VERSITA

CELLULAR & MOLECULAR BIOLOGY LETTERS
Volume 11, (2006) pp 80 – 89
http://www.cmbl.org.pl

DOI: 10.2478/s11658-006-0008-4

Received: 02 November 2005
Revised form accepted: 16 December 2005

© *2006 by the University of Wrocław, Poland*

THE IDENTIFICATION AND CHARACTERIZATION OF A TESTIS-SPECIFIC cDNA DURING SPERMATOGENESIS

YING CHEN[1], JIARUI HU[2], PING SONG[1] and WUMING GONG[1]
[1]Laboratory of Molecular Genetics and Developmental Biology, College of Life Science, Wuhan University, Wuhan 430072, China, [2]Department of Gynecology and obstetrics, Zhongnan Hospital, Wuhan University, Wuhan 430071, China

Abstract: Using bioinformatics and experimental validation, we obtained a cDNA (named s*rsf*) which was exclusively expressed in the mouse testes. RT-PCR analysis showed that s*rsf* mRNA was not expressed in the gonad during the sex determination period or during embryogenesis. In developing mouse testes, s*rsf* expression was first detected on post-natal day 10, reached its highest level on day 23, and then reduced to and remained at a moderate level throughout adulthood. *In situ* hybridization analysis demonstrated that *srsf* mRNA was expressed in pachytene spermatocytes and round spermatids in the testes. The predicted protein contains one RNA-binding domain (RBD) and a serine-arginine rich domain (RS), which are characterized by some splicing factors of SR family members. These findings indicate that *srsf* may play a role during spermatogenesis.

Key Words: Expression pattern, Spermatogenesis, Splicing factor

INTRODUCTION

Alternative mRNA splicing plays an important role in development and differentiation, but relatively little is known about the mechanisms of pre-mRNA splicing [1, 2]. Both constitutive and alternative splicing occur on spliceosomes,

* Corresponding author; e-mail: pingsongps@yahoo.com.cn, tel: 86-27-87663175, fax: 86-27-68752560.

Abbreviations used: A – type Aspermatogonia; In – intermediate spermatogonia; B – type B spermatogonia; PL – preleptotene spermatocytes; L – leptotene spermatocytes; Zs – zygotene spermatocytes; EPs – early pachytene spermatocytes; LPs – late pachytene spermatocytes; Ds – diplotene spermatocytes; MI – first meiosis; MII – second meiosis; Rs – round spermatids; Es – elongated spermatids; S16 – step 16 spermatids; E – early; L – late; SG – spermatogonia; SC – spermatocytes; ST – round spermatid.

complex particles composed of small nuclear ribonucleo proteins (snRNPs) and numerous non-snRNP proteins [3]. The most extensively studied non-snRNP splicing factors are the SR family proteins [4], which are characterized by one or two N-terminal RNA-binding domains (RBDs) and a C-terminal serine and arginine domain (RS), which can be extensively phosphorylated [5]. The C-terminal serine and arginine domain can mediate the sub-cellular localization of individual SR proteins and can also function as a splicing-activated module [6]. The SR proteins constitute a family of pre-mRNA splicing factors; about ten are currently known, such as ASF/SF2, SC35, and SRp20 [4-11]. Recent studies disclosed that SR proteins play critical roles in both constitutive and alternative pre-mRNA splicing, either as recruitment factors, bridging factors, enhancer factors or weakening factors [12]. In one study, a lack of ASF/SF2 was found to cause an accumulation of incompletely processed pre-mRNA and subsequently cell death [13]. Other RS-containing splicing factors like U2AF65, U2AF35, tra, tra-2, sx1 and SRm160/300 are structurally different, but roles have been proposed for them in splice-site recognition, enhancer binding and the promotion of a network of interactions whereby their activities are modulated by phosphorylation/dephosphorylation of the serines [11-13]. Recent studies showed that SRm160 is not only a co-activator of pre-mRNA splicing, but is also involved in mRNA export as part of an exon-junction complex [14-15].
Mammalian spermatogenesis is a complex process that leads to the formation of male gametes. For future study of the processes involved in germ cell development and maturation, we set about isolating genes which are expressed predominantly in the testes and ovaries. In this study, we screened a clone (4932702H24) from the Riken database (http://fantom.gsc.riken.jp/). This clone is exclusively expressed in mouse testes, and it encodes a putative SR protein. Based on its profile of expression and characteristics, this gene was designated *srsf* (Spermatogenesis Related Splicing Factor).

