

# Mechanisms of translational repression of the *Smcp* mRNA in round spermatids

Danielle L Cullinane\*, Tamjid A Chowdhury\* and Kenneth C Kleene

Department of Biology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125-3393, USA

Correspondence should be addressed to K C Kleene; Email: Kenneth.kleene@umb.edu

\* (D L Cullinane and T A Chowdhury contributed equally to this work)

## Abstract

The protamine 1 (*Prm1*) and sperm mitochondria-associated, cysteine-rich protein (*Smcp*) mRNAs exemplify a widespread pattern of mRNA-specific regulation of mRNA translation in post-meiotic spermatogenic cells, spermatids. Both mRNAs are transcribed and initially stored in free-mRNPs in early spermatids, and translated on polysomes in late spermatids. In this study, we demonstrate that the 5' and 3'-UTRs and the 3' terminus of the *Smcp* 3'-UTR are required for normal repression of the *Smcp* mRNA in transgenic mice. RNA affinity chromatography and mass spectrometry sequencing identified Y-box protein 2 (YBX2/MSY2) as the major protein that interacts with the 3' terminus of the *Smcp* 3'-UTR and a Y-box recognition sequence, GCCACCU, in the translation control element that is necessary for *Prm1* mRNA repression. Depletion of YBX2 in *Ybx2*-null mice prematurely activates *Prm1* and *Smcp* mRNA translation in early spermatids. Fluorescent *in situ* hybridization reveals that the *Smcp* intron, the *Smcp* mRNA, and both *Smcp-Gfp* transgenic mRNAs are strongly concentrated in the chromatoid body, and that the *Ybx2*-null mutation does not eliminate the *Smcp* mRNA from the chromatoid body. This and previous findings suggest that the *Smcp* pre-mRNA is spliced and associates with YBX2 in the chromatoid body, and that repressed free-mRNPs are stored in the general cytoplasm. As YBX2 is the predominant protein in testis free-mRNPs, it likely represses many mRNAs in early spermatids. The mechanisms by which YBX2 represses the *Smcp* and *Prm1* mRNAs are relevant to reproductive medicine because mutations in the human *YBX2* gene correlate with abnormal protamine expression and male infertility. *Reproduction* (2015) **149** 43–54

## Introduction

The developmental regulation of translational of specific mRNAs is important in controlling protein expression during the development of post-meiotic, haploid spermatogenic cells, spermatids, because transcription in late spermatids ceases due to chromatin remodeling (reviewed in Meistrich *et al.* (2003) and Kleene (2013)). This phenomenon is illustrated by the protamine 1 (*Prm1*) and sperm-mitochondria cysteine-rich protein (*Smcp*) mRNAs, which are transcribed in round spermatids stored as translationally inactive messenger ribonucleoprotein particles (free-mRNPs) for 3 and 7 days, respectively, before translation begins in transcriptionally compromised elongating and elongated spermatids (Kleene 1989, 2013). The initial block to mRNA translation avoids deleterious effects of premature PRM1 expression on male fertility (Lee *et al.* 1995).

mRNA-specific translational regulation usually involves *cis*-elements which bind *trans*-factors, RNA-binding proteins (RBPs) or small noncoding RNAs, which activate or repress translation (reviewed in Jackson *et al.* (2010)). About 20 RBPs and several miRNAs have been

reported to be implicated in translational control in spermatids by studies of knockout mice and overexpression of RBPs (Kleene 2013, Kotaja 2014). However, RBPs and miRNAs often interact with huge numbers of mRNA targets and are expressed for prolonged periods in spermatogenic cells. This complexity creates uncertainties whether the effects of individual factors on target mRNAs are direct or indirect and whether the factor produces major or minor changes in translational activity (Kleene 2013). The functions of factors in translational regulation can be accurately defined by analyzing mutations in *cis*-elements that abrogate binding of factors. Unfortunately, this approach is rarely used because mutations are prone to negative results and the only system in which mutations can be studied in spermatids, transgenic mice, is expensive and laborious.

One of the atypical features of post-transcriptional gene expression in spermatids is the chromatoid body, a prominent, ~1 µm diameter, perinuclear RNP granule containing thousands of mRNA species, and a plethora of RBPs and noncoding RNAs with diverse functions in RNA biology (Meikar *et al.* 2014). Gene knockouts that

deplete chromatoid body constituents produce male infertility, suggesting that the chromatoid body coordinates regulatory processes that are necessary for spermatid development (Meikar *et al.* 2011). However, it is unclear whether the chromatoid body functions as a remodeling center, which regulates the interactions of RBPs and miRNAs with mRNAs, or a realization center, which carries out the consequences of these interactions. An especially relevant question here is whether translationally repressed free-mRNPs are formed and stored in the chromatoid body or formed in the chromatoid body and stored in the general cytoplasm (reviewed in Kleene & Cullinane (2011)).

At present, the *Prm1* and *Smcp* mRNAs are the only mRNAs in which multiple mutations have been analyzed in transgenic mice. A remarkable series of deletion and point mutations identify a 17-nucleotide (nt) translational control element (TCE) in the *Prm1* 3'-UTR immediately upstream of the AAUAAA polyadenylation signal as necessary for repression in step 7–9 spermatids (Zhong *et al.* 2001). However, the mechanism by which the TCE represses translation is unknown because a factor that binds the TCE has not been identified.

Previous studies using the GFP reporter suggest that the week-long repression of the *Smcp* mRNA in step 3–10 spermatids is regulated by multiple mechanisms involving both UTRs (Hawthorne *et al.* 2006, Bagarova *et al.* 2010). The *Smcp* 5'-UTR alone delays GFP expression until step 5, the *Smcp* 3'-UTR alone delays GFP expression until step 9, and a mutation in the *Smcp* 3'-UTR partially releases translational repression (Bagarova *et al.* 2010). Clearly, *cis*-elements and factors that repress the *Smcp* mRNA until step 11 spermatids remain to be identified.

This study continues our objectives of identifying the *cis*-elements and *trans*-factors that repress the *Smcp* mRNA. We have analyzed two new transgenes. The first transgene contains the *Smcp* 5'-UTR and 3'-UTR to test the proposition that both UTRs are necessary to delay translational activation until step 11. The second transgene mutates a conserved segment in the *Smcp* 3'-UTR downstream of the first AAUAAA polyadenylation signal (Chowdhury & Kleene 2012). The unusual position of this mutation is based on the evidence that elements which repress *Prm1* mRNA translation are located at the 3' terminus of the 3'-UTR, referred to as the 3T3U below (Giorgini *et al.* 2001, Zhong *et al.* 2001, RE Braun 2013, personal communication). We also use RNA affinity chromatography and mass spectrophotometry sequencing to demonstrate that Y-box protein 2 (YBX2/MSY2) binds the *Prm1* TCE and the *Smcp* 3T3U, and demonstrate that the *Ybx2*-null mutation results in premature activation of *Prm1* and *Smcp* mRNA translation. Finally, a highly sensitive fluorescent *in situ* hybridization protocol reveals that natural and transgenic mRNAs with widely different translational activities are concentrated in the chromatoid body.

## Materials and methods

### Animal research

Protocols for the maintenance and usage of mice in this study were reviewed and approved by the University of Massachusetts Boston IACUC, Assurance # A3383-01, and are in accordance with the 2011 NIH 'Guide for the Care and Use of Experimental Animals'. CD-1 and C57BL/6 mice were maintained on a 12-h light:12 h darkness cycle, provided with food and water *ad libitum*, and killed by CO<sub>2</sub> hypoxia.

### Construction of the *S<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* and *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>-mut2* transgenic mice

The *S<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* transgene was constructed from *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* and *S<sup>5</sup>G<sup>3</sup>G<sup>3</sup>* transgenes described previously (Hawthorne *et al.* 2006). Briefly, the *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* and *S<sup>5</sup>G<sup>3</sup>G<sup>3</sup>* transgenes were digested with *Bsr*I and *Afl*II, and the small fragment containing the *Smcp* 3'-UTR (*S<sup>3</sup>*) was ligated into the large fragment containing the *Smcp* 5'-UTR (*S<sup>5</sup>*) and EGFP-coding region (*G<sup>C</sup>*).

The *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>-mut2* transgene was constructed from the *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* and *G<sup>5</sup>G<sup>C</sup>G<sup>3</sup>* transgenes in several steps. A *Swa*I site was inserted upstream of the first *Smcp* poly(A) signal with an overlap extension PCR in the *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>* transgene (Higuchi *et al.* 1988). Next, the *Swa*I–*Afl*II fragment from the *G<sup>5</sup>G<sup>C</sup>G<sup>3</sup>* transgene was inserted into the *Swa*I–*Afl*II sites of *G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>*. Finally, the *Swa*I site was reversed to that of the natural *Smcp* 3'-UTR with a second round of overlap extension PCR. The sequences of both transgenes are presented in Supplementary data 1, see section on supplementary data given at the end of this article.

Transgenic mice were generated by pronuclear injection (University of Massachusetts Medical Center Transgenic and Knockout Animal Core Facility Worcester, MA, USA). The techniques for the breeding and genotyping of transgenic mice have been described previously (Bagarova *et al.* 2010).

### Analysis of GFP expression in the squashes of seminiferous tubules

The stage of GFP expression was analyzed in the squashes of living spermatogenic cells (Kotaja *et al.* 2004, Bagarova *et al.* 2010). Adult or immature mice were killed with CO<sub>2</sub> hypoxia, the testes were dissected out and the tunica albuginea was removed. The seminiferous tubules from adult testes were teased apart in PBS and visualized using a dissecting microscope and transillumination to identify tubule segments from adult mice of potential interest. The stages of spermatids were identified in squashes of 0.5 mm tubule segments by phase-contrast microscopy using an Olympus BX51 microscope equipped with a Plan Fluorite 100× phase objective (NA 1.3), 100 W mercury burner, and SPOT XPLORER monochrome camera, SPOT image processing software (Diagonistic Instruments, Sterling Heights, MI, USA). EGFP fluorescence was excited at 470 nm and emitted light at 525 nm was photographed at a manual setting of 3 s and  $\gamma = 1$ . ImageJ 1.45S was downloaded from the NIH (Bethesda, MD, USA) and used to quantify the levels of GFP expression in various cell types.

