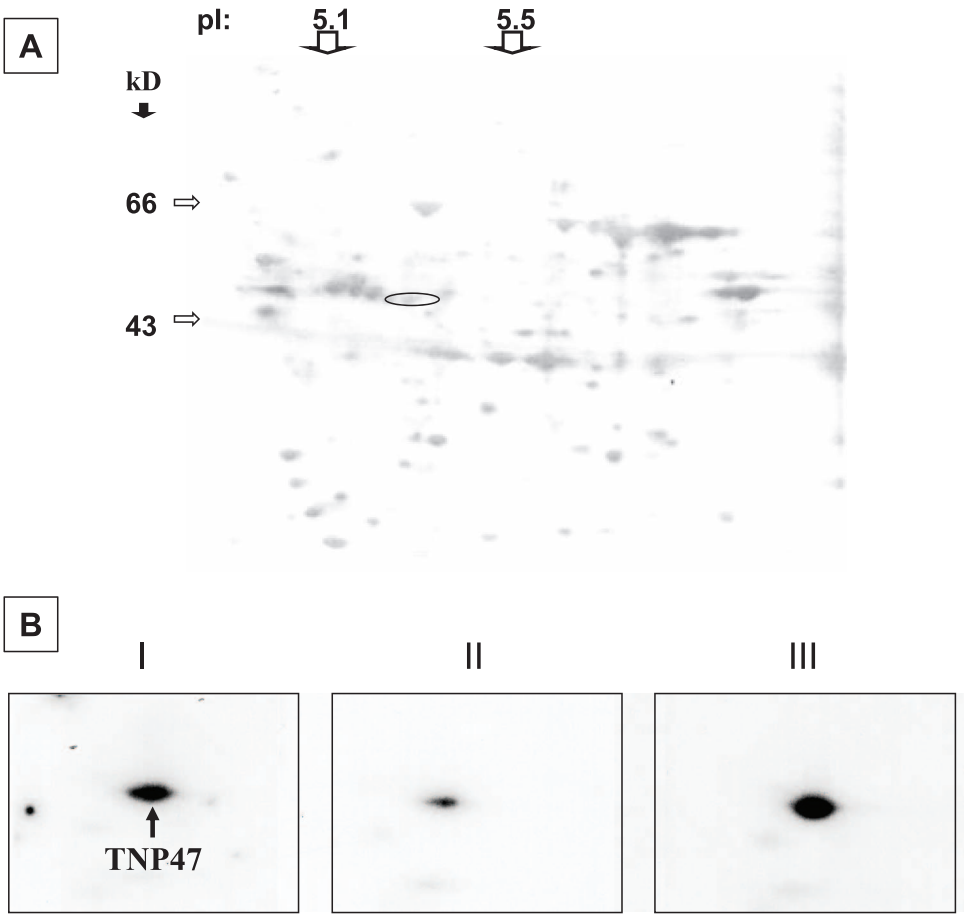

Figure 6. (A) 2D-SWA showing the isoelectric point (pI) and specificity of TNP47. TNPs separated by 2-D electrophoresis and silver stained (A). The circled area represents the location of TNP-47 based on hybridization signal obtained on probing with radiolabeled −186/−148 fragment. (B) Sequence specificity of the −186/−148 probe binding. TNP47 hybridization signal when probed with labeled −186/−148 fragment in the absence of competitor (I), in the presence of 150-fold excess unlabeled −186/−148 (II), or UR DNA of similar length (III) as competitor.

were not found with any known TF binding sites, several with one nucleotide mismatch were found. These include SRY, LYF1, STAT, HSF1, and HSF2 on the plus strand and MZF1, CEBP, and YY1 on the minus strand of the 50-bp enhancer (Figure 7). A complete list of TF sites can be found in Supplemental Table 2. Whether one or more of these factors bind to the 50-bp enhancer in vivo will be investigated by chromatin immunoprecipitation experiments in the future. To find out whether the 50-bp enhancer shares common elements with other spermiogenesis-specific genes, we focused on the consensus cis-elements listed for round spermatid-expressed genes. A recent study used single cell RNA seq to identify genes highly expressed in various types of testicular germ cells of the mouse and curated the over-represented cis-elements from the promoters of those genes [28]. Interestingly, out of 11 such sequences, listed for round spermatid expressed genes (gene group 5 in Figure 3B of Ref. [28]), we found 100% match with two cis-elements (ATTCTA and CACAGAA) and a partial matchwith two more (ACAAGAGA and HSF2) located within the 50-bp enhancer (Figure 7). Except for HSF2, the above promoter elements are unknown motifs, i.e., they do not match with any known TF binding site. It is interesting to note that HSF-2 site was identified within the 50-bp enhancer by the TF database search also.