MATERIALS AND METHODS

Animals and general methods

The BalB/c mice used in this study were provided by the Animal Experimental Center of Disease Prevention and Control, of Hubei Province, China. All the animals were maintained, killed and dissected according to the guidelines of Wuhan University.

RNA isolation and cDNA synthesis

All of the mice RNA was isolated using guanidinium isothiocyanate extraction at each stage of testes and embryo development. 2 μg of RNA was reverse transcribed into the first strand cDNA using 3'-AP-olig-dT [12-18] and M-MLV reverse transcriptase (Promega). Pre-natal days were defined as the number of days following the identification of a vaginal plug (E0.5) and post-natal days were defined as the number of days after birth.

RT-PCR analysis

Reverse transcription PCR was used to amplify genes from different tissues of adult mice, from embryos at various stages of development, and from mouse testes at various stages of post-natal development. Touch down PCR was performed under the following conditions: 45 s at 94ºC, 45 s at 59ºC and 2 min at 72ºC for 5 cycles; 45 s at 94ºC, 45 s at 57ºC and 2 min at 72ºC for 5 cycles; 45 s at 94ºC, 45 s at 55ºC and 2 min at 72ºC for 24 cycles; and 7 min at 72ºC. Stratagems Eagle Eye software was used to compare band density. The primers were as follows: RTF and RTR for the *srsf* gene, BAF and BAR for the ß-actin, and TRF2-F and TRF2-R for the TRF2 gene. (BAF 5'-CCA TGT ACG TAG CCA TC-3'; BAR 5'-GTA CCA CCA GAC AGC A-3'; RTF 5'-GGA GCC CAC TGG CAG GTT TA-3' RTR; 5'-GAG CGG CGA GTC CGT GAT TG-3'; TRF2-F 5'-TCT AAA CTA CCC CAA TGG ATG C-3'; TRF2-R 5'-ACG GTG CTC AGG TGG AGA CTA A-3'.

***In situ* hybridization**

The cDNA frame probe was PCR-amplified and subcloned into the pGEM-T vector (Promega), and then used to transcribe DIG-labeled sense and antisense probes *in vitro*, respectively using SP6 or T7 RNA polymerase. The testes were cut into 12-μm sections. After fixation, hybridization and washing, the signal was detected with NTB-emulsion [16].

Bioinformatics analysis of *srsf*

The analysis of both the nucleotide sequence and the deduced amino acid sequence were done online, respectively using the NCBI Blast-N server and Ensembl Blast-X server. Multiple alignments were performed with CLUSTALW [17].

RESULTS

Cloning a fragment of *srsf* cDNA and analyzing the genomic structure

After a large-scale screening of the cDNA pool of multiple tissues from the Riken database, we obtained a testis-specific expressed clone (Riken ID: 4932702H24). After devising primers according to the ORF (open reading frame) and sequencing the PCR product, we found that the amplified cDNA fragment of about 435 bp was identical to the sequence in the database (Riken: 4932702H24) (Fig. 1).

BLAST-N searches against the Nr database showed no identity with any known gene in Genbank. When analyzed in ENSEMBL (http://www.ensembl.org), the cDNA was found to belong to GENSCAN 00000074293, locating to the mouse 4th chromosome E1 (141,875,120-141,877,671 bp).

Fig. 1. The cDNA and the deduced amino acid sequence of ***srsf***. The regions of the primers used are shaded. The proline-rich motif is shaded and underlined. * indicates the stop codon.