### Sucrose and Nycodenz gradient analysis of polysomal loading

The cytoplasmic extracts of 21 dpp prepubertal and adult testes were prepared by dissecting the testes of adult and 21-dpp mice as described earlier. The subsequent procedures were carried out at 4 °C. The testes were homogenized in 300 µl HNM buffer (20 mM HEPES, pH 7.4, 0.1 M NaCl, 3 mM MgCl<sub>2</sub>) containing 0.5% Triton X-100 and 1 unit/µl RNasin Plus (Promega Biotech; Kleene *et al.* 2010). The extract was centrifuged at 13 000 *g* for 2 min, and 250 µl of the supernatant was layered on either a 3.8 ml linear 15–40% sucrose gradient in HNM buffer (w/w) or a 3.8 ml of 20–60% (w/v) Nycodenz gradient prepared by layering 760 µl of 60, 50, 40, 30, and 20% Nycodenz (Accurate Scientific Chemical Corporation, Westbury, NY, USA) in HNM in polyallomer centrifuge tubes for the Beckman SW60 rotor. The sucrose gradients were centrifuged for 80 min at 125 000 *g*, and ~0.4 ml the fractions were collected onto 0.3 g guanidine thiocyanate (Sigma–Aldrich) and Nycodenz gradients were centrifuged for 24 h at 140 000 *g*, and ~0.2 ml fractions were collected onto 0.15 g guanidine thiocyanate. RNA was extracted from the pellet and fractions and the levels of the *Ldhc*, *Smcp*, and transgenic mRNAs in each fraction were determined by RT-qPCR as described previously (Bagarova *et al.* 2010, Kleene *et al.* 2010).

### RNA affinity chromatography

Briefly, 20 µg 5'-biotinylated RNA (Sigma–Aldrich, Co.) was heated to 70 °C for 5 min in 400 µl binding buffer (20 mM HEPES, 3 mM MgCl<sub>2</sub>, 40 mM KCl, and 1 mM dithiothreitol, pH 7.6) containing protease inhibitor cocktail (Roche 11836170001) and 5% glycerol, slowly cooled, and incubated with ~500 µg extract proteins and 5 µg *Escherichia coli* tRNA for 30 min at 25 °C. The reactions were treated with 2 µl of heparin (200 mg/ml) for 10 min and incubated with pre-washed streptavidin–agarose (Pierce 20347, Rockford, IL, USA) on a rotating disc for 2 h at 4 °C. After 5, 15 min, 1 ml washes with a binding buffer with protease inhibitors at 4 °C, the proteins were released by boiling for 5 min in 50 µl SDS sample buffer, separated by SDS–PAGE, and visualized by silver staining (Pierce 24600). The bands were excised and identified with trypsin digestion and mass spectrometry sequencing at the Taplin Mass Spectrometry Facility (Harvard Medical School, Boston, MA, USA).

### Analysis of *Prm1* and *Smcp* mRNA translation in *Ybx2*-null testis

Male and female mice bearing heterozygous WT and knockout alleles at the *Ybx2* gene (Yang *et al.* 2007) were mated and *Ybx2*<sup>+/+</sup>, *Ybx2*<sup>+/-</sup>, and *Ybx2*<sup>-/-</sup> progeny were identified in DNA extracted from tail biopsies (Bagarova *et al.* 2010). The WT and knockout *Ybx2* alleles were distinguished by PCR amplification using a reverse primer for both alleles, GCAGAACAGGATGGGTTGTT, and forward primers specific for the WT allele, GGAGGGGAAGGGGACATT, and the knock-out allele, GCAGAACAGGATGGGTTGTT. The samples were denatured at 94 °C for 2 min, amplified with 27 cycles

(94 °C 50 s, 57 °C 50 s, 72 °C 30 s) followed by 72 °C 7 min. The sequences of the primers and PCR program were kindly provided by Drs RM Schultz and S Medvedev.

### Immunocytochemistry and RNA-fluorescent *in situ* hybridization

Stage II–VI seminiferous tubule segments were identified by transillumination and dissected in DEPC-treated PBS mechanically dispersed in DEPC-treated 100 mM sucrose and fixed and spread as dried-down preparations on slides dipped in 0.05% Triton X-100 and freshly prepared 4% formaldehyde (EM Sciences, Hatfield, PA, USA) (Kotaja *et al.* 2004).

For RNA-fluorescent *in situ* hybridization (RNA-FISH), 20-nt oligo probe sets for *Smcp* coding region, the *Smcp* intron, the *Gfp* coding region, and the *Ldhc* mRNA and introns were selected using the Stellaris Probe Designer at the Biosearch Technologies (Petaluma, CA, USA) website. The probe sets, consisting of 24–48 oligos, were fluorescently labeled with Quasar 570 or Quasar 670. The sequences of the probe sets and additional details of their selection are described in Supplementary data 2, see section on supplementary data given at the end of this article. The cells were permeabilized with 70% ethanol at room temperature (RT) for 1 h, rinsed twice in wash buffer (2× SSC, 10% deionized formamide (Ambion AM9342, Austin, TX, USA) for 5 min at 37 °C, and incubated overnight at 37 °C with a 1:50 dilution of each probe-set in a hybridization buffer (10% dextran sulfate (Sigma–Aldrich D8906) in wash buffer). After hybridization, the slides were washed 3× for 30 min at 37 °C in wash buffer, rinsed with DEPC-treated PBS, and mounted in Prolong Gold Antifade reagent with DAPI (Life Technologies P36931).

For immunocytochemistry, dried-down preparations were rehydrated in PBS, blocked with 10% normal goat serum (S-1000, Vector Laboratories, Youngstown, OH, USA) for 20 min at RT, washed with DEPC PBS, and incubated with rabbit polyclonal antibody to mouse vasa homologue, MVH (Abcam, Eugene, OR, USA, ab13840, 1:200, overnight at 4 °C), or affinity-purified rabbit polyclonal antibody to Y-box protein 2, YBX2 (Yu *et al.* 2003) (1:200, 1 h, RT), washed with PBS, reacted with goat anti-rabbit secondary antibody (Alexa Fluor 488, A11008) or 594 (A11037) (1:500, 30 min, RT), washed in PBS, mounted, and counterstained with DAPI as described earlier.

For combined RNA-FISH and immunocytochemical studies, RNA-FISH was carried out through the final washing in PBS, followed by blocking and reaction with primary and secondary antibodies as described earlier.

The cells were photographed using an Olympus BX51 microscope 100× plan fluorite objective equipped with Olympus filters, U-N31000 (excitation at 360 nm and emission at 460 nm) and U-N31004 (excitation at 560 nm and emission at 630 nm), or scanned at 0.8 µm using a Zeiss LSM 510 confocal microscope equipped with a Zeiss Plan-apo 63× NA oil objective. The wavelengths (nm) used for the confocal excitation and emissions of the fluoros in this study are as follows: DAPI, 405 and 460; Quasar 570, 547, and 570; Quasar 670, 644, and 670; Alexa fluor 488, 488, and 525; Alexa fluor 594, 594, and 617.

## Statistical analyses

Student's paired, two-sided *t*-test was used to calculate *P* values for pixel intensities measured with ImageJ (Instat3, GraphPad Software, La Jolla, CA, USA).

## Results

### Developmental expression of GFP fluorescence in $S^5G^CS^3$ and $G^5G^CS^3$ -mut transgenic spermatids

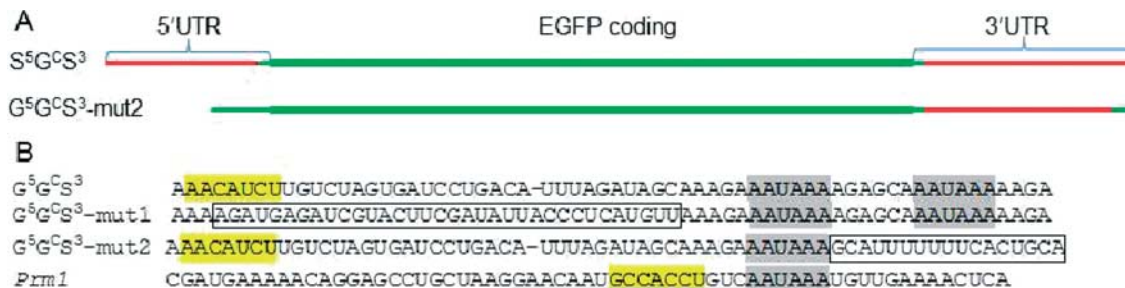
Two transgenes were studied to clarify the functions of the *Smcp* UTRs in the developmental regulation of translation in spermatids. Both transgenes contain 511 nt of *Smcp* 5' flanking region, which directs the transcription of *Smcp*-*Gfp* transgenic mRNAs in round spermatids at the same start site as the natural *Smcp* mRNA (Hawthorne *et al.* 2006). The segments of the transgenic mRNAs that are derived from *Smcp* and pEGFP are designated S and G respectively, and superscripts indicate the 5'-UTR, coding region and 3'-UTR. The first transgene,  $S^5G^CS^3$ , contains the complete *Smcp* 5'-UTR and 3'-UTRs. The second,  $G^5G^CS^3$ -mut2, contains the pEGFP 5'-UTR and the *Smcp* 3'-UTR, in which the 16 nt downstream of the first *Smcp* AAUAAA poly(A) signal has been replaced with the 17 nt downstream of the poly(A) signal in the pEGFP 3'-UTR. The full-length pEGFP 3'-UTR does not repress translation in round spermatids (Hawthorne *et al.* 2006, Bagarova *et al.* 2010). The structures of the transgenes and the sequences of the 3T3Us of the  $G^5G^CS^3$ ,  $G^5G^CS^3$ -mut1, and  $G^5G^CS^3$ -mut2 transgenes are depicted in Fig. 1, and their complete sequences are presented in Supplementary data 1.