Discussion

Studies on delineating specific enhancer and promoter regions required for testis- and germ cell-specific transcription in the mammalian system are limited. This is an important area of study toward understanding the transcriptional regulatory mechanisms that ensure normal progression of spermatogenesis. Using transgenic mice as a reporter system, here we show that a 50-bp proximal promoter segment (−186/−135) of the mouse *Acrv1* gene functions as an enhancer and activates transcription in round spermatids, and that the enhancer shows core promoter selectivity. Using biochemical approaches, we show that the enhancer region assumes a nucleosome-free state at the endogenous *Acrv1* locus in vivo to facilitate transcription, and that it binds to a 47 kDa factor present in testis nuclear extracts.

Transgene bearing three copies of the 50-bp enhancer placed upstream of the −91/+28 *Acrv1* core promoter activated testis-specific reporter expression. The assay tested the transcription-activation property of the *Acrv1* enhancer outside its endogenous DNA sequence and chromatin environment. It is noteworthy that all four lines of transgenic mice expressed luciferase in the testis (Figure 2). Expression without fail from four separate heterologous chromatin locations suggests that the 50-bp enhancer is hardwired

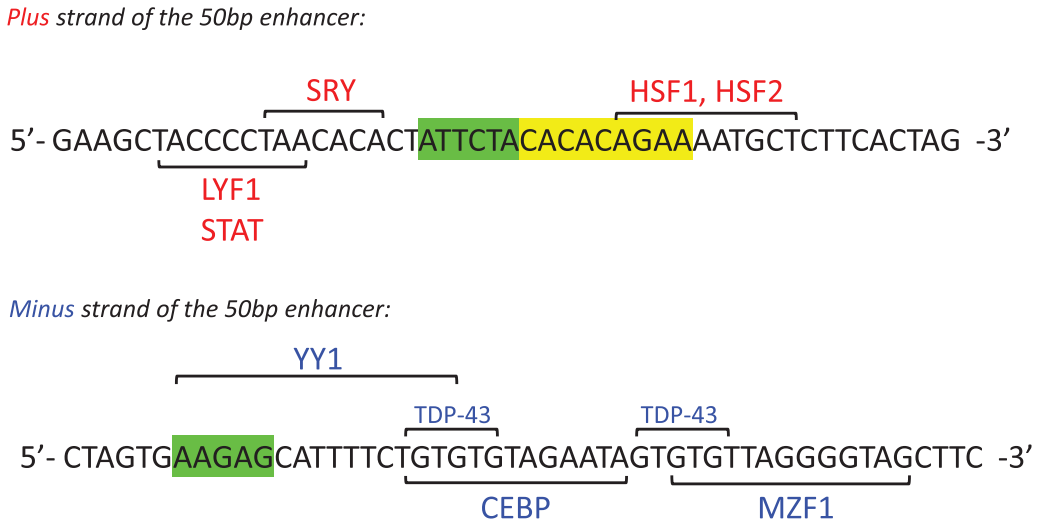

Figure 7. TF database search revealed the presence of binding sites (with 1 mismatch) for TF on the plus and minus strands of the 50-bp enhancer (see Supplemental Table 1). The green and yellow highlighted cis-elements are common to several other round spermatid-expressed genes of the mouse testis listed in Ref. [28]. TDP-43 binding sites have been previously confirmed by ChIP [30].

for testis expression and argues in favor of pioneer TF(s) binding to the 50-bp *Acrv1* enhancer within male germ cells. Variable levels of testis expression seen among the four transgenic lines (Figure 2C), however, suggests that the chromatin area surrounding the site of integration still influenced the enhancer activity. It is well known that the site of integration could both positively as well as negatively influence reporter gene expression [29]. In this context, lack of luciferase expression in any of the somatic tissues tested (Figure 2B) is consistent with the 50-bp fragment's insulator activity reported previously [19]. Based on biochemical studies of the endogenous *Acrv1* gene, we had previously postulated that tethering of the proximal promoter to the nuclear matrix silences the gene in somatic cells, whereas its release in germ cells permits transcription. Reporter expression data from this study (Figure 2) suggest that the −186/−135 *Acrv1* enhancer is capable of mediating these dual functions.