Analysis of predicted protein

The cDNA contains an open reading frame (ORF) of 435 bp. The protein deduced from this ORF contains 145 amino acid residues and has a molecular weight of 14.9 KD, and a PI (isoelectric point) of 12.24; thus, it is a basic

Fig. 2. The domain organization of the *srsf* protein. The primary sequences of the *srsf* RBD (A) and RS (B) domains were aligned with those from the various proteins of the SR family [18]. The identical residues in all the sequences are blue. The GenBank numbers are as follows: gi:3929402(D-B52), gi:3929384 (D-TRA2), gi:730773(H-SF2), gi:266992 (H-SC35), gi:267188(H-U2AF65), gi:48474242 (M-THO4), gi:136628(M-U2AF65), gi:6226906(E-U2AF65), gi:5702013(T-TCSR), gi:1708417 (Y-IF41). H – Human; D – *Drosophila*; E – *Caenorhabdites elegans*; M – *Musmuscle*; T – *Trypanosomacruzi*; Y – *yeast.*

protein. A protein motif search revealed that the putative protein has 1 N-glycosylation site, 4 protein kinase C-phosphorylation sites and a proline-rich region profile at the C-terminal (http://www.expasy.org/tools/scanprosite). BLAST-N against the Swissport database shows that the predicted protein of *srsf*, like the other SR family proteins, contains a typical RBD and RS domain. We performed the alignment of these two structures with some members of the SR protein family using software. The results are as follows. In terms of RBD, the identity within one protein which contains two or three RBDs is from 23.3% to 29.7%, the identity between the different proteins is from 7.6% to 38%, and the identity between *srsf* and the other SR proteins ranges from 15.9% to 25.1% (Fig. 2A). In terms of RS, the identity within one protein which contains two or three RS ranges from 21.8% to 61.8%, that between different proteins is from 25.8% to 61.8%, and the identity between *srsf* and *tra-2* (*Drosophila*) is the highest. Moreover, the conserved residues of this domain occur in *srsfs* such as serine and arginine.

RT-PCR analysis

To understand the expression pattern of *srsf* mRNA during embryonic development, total RNA extracted from embryos at different stages (between embryonic day 8.5 and 18.5) was subjected to RT-PCR analysis. As shown in Fig. 3A, no expression was detected. A variety of adult mouse tissues including

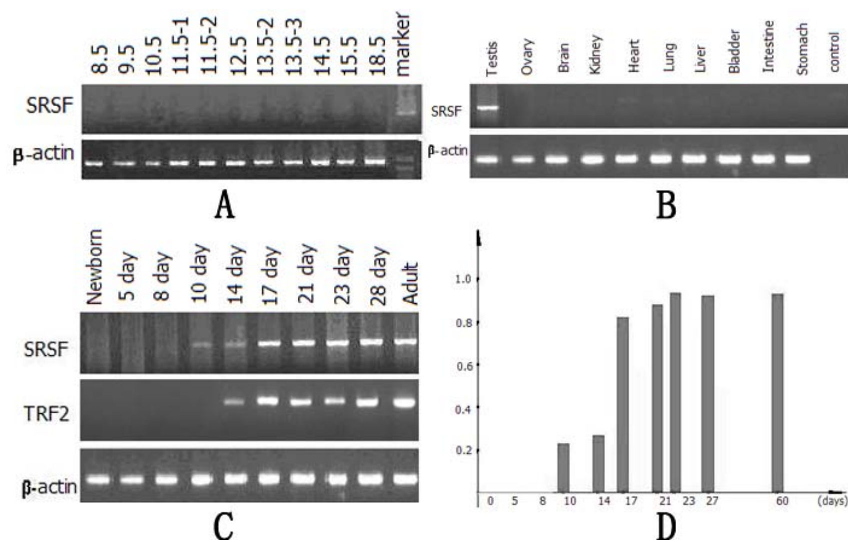

Fig. 3. The RT-PCR analysis of *srsf* mRNA expression. (A) The expression of *srsf* mRNA at different stages of embryonic development. (B) *srsf* mRNA is specifically expressed in the adult testes. (C) The expression of *srsf*, TRF2, and β-actin at different stages of testicular development. (D) A graphical representation of *srsf* mRNA comparative expression from RT-PCR data at different stages of testicular development, normalized to β-actin.