The developmental expression of GFP fluorescence was analyzed with phase-contrast and fluorescence microscopy of squashes of short segments of living seminiferous tubules (Fig. 2). The morphology of squashed spermatogenic cells is clearly visualized using phase-contrast microscopy (Kotaja *et al.* 2004). The visualization of weak GFP fluorescence in  $G^5G^CS^3$ -mut2 round spermatids was facilitated by analyzing sexually

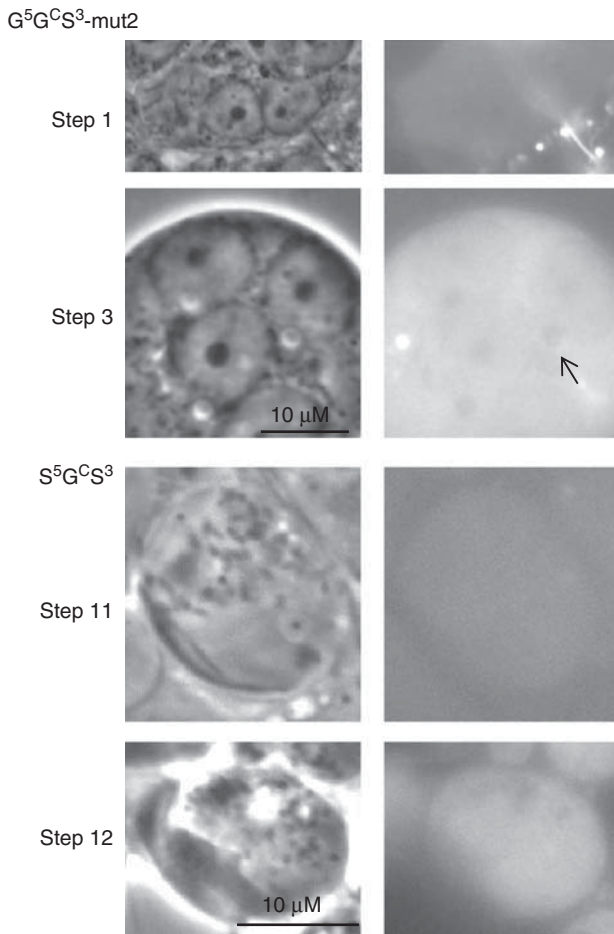
immature 25–28 day *post-partum* (dpp) testes which lack intensely fluorescent-elongated spermatids.

GFP fluorescence was obvious in the nuclei and cytoplasm of step 3 round  $G^5G^CS^3$ -mut2 spermatids and the cytoplasm of step 12  $S^5G^CS^3$  elongated spermatids (Fig. 2). The uniform GFP fluorescence throughout the nuclei and cytoplasm of step 3 spermatids and the cytoplasm of step 12 spermatids is distinct from the faint mottled, cytoplasmic background fluorescence in nontransgenic spermatids and spermatocytes (Supplementary Fig. 1, see section on supplementary data given at the end of this article). The exclusion of EGFP from the nuclei of step 12 spermatids and the acrosomes of round spermatids has been noted previously (Ventelä *et al.* 2000, Bagarova *et al.* 2010).

To document the developmental regulation of translational activity of the transgenic mRNAs more precisely, the intensities of GFP fluorescence of cell-free areas, spermatocytes, and spermatids were quantified using ImageJ. To reduce the effects of uneven illumination, the average pixel intensities of spermatids in close proximity to spermatocytes and cell-free areas were compared with Student's two-sided paired *t*-test (Supplementary Fig. 1 and Table 1, see section on supplementary data given at the end of this article). The pixel intensities of spermatocytes in various stages of meiosis were indistinguishable from background of cell-free areas ( $P \geq 0.26$ ). The pixel intensities of step 1  $G^5G^CS^3$ -mut2 and step 11  $S^5G^CS^3$  spermatids were on an average ~10% greater than those of spermatocytes or cell-free areas, and *t*-tests inconsistently produced *P* values <0.05. The pixel intensities of step 3  $G^5G^CS^3$ -mut2 and step 12 spermatids were ~1.5- to fourfold greater than those of spermatocytes and cell free areas and consistently produced *P* values <0.0001. Evidently, the initial repression of  $G^5G^CS^3$ -mut2 and  $S^5G^CS^3$  mRNAs is leaky followed by translational activation in step 3 and step 12 spermatids respectively. The marked increase in GFP fluorescence in step 3  $G^5G^CS^3$ -mut2 and step



**Figure 1** Structure and 3T3U sequences of transgenes. (A) The structures of the  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 transgenes are diagrammed. Segments derived from the pEGFP mRNA are green and segments derived from the *Smcp* mRNA are red. (B) Sequence of the 3' termini of natural and mutant *Smcp* and *Pzm1* 3'-UTRs. Y-box recognition sequences are highlighted yellow, canonical AAUAAA poly(A) signals are highlighted grey, and the 3' terminal A in each sequence is the poly(A) addition site that has been determined with 3'RACE or expressed sequence tags (Bagarova *et al.* 2010, Chowdhury & Kleene 2012). The boxed sequence in the  $G^5G^CS^3$ -mut1 transgene is randomized and eliminates the CAUC element that is essential for YBX2 binding (Bagarova *et al.* 2010). The boxed sequence in the  $G^5G^CS^3$ -mut2 transgene is derived from the early SV40 polyadenylation signal in the pEGFP plasmid (Kessler *et al.* 1986). The bold underlined sequence in the mouse *Pzm1* 3T3U is the translation control element (Zhong *et al.* 2001).



**Figure 2** Stage of first detection of GFP fluorescence in  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 transgenic round spermatids. Squashes of microdissected segments of seminiferous tubules were visualized by phase-contrast microscopy (left) to identify cell types and fluorescence microscopy (right) to monitor GFP expression. The contrast and brightness were adjusted to facilitate visualizing the morphology of spermatids and GFP fluorescence. The steps of spermatids were identified by cell associations and morphological criteria described by Russell *et al.* (1990). The arrow in the step 3 spermatid fluorescent cytoplasm points to an acrosome which appears dark because EGFP is excluded. The images demonstrate that dramatic increases in GFP expression occur in step 12  $S^5G^CS^3$  and step 3  $G^5G^CS^3$ -mut2 spermatids.

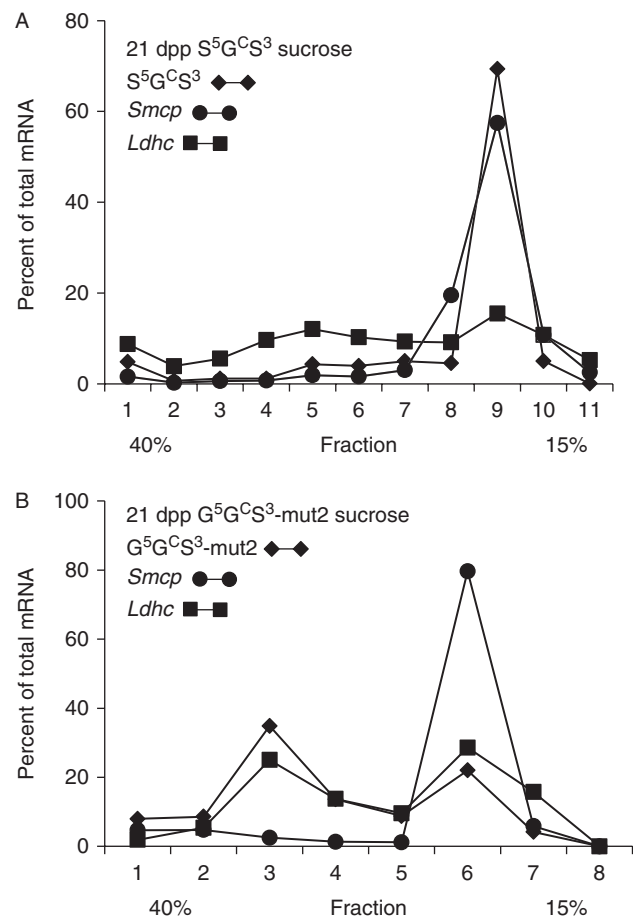
12  $S^5G^CS^3$  spermatids was documented in three and five independent lines and founders respectively.

### Sucrose gradient analysis of translational activity

To determine whether the differences in the developmental expression of GFP-fluorescence described above represent differences in translational activity, the translational activity of the  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 mRNAs was analyzed by sedimentation of cytoplasmic extracts from 21 dpp mice on sucrose gradients (Fig. 3). The gradients were collected as fractions including the pellets, the fractions were extracted with a protocol

that recovers equivalent proportions of RNA from each fraction (Kleene *et al.* 2010), and the levels of  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 mRNAs in each fraction were determined by RT-qPCR. The distribution of the transgenic mRNAs in each gradient was compared with those of the *Smcp* and *Ldhc* mRNAs. The *Ldhc* mRNA, which encodes the testis-specific isoform of lactate dehydrogenase, provides a control for mRNA recovery and polysome integrity because it exhibits constant polysome loading in 21 dpp and adult testes (Bagarova *et al.* 2010, Kleene *et al.* 2010).