Apart from being testis-specific, it is remarkable that the 50-bp enhancer mimicked the precise spatiotemporal expression of the endogenous *Acrv1* gene. All four lines expressed the reporter gene exclusively within the round spermatids (Figure 3). Our previous work indicated that round spermatid-specific expression of *Acrv1* might be a two-step process involving pausing of transcription in spermatocytes followed by transcriptional activation in round spermatids [30]. We also reported that mutation of two ACACAC motifs within the −186/−135 region causes premature reporter expression in spermatocytes [17]. Taken together with the round spermatid-specific reporter activity seen in this study (Figure 3), it is reasonable to suggest that cis-elements required for the above two steps are located within the −186/−135 region of the mouse *Acrv1* gene.

Chromatin accessibility by MNase at the endogenous *Acrv1* promoter shed further light on testis and round spermatid-specific transcription of this gene. The proximal promoter region exhibited a higher order chromatin structure within the liver as expected because the *Acrv1* gene is not expressed in this tissue (Figure 4). However, in both meiotic spermatocytes (PND18) as well as post meiotic round spermatids (PND21), the promoter assumed an open configuration. The open chromatin state observed at PND18 is consistent with the bivalent histone marks and paused RNA Pol II shown at the *Acrv1* promoter in spermatocytes [30]. Chromatin being open at PND21 is reflective of active *Acrv1* mRNA transcription in round spermatids. This study showed that the −1 nucleosome is relatively nucleosome-free, whereas the downstream +1 nucleosome appeared stable (Figure 4A and B). It has been reported that stable +1 nucleosome are seen at promoters, which exhibit Pol II pausing [31]. NFR is commonly found at a majority of eukaryotic promoters near their TSS [32]. Typically, pioneer TF with nucleosome-binding properties bind to the regulatory promoter and recruit chromatin modifiers for unraveling the nucleosomes, thus allowing access for TF binding [33]. Nucleosome-free state of the −186/−47 region seen here is indicative of occupancy by TF.

Additional factors could be contributing to the nucleosome-free state of the endogenous *Acrv1* gene promoter within the germ cells. It has been shown that dA:dT stretches on either strand will resist nucleosome formation and that a 16 bp dA:dT with four mismatches can induce nucleosome-free state [34]. The mouse and human *Acrv1* promoters contain a 16-bp stretch of dT with 2 and 1 mismatches at the −202/−187, and − 231/−216 positions, respectively [10]. While this may aid in creating a nucleosome-free state at the endogenous locus, it is worth mentioning that the transgenes used in the present study did not include the above dT stretch. This makes the case stronger for the 50-bp enhancer recruiting pioneer TF(s), which can induce a chromatin state permissive for gene transcription within the male germ cells.

Our previous attempts to identify TF binding to the functional *Acrv1* promoter included the expression cloning method, whereby a testis cDNA library was screened with radiolabeled *Acrv1* promoter fragments [17]. That method yielded TDP-43, PURA, and Musashi, all of which are ubiquitously expressed TF/RNA binding proteins. Interestingly, the SWA method used in the present study identified a relatively testis-specific nuclear protein of 47kDa (Figure 5). The molecular weight and isoelectric point of TDP-43, PURA, and Musashi differ from those of TNP47, thus indicating the novelty of TNP47. Proteomics approach to identify TNP47 yielded tubulin beta-4B (Supplemental Figure 1B). We believe that it is an artifact caused by the migration of tubulin beta to the same location as TNP47 in 2-D gels. The relative abundance of