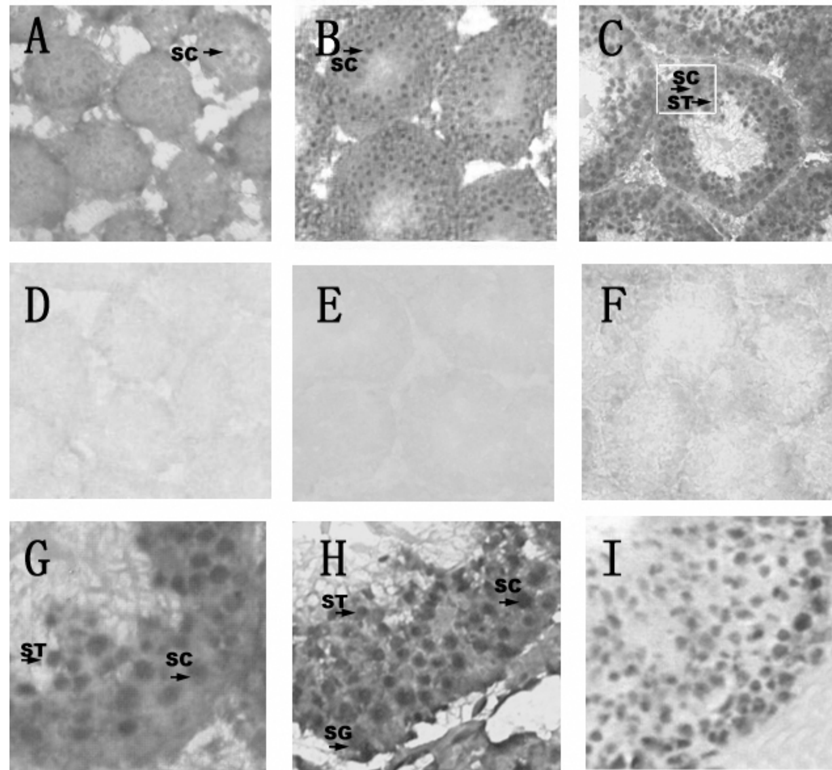

Fig. 4. The cellular location of srsf mRNA in the testes as detected via in situ hybridization. A, B, G, H - with an antisense probe. D, E, F - are control testes hybridized with a sense probe. H was hybridized with an mvh antisense probe and I was stained with hematoxylin. A, D - are 14-day old testes, B, E - are 21-day old testes, and C, F G, H, I - are adult testes. The black arrows mark the positive signals, while the pane indicates the amplification part. A-F×100; G-I×400.

brain, kidney, heart, lung, liver, bladder, intestine, stomach, testes and ovary were tested for srsf mRNA expression. As shown in Fig. 3B, srsf mRNA was exclusively detected in the testes.

We investigated *srsf* mRNA expression during the course of pubertal testes development to further explore its expression profile at this stage. The expression was first detected on post-natal day 10, increased abruptly on post-natal day 17, reached its highest level on day 23, and then reduced to and maintained a moderate level throughout adulthood (Fig. 3C and 3D). To clearly elucidate the *srsf* expression pattern, we also amplified *TRF2* [19], a TBP-related factor, the expression of which was confined to either pachytene spermatocytes or round spermatids during mouse spermatogenesis, in the same samples as in the control. There was a high expression of *TRF2* mRNA from post-natal day 17; this was similar to the *srsf* mRNA expression profile.