The  $S^5G^CS^3$  and *Smcp* mRNAs displayed negligible levels of polysomal mRNA in sucrose gradients in 21 dpp testes, in which the most advanced cells are step 4



**Figure 3** Sucrose gradients analysis of the translational activity of the  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 mRNAs in 21 dpp sexually immature testes. The cytoplasmic extracts were sedimented on sucrose gradients, fractions were collected from the bottom, RNAs were extracted using techniques that recover virtually identical proportions of RNA from each fraction, and the amounts of mRNA in each fraction were determined with RT-qPCR. The results are presented as graphs of the percentage of total mRNA on the gradient in each fraction. The symbols that identify the *Ldhc*, *Smcp*,  $S^5G^CS^3$ , and  $G^5G^CS^3$ -mut2 mRNAs are represented in each panel. The gradients demonstrate that the  $S^5G^CS^3$  mRNA (A) is strongly repressed and that the  $G^5G^CS^3$ -mut2 mRNA (B) is active in step 1–4 spermatids.  $S^5G^CS^3$  mRNA (Part A) is strongly repressed and the  $G^5G^CS^3$ -mut2 mRNA (Part B) is active.

spermatids (Kleene *et al.* 2010). In contrast, the  $G^5G^CS^3$ -mut2 mRNA exhibits high polysomal loading in 21 dpp testes, consistent with active translation and GFP expression. These findings demonstrate that the differences in the efficiency of translational initiation control the stages of first detection of GFP-fluorescence of the  $S^5G^CS^3$  and  $G^5G^CS^3$ -mut2 mRNAs. These inferences are supported by Nycodenz-gradient analysis shown in Supplementary Fig. 2, see section on supplementary data given at the end of this article, a technique that separates free-mRNPs and polysomes by differences in buoyant density (Kleene *et al.* 2010). In addition, sucrose and Nycodenz gradient analysis demonstrate that all four mRNA species exhibit substantial levels of polysomal mRNA in adult testis (Supplementary Fig. 2), implying that the  $S^5G^CS^3$  and *Smcp* mRNAs are translationally active in elongated spermatids.

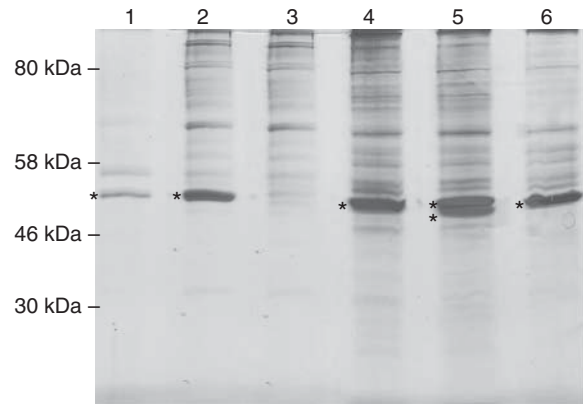
### RNA affinity chromatography of proteins binding to the 3' termini of the *Prm1*, *Smcp*, and $G^5G^CS^3$ -mut2 3'-UTRs

To identify proteins that bind the 3' termini of the *Smcp*,  $G^5G^CS^3$ -mut2, and *Prm1* 3'-UTRs, 5' biotinylated RNA probes were incubated with testis protein extracts, treated with heparin to reduce nonspecific binding by electrostatic interactions, and protein-RNA complexes were captured with streptavidin-agarose resin (Fig. 4). After extensive washing, the bound proteins were eluted in SDS sample buffer, resolved by SDS-PAGE, and unique bands were identified by mass-spec sequencing.

Lanes 2, 4, and 6 display a single, prominent band at ~52 kDa. Mass spec sequencing of these bands identifies two Y-box proteins, Y-box protein 2, YBX2/MSY2 (NP\_058571.2), and the long isoform of Y-box protein 3, YBX3L/MSY4 (NP\_620817.2). YBX2 and YBX3L have nearly identical calculated molecular weights, 38.0 and 38.8 kDa, respectively, but migrate at ~52 kDa due to the anomalous mobility of Y-box proteins (Davies *et al.* 2000, Mastrangelo & Kleene 2000, Skabkin *et al.* 2006). Spectral counts reveal that YBX2 is more abundant than YBX3L. In addition, the average precursor intensities of the four most abundant peptides demonstrate that YBX2 is nine- to 14-fold more abundant than YBX3L.

The levels of the ~52 kDa band are drastically reduced in the incubations of testis extract with streptavidin beads without biotinylated RNA (lane 1). Mass spec sequencing reveals that this protein is  $\beta$ -tubulin 4A (NP\_033477.2).

The segment of the *Smcp* 3'-UTR in lane 4 contains a Y-box 2 recognition sequence (YRS), AACAUCU, that has been analyzed with more than 40 single and multiple point mutations (Bouvet *et al.* 1995, Giorgini *et al.* 2001, Bagarova *et al.* 2010, Chowdhury & Kleene 2012). These studies define YBX2 binding site as a 7 nt element, [ACGU][AC][CA][UC][ACU]. Most permutations of alternative bases in brackets have little effect on YBX2 binding. The binding of YBX2 by the *Prm1* TCE, lane 2,



**Figure 4** Identification of proteins binding to the *Prm1*, *Smcp*, and  $G^5G^CS^3$ -mut2 3T3Us by RNA affinity chromatography and mass spectrometry sequencing. Total testis cytoplasmic extract was reacted with 5'-biotinylated RNA probes with heparin, bound to streptavidin-agarose, and eluted with SDS sample buffer. Proteins were resolved by SDS-PAGE, visualized with silver staining, and bands marked with asterisks were analyzed by mass spectrometry sequencing. YRSs are boxed, AAUAAA poly(A) signals are highlighted grey, and mutated sequences are double underlined. Lane 1, protein extract only; lane 2, WT *Prm1* TCE and poly(A) signal, GAACAAU[GCCACCU]GU-CAAUAAA; lane 3, mutated *Prm1* TCE and poly(A) signal, GAACAAUGACGACUGUCAAUAAA; lane 4, *Smcp* 3'-UTR YRS, AAAGGAUAGAAAACAUCUUGUCUAGUGAUCCUG; lane 5,  $G^5G^CS^3$ -mut2 3T3U, UGACAUUUAGAUAGCAAAGAAAUAAA-GCAUUUUUUUUCACUGC; lane 6, *Smcp* 3T3U, UGACAUUUAGAUAGCAAAGAAAUAAAAGAGCAAUAAA.

is eliminated by a GCCACCU → GCACGAU mutation, lane 3. This mutation has previously been demonstrated to abrogate YBX2 binding *in vitro* and translational repression by the *Prm1* YRS *in vivo* (Giorgini *et al.* 2001). The binding of YBX2 to the *Smcp* 3T3U in lane 6 is unexpected because it does not contain an YRS, suggesting that YBX2 binds sequences that are not recognizable in the 7 nt degenerate YRS above.

The probe for the 3' end of the  $G^5G^CS^3$ -mut2 3T3U binds two bands at ~51 and ~52 kDa (lane 5). The most abundant proteins in the 52 kDa band are YBX2 and YBX3L, but YBX2 is only 2.8-fold more abundant than YBX3L. The most abundant protein in the ~51 kDa band is the mouse homolog of the Lupus antigen (NP\_001103615.1), which binds oligo(U) sequences (Alfano *et al.* 2004), presumably the U-heptamer in the  $G^5G^CS^3$ -mut2 probe. Surprisingly, the bound proteins did not include ELAV1, another RBP with high affinity for U-rich elements and important functions in spermatogenesis (Chi *et al.* 2011).

The idea that YBX2 is the elusive factor that binds the *Prm1* TCE YRS, GCCACCU, is supported by u.v.-cross-linking assays shown in Supplementary Fig. 3, see section on supplementary data given at the end of this article. These assays demonstrate that YBX2 binds GCCACCU in assays in which the TCE probe is crosslinked with protein before treatment with RNase T1. The observation that digestion with RNase T1 before crosslinking prevents detection of the complex

with GCCACCU explains difficulties in the identification of YBX2 as the protein that binds the *Prm1* TCE.

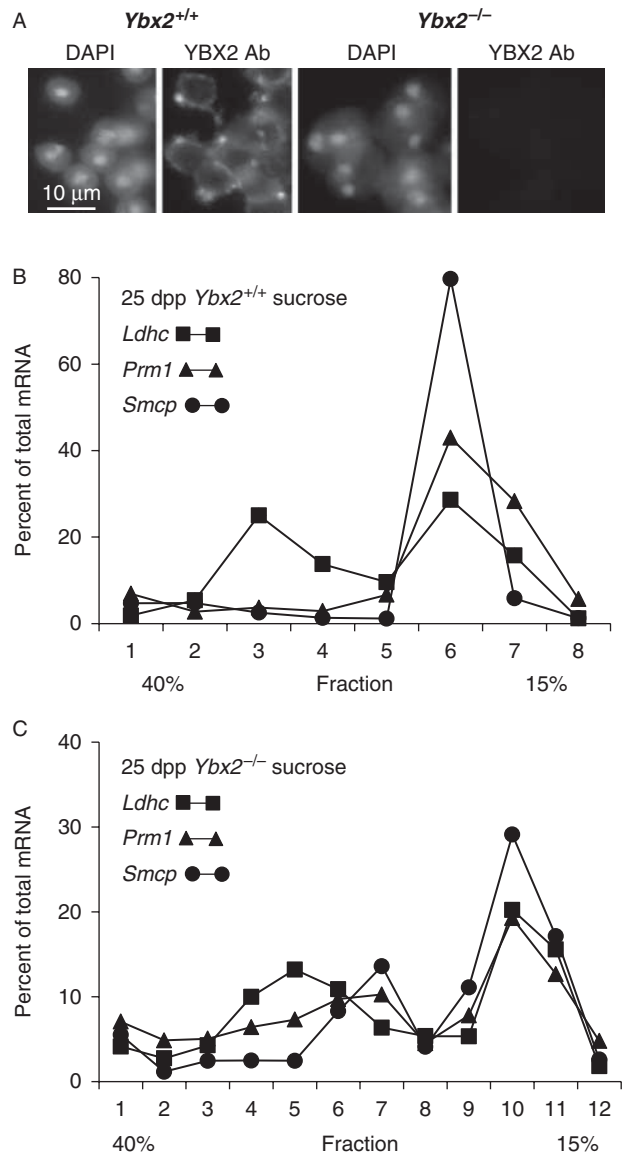
### Depletion of YBX2 results in premature recruitment of the *Prm1* and *Smcp* mRNAs onto polysomes in round spermatids

The evidence that YBX2 is the predominant protein that binds the *Smcp* 3T3U and *Prm1* TCE suggests that depletion of YBX2 with the *Ybx2*-gene knockout will prematurely activate *Prm1* and *Smcp* mRNA translation in round spermatids. Heterozygous *Ybx2*-knockout mice were mated (Yang *et al.* 2007), *Ybx2*<sup>-/-</sup> progeny were identified with PCR, and confirmed by the absence of YBX2 in immunocytochemistry in dried-down preparations of seminiferous tubules from 25 dpp *Ybx2*<sup>+/+</sup> and *Ybx2*<sup>-/-</sup> mice, Fig. 5A. DAPI-stained round spermatids are distinguished by round nuclei with bright central chromocenters. The YBX2-antibody stains the cytoplasm of *Ybx2*<sup>+/+</sup>-round spermatids with a higher concentration in a small perinuclear spot, the chromatoid body, in agreement with a previous immunohistochemical study of YBX2 in mouse testis (Oko *et al.* 1996). The YBX2-antibody did not stain *Ybx2*<sup>-/-</sup> round spermatids.