tubulin likely crowded the gel area leading to poor representation of TNP47 for proteomic identification. The characteristics of TNP47 shown here, i.e., testis-specificity and increased abundance in round spermatids, coupled with binding to the *Acrv1* enhancer, suggest the potential for it to be a pioneer TF for acrosomal biogenesis and round spermatid differentiation. Future studies will be aimed at determining the identity of TNP47. In silico analysis showed several TF as potential binding proteins of the 50-bp enhancer (Figure 7). Of these, SRY and HSF are of particular interest. Sox proteins share DNA binding domain with SRY. Sox5 and Sox30 have been shown to be expressed in round spermatids and knockout showed the latter to be essential for spermiogenesis [35,36]. Similarly, dual knockout of HSF1 and HSF2 leads to arrest of spermatogenesis at step3 round spermatid stage; these double knockout mice showed 9.8-fold reduction in the expression of the Acrv1 gene [37]. This attaches functional significance to the presence of HSF binding sites on the 50-bp enhancer of the *Acrv1* gene. We also sought to determine whether the 50-bp enhancer shares common elements with other spermiogenesis-specific genes. Interestingly, out of the 11 such sequences listed in a recent study [28], we found 100% match with two cis-elements (ATTCTA, and CACAGAA) and partial match with two more (ACAAGAGA, HSF2) located within the 50-bp enhancer (Figure 7). Overall, the 50-bp enhancer shares cis-elements in common with a number of genes expressed in round spermatids. The three unknown motifs (ATTCTA, CACAGAA, and ACAAGAGA) found within the functional 50 bp round spermatid enhancer are candidates for future identification of novel cognate TF.

One interesting observation that emerged from the current study is that the 50-bp *Acrv1* enhancer was functional when used in combination with its own core promoter but not with the heterologous HSV TK core promoter. The HSV TK core promoter used in the present study is a TATA box containing promoter whereas the *Acrv1* core promoter is a TATA-less and initiator-containing promoter. It is possible that TF recruited by the 50-bp enhancer require the initiator and other core elements present in the *Acrv1* core promoter to assemble the preinitiation complex. Recent studies using genome-wide approaches in the fruit-fly have shown that developmental enhancers work in combination with specific core promoters but not with house-keeping core promoters [15]. The idea of core promoter specificity for testis expression is supported by reports of testis-specific general TF [38–41]. An alternative explanation for why the 50-bp enhancer did not work with the HSV TK core promoter ($n = 4$) or with the native promoter ($n = 1$) might be that the 50-bp enhancer (−186/−135) is required to be located at a certain distance from the core promoter in order to be functional. Within the −186/+28 *Acrv1* promoter, shown to be functional in transgenic mice, the −186/135 enhancer is separated by 44 bp from the −91/+28 core promoter. In the current experiment, the 50-bp enhancer was juxtaposed to the core promoter. Future experiments will test whether a spacer of 44 bp would result in one copy of the 50-bp enhancer activating round spermatid-specific transcription from either the native or generic core promoter in vivo.

The property of imparting round spermatid expression by a 50-bp enhancer adds to the intrigue generated by previous studies showing that short proximal promoter regions (100–300 bp) are sufficient for testis-specific expression. The question of how they mechanistically execute the complex task of spatiotemporal expression within the seminiferous epithelium in vivo, however, is not well-understood. One possibility is that the enhancers of similarly expressed genes may form a chromatin hub for coordinate regulation [42–44]. The *Acrv1* gene promoter model suggests that short promoters could be easily sequestered from the transcription machinery in somatic cells and that they are hardwired to assume open chromatin configuration in male germ cells, likely by recruiting pioneer TF. RNA Pol II pausing or other repressor mechanisms are expected to then play a role in further fine-tuning gene transcription to suit the precise spatiotemporal requirement within the seminiferous epithelium.

Finally, the 50-bp *Acrv1* enhancer identified here could be used for the development of Cre-deleter mouse lines for round spermatid-specific knockout of genes. One potential future application could be the use of the enhancer to deliver therapeutic proteins to haploid germ cells to ameliorate male infertility.

Supplementary data

Supplementary data are available at *BIOLRE* online.

Acknowledgment

We gratefully acknowledge the services of Maggie Ober of the Transgenic Core Facility, and N. Sherman of the Biomolecular Research Facility and W.M. Keck of the Biomedical Mass Spectrometry Laboratory of the University of Virginia. KKA and PC were recipients of postdoctoral fellowship from the Fogarty International Center.