***In situ* hybridization**

To identify which cells express *srsf* mRNA in the mouse testes, we performed *in situ* hybridization on the testes taking *mvh* and hematoxylin as controls. In 2-week old testes, a faint positive signal was detected (Fig. 4A). In 3-week old testes, there was a stronger signal (Fig. 4B). However, on sections of adult testes (Fig. 4C and 4G), an intense signal was found to localize in the inner portion of the seminiferous tubules and to extend from the middle layer toward the lumen. Among the readily identifiable cells are pachytene spermatocytes and round spermatids, whereas other sperm cells, including the spermatogonia and elongating spermatids were all negative. As expected, a positive signal was detected in all the germ cells present in the testes when *mvh* was used as a probe [20] (Fig. 4H). The sense RNA probes of *srsf* labeled no cells in the testes (Fig. 4D, 4E and 4F).

DISCUSSION

In this study, we identified and characterized a cDNA encoding a novel protein which we named *srsf*. It contains, from the N-terminus to the C-terminus, one RBD and one RS domain. In the RS domain, there are two proline-rich regions; this is a multi-functional protein-protein interaction module that plays important roles in clustering proteins and organizing signal transduction [13]. The prolines in the proline-rich regions can be extensively phosphorylated, like serine and arginine [13]. Thus, we predict that *srsf* may be a splicing factor.

The characteristics of *srsf* suggest that this predicted protein is likely to function as an SR protein participating in the process of pre-mRNA splicing. Recent studies have also found that some SR proteins play crucial roles in proper sperm development. For example, the RNAi of both the *srp-4* and *srp-5* genes, which respectively encode the Srp-4 and Srp-5 proteins in *Caenorhabdites elegans,* caused a slight decrease in the sperm growth rate and some abnormal spermatogenesis [22]. Another example is tra-2, a testis-specific splice in *Drosophila.* Evidence shows that tra-2 has an essential role not only in controlling normal female sexual differentiation but also in normal spermatogenesis [23]. Considering the roles of the SR proteins in pre-mRNA splicing elucidated so far, we hypothesize that *srsf* may be responsible for the correct regulation, possibly at the level of splicing, of special genes which are important in germ cell development and maturation.

In our study, the initial expression of *srsf* mRNA in the testes was detected on post-natal day 10, when the first wave of spermatogenetic cells had entered meiosis I [24] (Fig. 5A). However, there was an abrupt increase in *srsf* mRNA expression on post-natal day 17, when the spermatocytes mature to the late pachytente stages, while meiosis is completed and early spermatids appear by post-natal day 21 [24]. The results of RT-PCR analysis in developing testes revealed that *srsf* mRNA was highly expressed in the late stages of meiosis, and showed the complex morphological changes of spermatogenesis (Fig. 5A, 5B).

This restricted profile of *srsf* expression was also confirmed by the analysis of *in situ* hybridization, which demonstrated that *srsf* mRNA was expressed in pachytene spermatocytes and round spermatids but not in germ cells at earlier stages (Fig. 4C, 4G). Thus, *srsf* participates in post-natal testis development.