The translational activity of the *Prm1* and *Smcp* mRNAs was analyzed in sucrose gradients using the testis extracts of 25 dpp *Ybx2*<sup>+/+</sup> and *Ybx2*<sup>-/-</sup> mice, an age when the most advanced cells in the testis are step 9 spermatids and both mRNAs are strongly repressed (Braun *et al.* 1989, Fajardo *et al.* 1997, Kleene *et al.* 2010). Figure 5B reveals the absence of peaks of *Prm1* and *Smcp* mRNAs in small polysomes in *Ybx2*<sup>+/+</sup> 25 dpp testes, consistent with strong repression, and the presence of obvious peaks in small polysomes in *Ybx2*<sup>-/-</sup> testes (Fig. 5C), consistent with active translation. The high levels of polysomal *Ldhc* mRNA in both gradients validate the integrity of the polysomes and suggest that YBX2 is not an important repressor of the *Ldhc* mRNA.

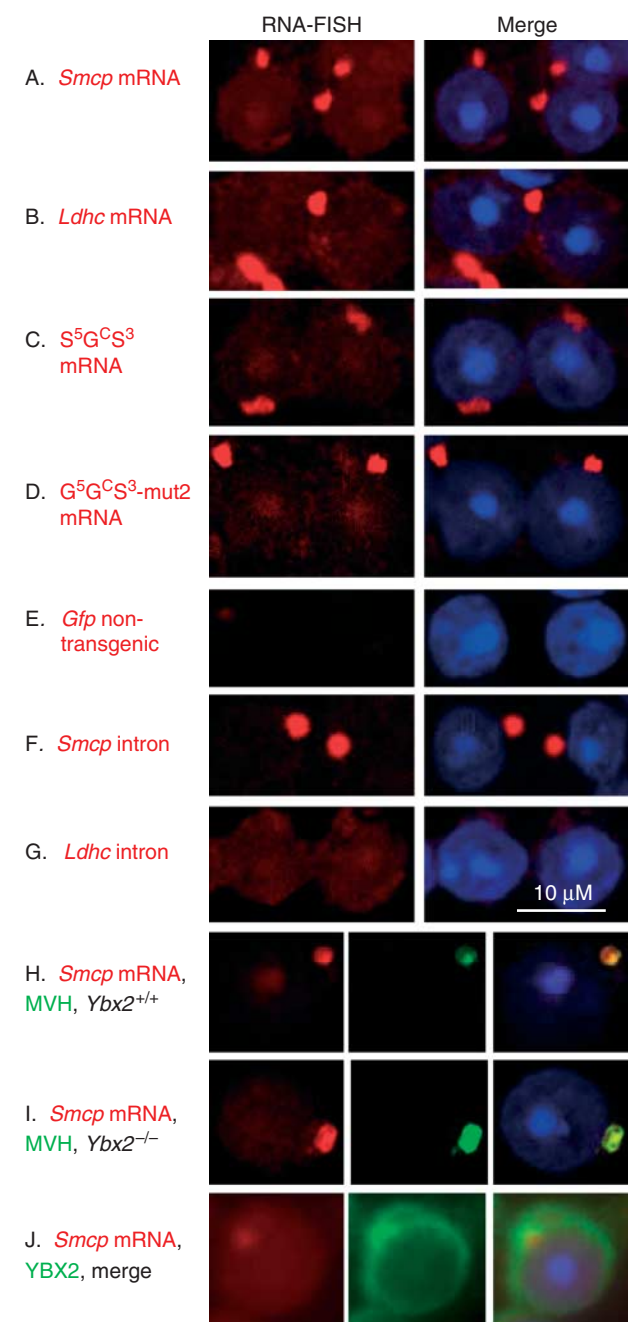
### Fluorescent in situ hybridization visualization of natural and transgenic mRNAs in chromatoid bodies in round spermatids

The dramatic differences in translational activity of the 5<sup>5</sup>G<sup>C</sup>S<sup>3</sup>, G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>-mut2, *Smcp*, and *Ldhc* mRNAs in round spermatids provide an opportunity to analyze the localization of mRNAs in the chromatoid body, an RNP-granule that has been postulated to function in storing translationally repressed mRNAs in round spermatids (reviewed in Kleene & Cullinane (2011)). We used an *in situ* hybridization protocol in which 24–48 tiled 20-base oligonucleotides specific for individual RNA species are 5' end labeled with fluorochromes (Raj *et al.* 2008). This protocol achieves high sensitivity by combining excellent probe penetration, extensive target coverage, and a high ratio of fluorochromes to bases.



**Figure 5** The *Ybx2*-null mutation prematurely recruits the *Prm1* and *Smcp* mRNAs onto polysomes in 25 dpp testes. (A) Dried down preparations of stage II–VI seminiferous tubules from 25 dpp *Ybx2*<sup>+/+</sup> and *Ybx2*<sup>-/-</sup> mice were stained with YBX2 antibody and DAPI. (B and C) Cytoplasmic extracts of *Ybx2*<sup>-/-</sup> and *Ybx2*<sup>+/+</sup> 25 dpp testes were sedimented on sucrose gradients, the gradients were fractionated and the relative levels of the *Prm1*, *Smcp*, and *Ldhc* mRNA in each fraction were analyzed by RT-qPCR. The symbols for the *Ldhc*, *Prm1*, and *Smcp* mRNAs are represented in B and C.

Confocal scanning microscopy with probes for the *Smcp* coding region and *Ldhc* mRNA (Fig. 6A and B) detected intense *in situ* hybridization signals in a ~1 μm diameter irregular perinuclear spot in dried down preparations from stage II–VI seminiferous tubules. The *Smcp* mRNA RNA-FISH co-localizes with immunocytochemical staining of mouse vasa homolog, MVH/DDX4 (Fig. 6H), an established marker for the chromatoid body (Yokota 2012, Meikar *et al.* 2014). The specificity of the 27 oligo



**Figure 6** RNA-FISH analysis of the localization of various mRNAs and introns in the chromatoid body. Probes for the *Smcp* coding region, *Gfp* coding region, *Ldhc* mRNA, *Smcp* intron, and *Ldhc* intron and mRNA were hybridized *in situ* to dried down cells from stage II–VI seminiferous tubules from adult testes. In some experiments, RNA-FISH was followed by immunocytochemistry with antibody to MVH/DDX4 or YBX2. The RNA-FISH staining was detected by laser scanning confocal microscopy (A, B, C, D, E, F, G, H and I) or fluorescence microscopy (J). The nuclei of DAPI-stained round spermatids are identified by their round shape and bright central chromocenter. Most of the probes strongly stain the chromatoid body, an irregular  $\sim 1 \mu\text{m}$  diameter organelle that is located adjacent to nuclei and is intensely stained with MVH/DDX4 antibody.

*Smcp* mRNA probe is supported by findings that the subsets consisting the 14 odd and 13 even numbered oligos exhibit the identical patterns of hybridization (A Ortajo 2012, personal communication). In addition, RNAseq shows that the *Ldhc* and *Smcp* mRNAs are abundant in purified chromatoid bodies (Meikar *et al.* 2014).

The intense RNA-FISH staining of the translationally active *Ldhc* mRNA and the repressed *Smcp* mRNA in the chromatoid body implies that localization in the chromatoid body is independent of translational activity. The repressed  $S^5G^CS^3$  and active  $G^5G^CS^3$ -mut2 mRNAs are also highly concentrated in the chromatoid body (Fig. 6C and D). The virtual absence of *Gfp* RNA-FISH signals in chromatoid bodies of nontransgenic mice demonstrate that the *Gfp* probe is specific for *Smcp*–*Gfp* transgenic mRNAs (Fig. 6E). The derepressed *Smcp* mRNA in *Ybx2*-null testis is also concentrated in the chromatoid body (Fig. 6I).

The probes for the *Smcp* and *Ldhc* introns were used with the goal of detecting pulsatile pre-mRNA transcription in round spermatids in the absence of the intense chromatoid body RNA-FISH signal (Raj *et al.* 2006). Unexpectedly, the *Smcp* intron probe strongly stained the chromatoid body (Fig. 6F), while the *Ldhc* intron probe did not (Fig. 6G). The implications of these observations are considered in the Discussion.

Careful inspection of confocal microscope images does not reveal the presence of cytoplasmic mRNA (Fig. 6A, B, C, D, H and I). The relatively strong RNA-FISH signals in the chromatoid body can be explained by the fact that the chromatoid body represents  $<0.4\%$  of the volume of the cytoplasm (Kleene & Cullinane 2011). This number can be used to calculate that if 5% of an mRNA was in chromatoid body and 95% of that mRNA was in the cytoplasm, the concentration of the mRNA would be 13-fold higher in the chromatoid body. The difference in mRNA concentration in the chromatoid body and cytoplasm is further exaggerated by  $0.8 \mu\text{m}$  confocal optical sections which include many mRNAs in a small object, the chromatoid body, and fewer mRNAs in a thick object, the cytoplasm. Figure 6J demonstrates that the *Smcp* mRNA is detectable in the thin layer of cytoplasm surrounding round spermatid nuclei by conventional fluorescence microscopy, and that the *Smcp* mRNA co-localizes with YBX2, a marker for cytoplasmic free-mRNPs (Oko *et al.* 1996, Yang *et al.* 2007). The *Smcp* mRNA signal in nuclei may represent high levels of *Smcp* mRNA or the extremely diverse transcriptome in round spermatids (Soumillon *et al.* 2013).