Conflict of interest

The authors have declared that no conflict of interest exists.

References

1. Eddy EM. Male germ cell gene expression. *Recent Prog Horm Res* 2002; **57**:103–128.
2. Lee K, Haugen HS, Clegg CH, Braun RE. Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. *Proc Natl Acad Sci USA* 1995; **92**:12451–12455.
3. Zambrowicz BP, Harendza CJ, Zimmermann JW, Brinster RL, Palmiter RD. Analysis of the mouse protamine 1 promoter in transgenic mice. *Proc Natl Acad Sci USA* 1993; **90**:5071–5075.
4. Howard T, Balogh R, Overbeek P, Bernstein KE. Sperm-specific expression of angiotensin-converting enzyme (ACE) is mediated by a 91-base-pair promoter containing a CRE-like element. *Mol Cell Biol* 1993; **13**:18–27.
5. Iannello RC, Young JC, Sumarsono S, Tymms MJ, Kola I. Mouse testis Pdha-2 promoter upstream sequences confer tissue-and temporal-specific activity in transgenic mice. *Reprod Fertil Dev* 1994; **6**:599–604.
6. Widlak W, Markkula M, Krawczyk Z, Huhtaniemi I. Functional analysis of spermatocyte-specific hst70 gene promoter in transgenic mice. *Acta Biochim Pol* 1994; **41**:103–105.
7. Han S, Xie W, Kim SH, Yue L, DeJong J. A short core promoter drives expression of the ALF transcription factor in reproductive tissues of male and female mice. *Biol Reprod* 2004; **71**:933–941.
8. Zhang LP, Stroud JC, Walter CA, Adrian GS, McCarrey JR. A gene-specific promoter in transgenic mice directs testis-specific demethylation prior to transcriptional activation in vivo. *Biol Reprod* 1998; **59**:284–292.
9. Li S, Zhou W, Doglio L, Goldberg E. Transgenic mice demonstrate a testis-specific promoter for lactate dehydrogenase, LDHC. *J Biol Chem* 1998; **273**:31191–31194.
10. Reddi PP, Flickinger CJ, Herr JC. Round spermatid-specific transcription of the mouse SP-10 gene is mediated by a 294-base pair proximal promoter. *Biol Reprod* 1999; **61**:1256–1266.