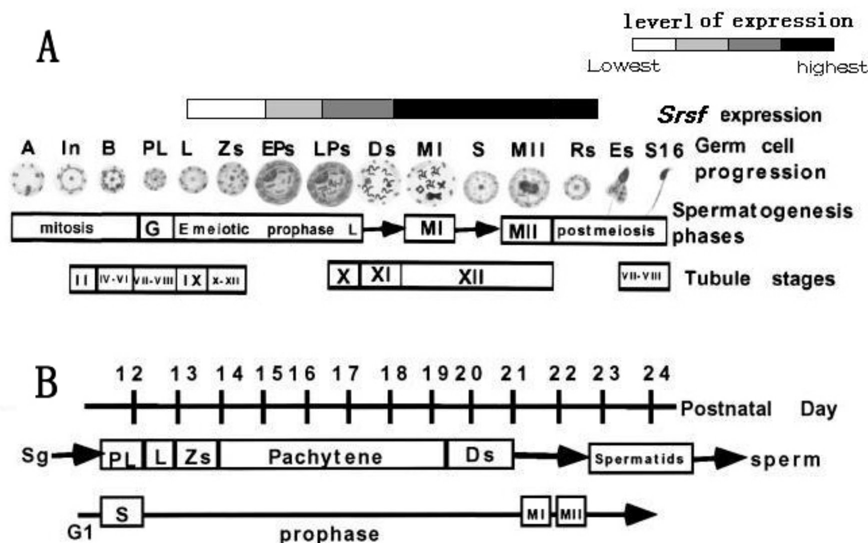

Fig. 5. The temporal and spatial expression of *srsf* during spermatogenesis is represented by the shaded boxes at the top. A - A diagram of germ cell progression, corresponding to the spermatogenesis phase. B - The timetable of the first wave of spermatogenesis, including the spermatogonia and all the prophase stages of germ cell differentiation [21].

Spermatogenesis is a complex process that requires specialized transcriptional regulation. The precise temporal-spatial expression of *srsf* mRNA suggests that this gene is subject to transcriptional regulation. Increasing evidence reveals that gene regulation mechanisms include a unique chromatin reorganization program and the use of distinct promoter elements and specific transcription factors during spermatogenesis [25]. Abundant studies show that various general transcription factors, such as TBP, TFIIB, and RNA polymeraseII, not only accumulate in much higher amounts in early haploid germ cells, but are also present in testis-specific isoforms [26]. Meanwhile, some meiosis-specific transcription factors were found to regulate the pachytene progression of male germ cells such as Ovol1 [27] in the mouse, and Ndt80 [28], the activity of which is essential for exiting the pachytene stage in yeast. Thus, it can be predicted that these specific transcription factors functionally cooperate to increase *srsf* mRNA expression at the late pachytene stage. Future studies on the expression and function of srsf protein would provide insights into the regulation of mammalian spermatogenesis.

In summary, our results show that *srsf* has a typical SR family structure, and that its mRNA is selectively expressed in the late pachytene spermatocytes and round

spermatids (Fig. 5A, 5B). The temporal and spatial expression of *srsf* reported in this study correlates well with the genetically defined role of *srsf* in male germ cell differentiation. Consequently, we suggest that *srsf* may play a role during spermatogenesis.

Acknowledgements. This study was supported by grants from the ICGEB (International Centre for Genetic Engineering and Biotechnology) (CRP/CHN02-01), National Nature Science Foundation of China (No. 30270675) and the National Basic Research Program of China (2004CB117400).

REFERENCES

1. Kleen, K.C. Patterns of transcriptional regulation in the mammalian testes. **Mol. Report Dev.** 423 (1996) 268-281.
2. Venables, J.P. Alternative splicing in the testes. **Curr. Opin. Genet. Dev.** 12 (2002) 615-619.
3. Fu, X.D. Specific commitment of different pre-mRNA to splicing by single SR proteins. **Nature** 365 (1993) 82-85.
4. Portal, D. and Joaqu´ın, M. An early ancestor in the evolution of splicing: a *Trypanosoma cruzi* serine-arginin-rich protein (TcSR) is functional in *cis*-splicing. **Mol. Biochem. Parasitol.**127 (2003) 37-46.
5. Roland, T. and James, L. Determinants of SR protein specificity. **Curr. Opin. Cell Biol.** 11 (1999) 358-362.
6. Krämer, A. The structure and function of proteins involved in mammalian pre-mRNA splicing. **Annu. Rev. Biochem.** 65 (1996) 367-409.
7. Fu, X.D. The superfamily of arginine/serine-rich splicing factors. **RNA** 1 (1995) 663-680.
8. Manley, J.L. and Tacke, R. SR proteins and splicing control. **Genes Dev.** 9 (1996) 284-293.
9. Valcárcel, J. and Green, M.R. The SR protein family: pleiotropic functions in pre-mRNA splicing. **Trends Biochem. Sci.** 21 (1996) 296-301.
10. Hastings, M.L. and Krainer, A.R. Splicing in the new millennium. **Curr. Opin. Cell Biol.** 13 (2001) 302-308.
11. Meissner, M., Lopato, S., Gotzmann, J., Sauermann, G. and Barta, A. Proto-oncoprotein TLS/FUS is associated to the nuclear mitrix and complexed with splicing factors PTB, SRm160, and SR proteins. **Exp. Cell Res.** 283 (2003) 184-195.
12. Zhu, J. and Krainer, A.R. Pre-mRNA splicing in the absence of a SR protein SR domain. **Genes** 14 (2000) 3166-3178.
13. Wang, J., Takegaki, Y. and Manliy, J.L. Targeted disruption of an essential vertebrate gene ASF/SF2 is required for cell viability. **Genes Dev.** 10 (1996) 2588-2599.
14. Stefan, W., Simion, C., Ivshina, M. and Nickerson, J.A. *In vitro* FRAP reveals the ATP-dependent nuclear mobilization of the exon junction complex protein SRm160. **J. Cell Biol.** 164 (2004) 843-850.