## Discussion

The evolutionarily unrelated *Prm1* and *Smcp* mRNAs are the only mRNAs in which translational repression in round spermatids has been analyzed with multiple mutations in transgenic mice. These studies reveal important differences and similarities in mechanisms.



The principle difference is that the 3-day repression of the *Prm1* mRNA is mediated entirely by the 3'-UTR TCE (Zhong *et al.* 2001), while the 8-day repression of the *Smcp* mRNA requires both the 5'-UTR and 3'-UTR. The 2-day repression by the *Smcp* 5'-UTR alone requires upstream reading frames which produce small polyosomes, whereas the 5-day repression by the *Smcp* 3'-UTR alone blocks initiation producing free-mRNPs (Bagarova *et al.* 2010). Additional mutations will be required to distinguish whether the 8-day repression by the *Smcp* 5'-UTR and 3'-UTR together is mediated by the upstream reading frames in the 5'-UTR and the block to initiation by the 3'-UTR, or interactions between the 3'-UTR and unidentified 5'-UTR elements that prolong the block to initiation.

The similarities are that the *Prm1* and *Smcp* 3'-UTRs direct repression in free-mRNPs (Braun *et al.* 1989, Zhong *et al.* 2001, Bagarova *et al.* 2010). Significantly, YBX2 binds the 3T3Us of the *Prm1* and *Smcp* mRNAs and both mRNAs are prematurely activated in round spermatids by the *Ybx2*-null mutation. The importance of YRS position in the 3T3U is demonstrated by transgenes in which YRSs 3 and 16 nt upstream of the *Prm1* AAUAAA poly(A) signal respectively repress translation strongly and partially (Fajardo *et al.* 1997, Giorgini *et al.* 2001, Zhong *et al.* 2001). In contrast, YRSs in the *Prm1* and *Smcp* 5'-UTRs or >34 nt upstream of the poly(A) signal in the 3'-UTR produce little or no repression (Zhong *et al.* 2001, Bagarova *et al.* 2010, RE Braun 2013, personal communication). The effects of YRS position imply that co-factors that bind the poly(A) signal, the poly(A) tail, or unidentified 3T3U elements strengthen repression by YBX2.

The importance of the 3T3U in mRNA repression in round spermatids is further supported by findings that repression is released by replacing the 16 nt downstream of first *Smcp* poly(A) signal with the 17 nt downstream of the pEGFP poly(A) signal. However, the pathway by which this mutation activates translation is unclear. The 23 nt segment extending from the AAUAAA poly(A) signal to the poly(A) site is known as the early SV40 polyadenylation signal (Kessler *et al.* 1986). As far as we are aware, the only proteins that bind this segment are CPSF30 and Wrd33, the canonical nuclear polyadenylation factors that bind AAUAAA (Chan *et al.* 2014). Therefore, the finding that the LA Lupus autoantigen (La) protein binds the G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>-mut2 3T3U was unexpected. The LA protein is necessary for early embryonic development (Park *et al.* 2006), binds oligo(U) sequences, and activates mRNA translation (Alfano *et al.* 2004), and its mRNA is expressed at high levels in testis (Carter & Sarnow 2000). The LA protein creates uncertainties whether the G<sup>5</sup>G<sup>C</sup>S<sup>3</sup>-mut2 3T3U blocks the assembly of a complex that represses translation or stimulates translation by another pathway.

All of the mRNA species studied here are concentrated in the chromatoid body in round spermatids, regardless

of whether they are strongly repressed or translationally active. The prematurely translated *Smcp* mRNA in *Ybx2*-null round spermatids is also concentrated in the chromatoid body. Evidently, sequestration in the chromatoid body is not sufficient for prolonged repression, and sucrose gradient analysis suggests that free-mRNPs and polyosomes are present in the general cytoplasm (reviewed in Kleene & Cullinane (2011)). The transitory localization of repressed mRNAs in the chromatoid body is clearly demonstrated by the first detection of the *Tnp2* and *Prm2* mRNAs simultaneously in the chromatoid body and the cytoplasm (Saunders *et al.* 1992, Fukuda *et al.* 2013).

The striking concentration of the *Smcp* intron in the chromatoid body was unexpected. However, noncanonical splicing in the cytoplasm is well-known (reviewed in Buckley *et al.* (2013)), and introns and splicing-factors are enriched in the chromatoid body (Biggiogera *et al.* 1990, Moussa *et al.* 1994, Meikar *et al.* 2014). The presence of the *Ldhc* mRNA and the absence of its introns in the chromatoid body are potentially related with the cessation of transcription of the *Ldhc* mRNA after meiosis (Tang & Goldberg 2012). One explanation is that the *Ldhc* mRNA is continuously present in germ cell RNP granules in spermatocytes and round spermatids. Another explanation is that the *Ldhc* mRNA is imported from the cytoplasm into the chromatoid body in round spermatids.

Two Y-box protein isoforms, YBX2 and YBX3L(MSY4), have been proposed to repress mRNA translation in round spermatids. The levels of both isoforms are maximal in round spermatids, followed by drastic decreases in elongating and elongated spermatids. These changes in levels correlate with the repression of many mRNAs in round spermatids and their subsequent activation in elongating and elongated spermatids (Oko *et al.* 1996, Davies *et al.* 2000). Our findings differ from previous ideas about the functions of YBX2 and YBX3L in mRNA repression in round spermatids.

YBX3L has been postulated to repress many mRNAs in round spermatids (Giorgini *et al.* 2002). However, YBX2 appears to be the critical isoform because the *Ybx2*-null mutation totally blocks spermatid development, whereas the *Ybx3*-null mutation has slight effects on spermatid development and male fertility (Lu *et al.* 2006, Yang *et al.* 2007). The more deleterious phenotype of the *Ybx2*-null mutation is likely related with observations that YBX2 is the predominant RBP in testis free-mRNPs (Herbert & Hecht 1999), and is much more abundant than YBX3L. YBX2 is also expected to bind more mRNAs because it binds single YRSs, whereas YBX3L forms a dimer with YBX2 that binds double YRSs (Davies *et al.* 2000). Single YRSs are much more numerous than double YRSs in repressed mRNAs in round spermatids (Chowdhury & Kleene 2012).

Another model proposes that YBX2 is a nonspecific RBP that selects pre-mRNAs for repression in the

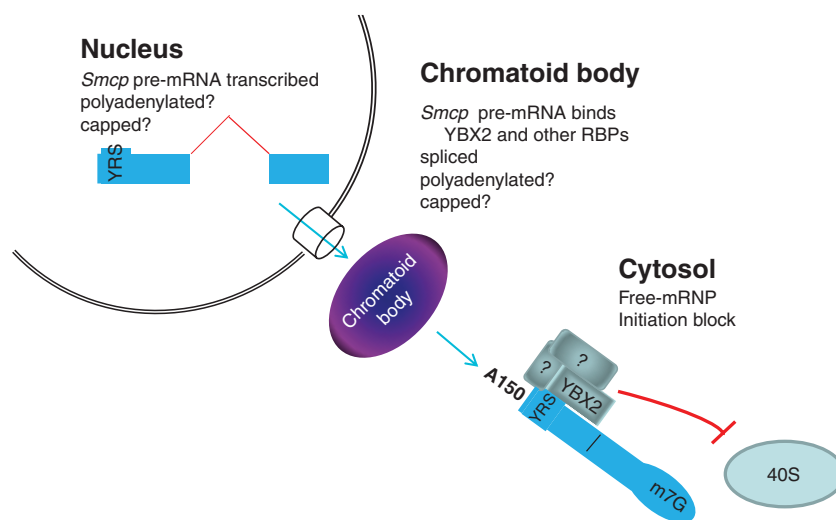
cytoplasm by binding Y-box promoter elements in the nucleus (Yang *et al.* 2005a,b). This creates a propensity for YBX2 to bind nascent mRNAs, and YBX2 remains bound to the mRNA and represses translation after the mRNA is exported to the cytoplasm. This model is incompatible with evidence that Y-box proteins bind specific RNA sequences (Bouvet *et al.* 1995, Giorgini *et al.* 2001, Chowdhury & Kleene 2012, Wei *et al.* 2012) and exhibit no specificity for Y-box promoter elements (Zasedateleva *et al.* 2002, Dolfini & Mantovani 2013). Significantly, mutations in the *Prm1* and *Smcp* mRNAs consistently demonstrate that mRNA sequences, not promoters, control repression in round spermatids (Braun *et al.* 1989, Schmidt *et al.* 1999, Zhong *et al.* 2001, Bagarova *et al.* 2010).

Understanding of the translational regulation of the *Smcp* mRNA is shown in Fig. 7. A bare-bones model maintains that the *Smcp* pre-mRNA is transcribed in the nucleus and exported through nuclear pores to the chromatoid body, where the pre-mRNA is spliced to produce the mature mRNA. YBX2 is predicted to bind 3T3U YRSs in the chromatoid body, because YBX2 is concentrated in the chromatoid body and undetectable in nuclei (Oko *et al.* 1996). The resulting mRNPs are exported to the cytoplasm and stored until the complex disassembles activating translation in step 12. The actual mechanism likely involves additional nuclear and cytoplasmic factors that facilitate the initial binding of YBX2 and its later dissociation from the YRS (Herbert & Hecht 1999, Zhong *et al.* 1999, Tsai-Morris *et al.* 2004, Matsumoto *et al.* 2005). Repression also likely requires factors that bind unidentified 3T3U elements and prevent binding of

free-mRNPs to the small ribosomal subunit. As transcription, splicing, polyadenylation, and capping are normally tightly coupled in the nucleus (Bentley 2014), splicing in the chromatoid body creates uncertainty whether the 7-methylguanosine 5' cap and 3' poly(A) tail are added to pre-mRNAs in the nucleus or the chromatoid body.