11. Banerji J, Rusconi S, Schaffner W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 1981; **27**:299–308.
12. Verrijzer CP, Chen JL, Yokomori K, Tjian R. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* 1995; **81**:1115–1125.
13. Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* 2018; **19**:621–637.
14. Vo Ngoc L, Wang YL, Kassavetis GA, Kadonaga JT. The punctilious RNA polymerase II core promoter. *Genes Dev* 2017; **31**:1289–1301.
15. Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, Stark A. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 2015; **518**:556–559.
16. Levine M, Cattoglio C, Tjian R. Looping back to leap forward: transcription enters a new era. *Cell* 2014; **157**:13–25.
17. Acharya KK, Govind CK, Shore AN, Stoler MH, Reddi PP. Cis-requirement for the maintenance of round spermatid-specific transcription. *Dev Biol* 2006; **295**:781–790.
18. Reddi PP, Shore AN, Shapiro JA, Anderson A, Stoler MH, Acharya KK. Spermatid-specific promoter of the SP-10 gene functions as an insulator in somatic cells. *Dev Biol* 2003; **262**:173–182.
19. Abhyankar MM, Urekar C, Reddi PP. A novel CpG-free vertebrate insulator silences the testis-specific SP-10 gene in somatic tissues: role for TDP-43 in insulator function. *J Biol Chem* 2007; **282**:36143–36154.
20. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975; **98**:503–517.
21. Bellve AR, Cavicchia JC, Millette CF, O'Brien DA, Bhatnagar YM, Dym M. Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. *J Cell Biol* 1977; **74**:68–85.
22. Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 1983; **11**:1475–1489.
23. Bowen B, Steinberg J, Laemmli UK, Weintraub H. The detection of DNA-binding proteins by protein blotting. *Nucleic Acids Res* 1980; **8**:1–20.
24. Jia Y, Nagore L, Jarrett H. Southwestern blotting assay. *Methods Mol Biol* 2015; **1334**:85–99.
25. Reddi PP, Naaby-Hansen S, Aguolnik I, Tsai JY, Silver LM, Flickinger CJ, Herr JC. Complementary deoxyribonucleic acid cloning and characterization of mSP-10: The mouse homologue of human acrosomal protein SP-10. *Biol Reprod* 1995; **53**:873–881.
26. al-Shawi R, Burke J, Wallace H, Jones C, Harrison S, Buxton D, Maley S, Chandley A, Bishop JO. The herpes simplex virus type 1 thymidine kinase is expressed in the testes of transgenic mice under the control of a cryptic promoter. *Mol Cell Biol* 1991; **11**:4207–4216.
27. Bellve AR, Millette CF, Bhatnagar YM, O'Brien DA. Dissociation of the mouse testis and characterization of isolated spermatogenic cells. *J Histochem Cytochem* 1977; **25**:480–494.
28. Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, Moritz L, Sultan C, Gurczynski SJ, Moore BB, Tallquist MD, Li JZ et al. Comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-Seq. *Dev Cell* 2018; **46**:651–667. e10.
29. Akhtar W, de Jong J, Pindyurin AV, Pagie L, Meuleman W, de Ridder J, Berns A, Wessels LF, van Lohuizen M, van Steensel B. Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell* 2013; **154**:914–927.
30. Lalmansingh AS, Urekar CJ, Reddi PP. TDP-43 is a transcriptional repressor; the testis-specific mouse ACRV1 gene is a TDP-43 target in vivo. *J Biol Chem* 2011; **286**:10970–10982.
31. Jimeno-Gonzalez S, Ceballos-Chavez M, Reyes JC. A positioned +1 nucleosome enhances promoter-proximal pausing. *Nucleic Acids Res* 2015; **43**:3068–3078.
32. Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* 2005; **309**:626–630.
33. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 2011; **25**:2227–2241.
34. Segal E, Widom J. Poly (dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol* 2009; **19**:65–71.
35. Roumaud P, Hache J, Martin LJ. Expression profiles of Sox transcription factors within the postnatal rodent testes. *Mol Cell Biochem* 2018; **447**:175–187.
36. Bai S, Fu K, Yin H, Cui Y, Yue Q, Li W, Cheng L, Tan H, Liu X, Guo Y, Zhang Y, Xie J, et al. Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. *Development* 2018; **145**; DOI 10.1242/dev.164855.
37. Wang G, Ying Z, Jin X, Tu N, Zhang Y, Phillips M, Moskophidis D, Mivechi NF. Essential requirement for both hsf1 and hsf2 transcriptional activity in spermatogenesis and male fertility. *Genesis* 2004; **38**:66–80.
38. Falender AE, Freiman RN, Geles KG, Lo KC, Hwang K, Lamb DJ, Morris PL, Tjian R, Richards JS. Maintenance of spermatogenesis requires TAF4b, a gonad-specific subunit of TFIID. *Genes Dev* 2005; **19**:794–803.
39. Martianov I, Velt A, Davidson G, Choukrallah MA, Davidson I. TRF2 is recruited to the pre-initiation complex as a testis-specific subunit of TFIIA/ALF to promote haploid cell gene expression. *Sci Rep* 2016; **6**:32069.
40. Chen X, Hiller M, Sancak Y, Fuller MT. Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. *Science* 2005; **310**: 869–872.
41. Oyama T, Sasagawa S, Takeda S, Hess RA, Lieberman PM, Cheng EH, Hsieh JJ. Cleavage of TFIIA by Taspase1 activates TRF2-specified mammalian male germ cell programs. *Dev Cell* 2013; **27**:188–200.
42. de Laat W, Grosveld F. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* 2003; **11**:447–459.
43. Felsenfeld G, Dekker J. Genome architecture and expression. *Curr Opin Genet Dev* 2012; **22**:59–61.
44. Allahyar A, Vermeulen C, Bouwman BAM, Krijger PHL, Verstegen MJAM, Geeven G, van Kranenburg M, Pieterse M, Straver R, Haarhuis JHI, Jalink K, Teunissen H et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet* 2018; **50**: 1151–1160.