15. Stefan, W., Simion, C. and Nickerson, J.A. The spatial targeting and nuclear matrix binding domains of SR m160. **Proc. Natl. Acad. Sci. U.S.A**. 100 (2003) 3269-3274.
16. Dijkman, H.B.P.M., Mentzel, S, de Jong, A.S., Assmann, K.J.M. RNA in situ hybridization using digoxigenin-labeled cRNA probes. **Biochemica** 2(1995) 21-25
17. Thompson, J.D., Higgins, D.G., Gibson, and T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res**. 11 (1994) 4673-4680.
18. Birney, E., Kumar, S. and Krainer, A.R. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. **Nucleic Acids Res.** 21 (1993) 5803-5816.
19. Zhang, D., Penttila, T.L., Morris, P.L. and Roeder, R.G. Cell- and stage-specific high-level expression of TBP-related factor 2 (TRF2) during mouse spermatogenesis. **Mech. Dev.** 106 (2001) 203-205.
20. Ina, S., Tsunekawa, N., Nakamura, A. and Noce, T. Expression of the mouse Aven gene during spermatogenesis, analyzed by subtraction screening using Mvh-knockout mice. **Gene Expr. Patterns** 5 (2003) 635-638.
21. Bellve, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.N. and Dym, M. Spermatogenic cells of the prepuberal mouse: isolation and morphological characterization. **J. Cell Biol.** 74 (1977) 68-85.
22. Taizo, K. and Masaki, F.T. Unique and redundant functions of SR proteins, a conserved family of splicing factors, in *Caenorhabditis elegants* development. **Mech. Dev.** 95 (2000) 67-76.
23. Mattox, W., McGuffin, M.E. and Baker, B.S. A Negative feedback mechanism revealed by functional analysis of the alternative isoforms of the *Drosophila* splicing regulator transformer-2. **Genetics** 143 (1996) 303-314.
24. Rugh, R. The mouse, its reproduction and development. Burgess Publishing Company, Minneapolis. (1968).
25. Sassone-Corsi, P. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. **Science** 296 (2002) 2176-2178
26. Kimmins, S., Kotaja, N., Davidson, I. and Sassone-Corsi, P. Testis-specific transcription mechanisms promoting male germ-cell differentiation. **Repro**. 128 (2004) 5-12
27. Baoan, L., Mahalakshmi, N. and Mackay, D.R. Ovol1 regulates meiotic pachytene progression during spermatogenesis by repressing Id2 expression. **Development** 132 (2005)1463-1473
28. Tung, K-S. and Bilanchone, V. The pachytene checkpoint prevents accumulation and phosphorylation of the meiosis-specific transcription factor Ndt80. **Proc. Natl. Acad. Sci. U.S.A**. 97 (2000) 12187-12192.