The mechanisms of polyadenylation are of special interest, because the poly(A) tails on translationally repressed *Prm1* and *Smcp* mRNAs in the spermatid cytoplasm are much longer than the poly(A) tails on active mRNAs, 150 As vs 30 As (Kleene 1989, Yang & Yen 2013). The unusual relationship between poly(A) length and translational activity in spermatids has been highlighted as a significant problem in eukaryotic post-transcriptional gene regulation (Jackson *et al.* 2010). It is unknown whether the length of poly(A) tails on repressed mRNAs is established by poly(A) lengthening and shortening in the nucleus, chromatoid body, or cytoplasm (Kim & Richter 2006).

Elucidating the mechanisms by which YBX2 and 3T3U YRSs repress translation in round spermatids in mice will have broad significance. The mechanisms of translational regulation of the mouse *Prm1* mRNA are at least partly conserved in trout and chicken (Iatrou & Dixon 1977, Rhim *et al.* 1995). The function of YBX2 as a direct repressor of *Prm1* mRNA translation would be solidified by demonstrating that a point mutation in the TCE abrogates YBX2 binding and the strong repression in round spermatids. The analogous studies of the *Smcp* mRNA should begin with a search for YRSs, because the *Smcp* 3T3U lacks a recognizable YRS. This search should avoid RNase, because RNase T1 artifactually



**Figure 7** Speculative model of translational repression of the *Smcp* mRNA. The *Smcp* pre-mRNA is transcribed in the nucleus and exported to the chromatoid body where it is spliced. YBX2 likely binds *Smcp* 3T3U YRSs in the chromatoid body, because YBX2 levels are highest in the chromatoid body and undetectable in nuclei (Oko *et al.* 1996). The complex of *Smcp* mRNA and YBX2 and other unidentified proteins transits to the cytoplasm and is stored as a repressed free-mRNP with 150-nt long poly(A) tail (A150) until dissociation of the complex enables translational activation and recruitment onto polysomes. The requirement of position of the YRS in the 3T3U for strong repression implies that YBX2 interacts with one or more RBPs which bind other 3T3U elements. It is unknown whether the 150-nt long poly(A) tails that are present on many repressed mRNAs in spermatids are lengthened or shortened in the nucleus, the chromatoid body or the cytoplasm. It is also unknown whether the 7-methylguanosine cap (m7G) is added in the nucleus, chromatoid body, or cytoplasm.

prevents detection of complexes between YBX2 and the *Prm1* TCE (Supplementary Fig. 3). Studies on mouse models should also produce insights into the mechanisms by which mutations in the *Ybx2* gene cause abnormal protamine expression and human male infertility (Hammoud *et al.* 2009).

## Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/REP-14-0394>.

## Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this research.

## Funding

This research was supported by grant MCB-0642128 to K C Kleene from the National Science Foundation and University of Massachusetts Boston Doctoral Dissertation Grants to D L Cullinane and T A Chowdhury.

## Acknowledgements

The authors thank Drs R M Schultz and S Medvedev for the donation of *Ybx2*<sup>+/-</sup> mice, the YBX2 antibody, and PCR protocol to distinguish the knockout and WT *Ybx2* alleles. They also thank Dr Alexey Veraksa for advice on quantification by mass spectrometry and the use of the confocal microscope.

## References

- Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S & Conte MR 2004 Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nature Structural & Molecular Biology* **11** 323–329. (doi:10.1038/nsmb747)
- Bagarova J, Chowdhury TA, Kimura M & Kleene KC 2010 Identification of elements in the Smcp 5' and 3' UTR that repress translation and promote the formation of heavy inactive mRNPs in spermatids by analysis of mutations in transgenic mice. *Reproduction* **140** 853–864. (doi:10.1530/REP-10-0323)
- Bentley DL 2014 Coupling mRNA processing with transcription in time and space. *Nature Reviews. Genetics* **15** 163–175. (doi:10.1038/nrg3662)
- Biggiogera M, Fakan S, Leser G, Martin TE & Gordon J 1990 Immunoelectron microscopical visualization of ribonucleoproteins in the chromatoid body of mouse spermatids. *Molecular Reproduction and Development* **26** 150–158. (doi:10.1002/mrd.1080260209)
- Bouvet P, Matsumoto K & Wolffe AP 1995 Sequence-specific RNA recognition by the *Xenopus* Y-box proteins. An essential role for the cold shock domain. *Journal of Biological Chemistry* **270** 28297–28303. (doi:10.1074/jbc.270.47.28297)
- Braun RE, Peschon JJ, Behringer RR, Brinster RL & Palmiter RD 1989 Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes and Development* **3** 793–802. (doi:10.1101/gad.3.6.793)
- Buckley PT, Khaladkar M, Kim J & Eberwine J 2013 Cytoplasmic intron retention, function, splicing, and the sentinel RNA hypothesis. *Wiley Interdisciplinary Reviews. RNA* **5** 223–230. (doi:10.1002/wrna.1203)
- Carter MS & Sarnow P 2000 Distinct mRNAs that encode La autoantigen are differentially expressed and contain internal ribosome entry sites. *Journal of Biological Chemistry* **275** 28301–28307. (doi:10.1074/jbc.M004657200)
- Chan SL, Huppertz I, Yao C, Weng L, Moresco JJ, Yates JR 3rd, Ule J, Manley JL & Shi Y 2014 CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes and Development* **28** 2370–2380. (doi:10.1101/gad.250993.114)
- Chi MN, Auriol J, Jegou B, Kontoyiannis DL, Turner JM, de Rooij DG & Morello D 2011 The RNA-binding protein ELAVL1/HuR is essential for mouse spermatogenesis, acting both at meiotic and postmeiotic stages. *Molecular Biology of the Cell* **22** 2875–2885. (doi:10.1091/mbc.E11-03-0212)
- Chowdhury TA & Kleene KC 2012 Identification of potential regulatory elements in the 5' and 3' UTRs of 12 translationally regulated mRNAs in mammalian spermatids by comparative genomics. *Journal of Andrology* **33** 244–256. (doi:10.2164/jandrol.110.012492)
- Davies HG, Giorgini F, Fajardo MA & Braun RE 2000 A sequence-specific RNA binding complex expressed in murine germ cells contains MSY2 and MSY4. *Developmental Biology* **221** 87–100. (doi:10.1006/dbio.2000.9658)
- Dolfini D & Mantovani R 2013 YB-1 (YBX1) does not bind to Y/CCAAT boxes *in vivo*. *Oncogene* **32** 4189–4190. (doi:10.1038/onc.2012.521)
- Fajardo MA, Haugen HS, Clegg CH & Braun RE 1997 Separate elements in the 3' untranslated region of the mouse protamine 1 mRNA regulate translational repression and activation during murine spermatogenesis. *Developmental Biology* **191** 42–52. (doi:10.1006/dbio.1997.8705)
- Fukuda N, Fukuda T, Sinnamon J, Hernandez-Hernandez A, Izadi M, Raju CS, Czaplinski K & Percipalle P 2013 The transacting factor CBF-A/Hnnpab binds to the A2RE/RTS element of protamine 2 mRNA and contributes to its translational regulation during mouse spermatogenesis. *PLoS Genetics* **9** e1003858. (doi:10.1371/journal.pgen.1003858)
- Giorgini F, Davies HG & Braun RE 2001 MSY2 and MSY4 bind a conserved sequence in the 3' untranslated region of protamine 1 mRNA *in vitro* and *in vivo*. *Molecular and Cellular Biology* **21** 7010–7019. (doi:10.1128/MCB.21.20.7010-7019.2001)
- Giorgini F, Davies HG & Braun RE 2002 Translational repression by MSY4 inhibits spermatid differentiation in mice. *Development* **129** 3669–3679.
- Hammoud S, Emery BR, Dunn D, Weiss RB & Carrell DT 2009 Sequence alterations in the YBX2 gene are associated with male factor infertility. *Fertility and Sterility* **91** 1090–1095. (doi:10.1016/j.fertnstert.2008.01.009)
- Hawthorne SK, Busanelli RR & Kleene KC 2006 The 5' UTR and 3' UTR of the sperm mitochondria-associated cysteine-rich protein mRNA regulate translation in spermatids by multiple mechanisms in transgenic mice. *Developmental Biology* **297** 118–126. (doi:10.1016/j.ydbio.2006.04.468)
- Herbert TP & Hecht NB 1999 The mouse Y-box protein, MSY2, is associated with a kinase on non-polysomal mouse testicular mRNAs. *Nucleic Acids Research* **27** 1747–1753. (doi:10.1093/nar/27.7.1747)
- Higuchi R, Krummel B & Saiki RK 1988 A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Research* **16** 7351–7367. (doi:10.1093/nar/16.15.7351)
- Iatrou K & Dixon GH 1977 The distribution of poly(A)<sup>+</sup> and poly(A)<sup>-</sup> protamine messenger RNA sequences in the developing trout testis. *Cell* **10** 433–441. (doi:10.1016/0092-8674(77)90030-7)
- Jackson RJ, Hellen CU & Pestova TV 2010 The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews. Molecular Cell Biology* **11** 113–127. (doi:10.1038/nrm2838)
- Kessler MM, Beckendorf RC, Westhafer MA & Nordstrom JL 1986 Requirement of A-A-U-A-A and adjacent downstream sequences for SV40 early polyadenylation. *Nucleic Acids Research* **14** 4939–4952. (doi:10.1093/nar/14.12.4939)
- Kim JH & Richter JD 2006 Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Molecular Cell* **24** 173–183. (doi:10.1016/j.molcel.2006.08.016)
- Kleene KC 1989 Poly(A) shortening accompanies the activation of translation of five mRNAs during spermiogenesis in the mouse. *Development* **106** 367–373.

- Kleene KC** 2013 Connecting *cis*-elements and trans-factors with mechanisms of developmental regulation of mRNA translation in meiotic and haploid mammalian spermatogenic cells. *Reproduction* **146** R1–19. (doi:10.1530/REP-12-0362)
- Kleene KC & Cullinane DL** 2011 Maybe repressed mRNAs are not stored in the chromatoid body in mammalian spermatids. *Reproduction* **142** 383–388. (doi:10.1530/REP-11-0113)
- Kleene KC, Bagarova J, Hawthorne SK & Catado LM** 2010 Quantitative analysis of mRNA translation in mammalian spermatogenic cells with sucrose and Nycodenz gradients. *Reproductive Biology and Endocrinology* **8** 155. (doi:10.1186/1477-7827-8-155)
- Kotaja N** 2014 MicroRNAs and spermatogenesis. *Fertility and Sterility* **101** 1552–1562. (doi:10.1016/j.fertnstert.2014.04.025)
- Kotaja N, Kimmins S, Brancorsini S, Hentsch D, Vonesch JL, Davidson I, Parvinen M & Sassone-Corsi P** 2004 Preparation, isolation and characterization of stage-specific spermatogenic cells for cellular and molecular analysis. *Nature Methods* **1** 249–254. (doi:10.1038/nmeth1204-249)
- Lee K, Haugen HS, Clegg CH & Braun RE** 1995 Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. *PNAS* **92** 12451–12455. (doi:10.1073/pnas.92.26.12451)
- Lu ZH, Books JY & Ley TJ** 2006 Cold shock domain family members YB-1 and MSY4 share essential functions during murine embryogenesis. *Molecular and Cellular Biology* **26** 8410–8417. (doi:10.1128/MCB.01196-06)
- Mastrangelo MA & Kleene KC** 2000 Developmental expression of Y-box protein 1 mRNA and alternatively spliced Y-box protein 3 mRNAs in spermatogenic cells in mice. *Molecular Human Reproduction* **6** 779–788. (doi:10.1093/molehr/6.9.779)
- Matsumoto K, Tanaka KJ & Tsujimoto M** 2005 An acidic protein, YBAP1, mediates the release of YB-1 from mRNA and relieves the translational repression activity of YB-1. *Molecular and Cellular Biology* **25** 1779–1792. (doi:10.1128/MCB.25.5.1779-1792.2005)
- Meikar O, Da Ros M, Korhonen H & Kotaja N** 2011 Chromatoid body and small RNAs in male germ cells. *Reproduction* **142** 195–209. (doi:10.1530/REP-11-0057)
- Meikar O, Vagin VV, Chalmel F, Sostar K, Lardenois A, Hammell M, Jin Y, Da Ros M, Wasik KA, Toppari J et al.** 2014 An atlas of chromatoid body components. *RNA* **20** 483–495. (doi:10.1261/rna.043729.113)
- Meistrich ML, Mohapatra B, Shirley CR & Zhao M** 2003 Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* **111** 483–488. (doi:10.1007/s00412-002-0227-z)
- Moussa F, Oko R & Hermo L** 1994 The immunolocalization of small nuclear ribonucleoprotein particles in testicular cells during the cycle of the seminiferous epithelium of the adult rat. *Cell and Tissue Research* **278** 363–378. (doi:10.1007/BF00414179)
- Oko R, Korley R, Murray MT, Hecht NB & Hermo L** 1996 Germ cell-specific DNA and RNA binding proteins p48/52 are expressed at specific stages of male germ cell development and are present in the chromatoid body. *Molecular Reproduction and Development* **44** 1–13. (doi:10.1002/(SICI)1098-2795(199605)44:1<1::AID-MRD1>3.0.CO;2-S)
- Park JM, Kohn MJ, Bruinsma MW, Vech C, Intine RV, Fuhrmann S, Grinberg A, Mukherjee I, Love PE, Ko MS et al.** 2006 The multifunctional RNA-binding protein La is required for mouse development and for the establishment of embryonic stem cells. *Molecular and Cellular Biology* **26** 1445–1451. (doi:10.1128/MCB.26.4.1445-1451.2006)
- Raj A, Peskin CS, Tranchina D, Vargas DY & Tyagi S** 2006 Stochastic mRNA synthesis in mammalian cells. *PLoS Biology* **4** e309. (doi:10.1371/journal.pbio.0040309)
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A & Tyagi S** 2008 Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods* **5** 877–879. (doi:10.1038/nmeth.1253)
- Rhim JA, Connor W, Dixon GH, Harendza CJ, Evenson DP, Palmiter RD & Brinster RL** 1995 Expression of an avian protamine in transgenic mice disrupts chromatin structure in spermatozoa. *Biology of Reproduction* **52** 20–32. (doi:10.1095/biolreprod52.1.20)
- Russell LD, Ettlin RA, Sinha Hikim AP & Clegg ED** 1990 *Histological and Histopathological Evaluation of the Testis*, St Louis, MO: Cache River Press.
- Saunders PT, Millar MR, Maguire SM & Sharpe RM** 1992 Stage-specific expression of rat transition protein 2 mRNA and possible localization to the chromatoid body of step 7 spermatids by *in situ* hybridization using a nonradioactive riboprobe. *Molecular Reproduction and Development* **33** 385–391. (doi:10.1002/mrd.1080330404)
- Schmidt EE, Hanson ES & Capecchi MR** 1999 Sequence-independent assembly of spermatid mRNAs into messenger ribonucleoprotein particles. *Molecular and Cellular Biology* **19** 3904–3915.
- Skabkin MA, Liabin DN & Ovchinnikov LP** 2006 Nonspecific and specific interaction of Y-box binding protein 1 (YB-1) with mRNA and posttranscriptional regulation of protein synthesis in animal cells. *Molekuliarnaia Biologiia* **40** 620–633. (doi:10.1134/S0026893306040078)
- Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthès P, Kokkinaki M, Nef S, Gnirke A et al.** 2013 Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Reports* **3** 2179–2190. (doi:10.1016/j.celrep.2013.05.031)
- Tang H & Goldberg E** 2012 A-MYB (MYBL1) stimulates murine testis-specific Ldhc expression via the cAMP-responsive element (CRE) site. *Biology of Reproduction* **86** 30. (doi:10.1095/biolreprod.111.095661)
- Tsai-Morris CH, Sheng Y, Lee E, Lei KJ & Dufau ML** 2004 Gonadotropin-regulated testicular RNA helicase (GRTH/Ddx25) is essential for spermatid development and completion of spermatogenesis. *PNAS* **101** 6373–6378. (doi:10.1073/pnas.0401855101)
- Ventelä S, Okabe M, Tanaka H, Nishimune Y, Toppari J & Parvinen M** 2000 Expression of green fluorescent protein under  $\beta$ -actin promoter in living spermatogenic cells of the mouse: stage-specific regulation by FSH. *International Journal of Andrology* **23** 236–242. (doi:10.1046/j.1365-2605.2000.00237.x)
- Wei WJ, Mu SR, Heiner M, Fu X, Cao LJ, Gong XF, Bindereif A & Hui J** 2012 YB-1 binds to CAUC motifs and stimulates exon inclusion by enhancing the recruitment of U2AF to weak polypyrimidine tracts. *Nucleic Acids Research* **40** 8622–8636. (doi:10.1093/nar/gks579)
- Yang CK & Yen P** 2013 Differential translation of Dazap1 transcripts during spermatogenesis. *PLoS ONE* **8** e60873. (doi:10.1371/journal.pone.0060873)
- Yang J, Medvedev S, Reddi PP, Schultz RM & Hecht NB** 2005a The DNA/RNA-binding protein MSY2 marks specific transcripts for cytoplasmic storage in mouse male germ cells. *PNAS* **102** 1513–1518. (doi:10.1073/pnas.0404685102)
- Yang J, Medvedev S, Yu J, Tang LC, Agno JE, Matzuk MM, Schultz RM & Hecht NB** 2005b Absence of the DNA-/RNA-binding protein MSY2 results in male and female infertility. *PNAS* **102** 5755–5760. (doi:10.1073/pnas.0408718102)
- Yang J, Morales CR, Medvedev S, Schultz RM & Hecht NB** 2007 In the absence of the mouse DNA/RNA-binding protein MSY2, messenger RNA instability leads to spermatogenic arrest. *Biology of Reproduction* **76** 48–54. (doi:10.1095/biolreprod.106.055095)
- Yokota S** 2012 Nuage proteins: their localization in subcellular structures of spermatogenic cells as revealed by immunoelectron microscopy. *Histochemistry and Cell Biology* **138** 1–11. (doi:10.1007/s00418-012-0962-z)
- Yu J, Hecht NB & Schultz RM** 2003 Requirement for RNA-binding activity of MSY2 for cytoplasmic localization and retention in mouse oocytes. *Developmental Biology* **255** 249–262. (doi:10.1016/S0012-1606(02)00094-5)
- Zasedateleva OA, Krylov AS, Prokopenko DV, Skabkin MA, Ovchinnikov LP, Kolchinsky A & Mirzabekov AD** 2002 Specificity of mammalian Y-box binding protein p50 in interaction with ss and ds DNA analyzed with generic oligonucleotide microchip. *Journal of Molecular Biology* **324** 73–87. (doi:10.1016/S0022-2836(02)00937-3)
- Zhong J, Peters AH, Lee K & Braun RE** 1999 A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. *Nature Genetics* **22** 171–174. (doi:10.1038/9684)
- Zhong J, Peters AH, Kafer K & Braun RE** 2001 A highly conserved sequence essential for translational repression of the protamine 1 messenger rna in murine spermatids. *Biology of Reproduction* **64** 1784–1789. (doi:10.1095/biolreprod64.6.1784)

---

Received 2 August 2014

First decision 10 September 2014

Accepted 21 October 2014