African Journal of Biotechnology Vol. 7 (5), pp. 513-522, 4 March, 2008
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2008 Academic Journals

*Full Length Research Paper*

# Molecular cloning and functional characterisation of the human fertilin β(*FTN*-β) and *SPAM1* promoters

**Banu Az-Zubair M.K.**

Department of Molecular Biology and Bioinformatics, National Biotechnology Development Agency, Area 11, Garki-Abuja, Nigeria. E-mail: azzubair@nabda.gov.ng.

Accepted 4 October, 2008

**Among the molecules that play an essential role in the early steps of the process of fertilisation, and in the interaction of the sperm with the egg, are the sperm adhesion molecule 1 (SPAM1) and fertilin β(*FTN*-β). This interaction, mediated by various molecules on the gametes, starts with sperm-egg adhesion and ultimately results in the fusion of the membranes. The human fertilin β(*FTN*-β) and *SPAM1* promoters were cloned and characterised to determine their similarity with other testis germ cells-specific promoters, and to assess their functionality and tissue specificity. Various fragments of the promoters were fused with GFP in an expression cassette for an *in situ* transcription assay using four different cell types. Transcription initiation sites (TSS) were mapped by RACE at -127 nucleotide position relative to the initiation codon (ATG) in *SPAM1* and at -64 nucleotide position in *FTN*-β. Neither the *SPAM1* nor *FTN*-β promoter have consensus TATA or CCAAT boxes; however, they have the initiator (Inr) motif at the transcription initiation sites and the TSF (Testis Specific Factor) motif pair, one in either orientation. In addition *FTN*-β has a DPE (downstream promoter element) at +30 position relative to the transcription start site, while *SPAM1* has CRE (Cyclic-AMP Response Element) motifs at positions -28 and -70.**

**Key words:** Spermatogenesis, germ cells, promoter, transcription.

## INTRODUCTION

Perhaps the most crucial and essential of the early steps in the process of fertilisation is the interaction of the sperm with the egg. This interaction, which is mediated by various molecules on the gametes, starts with sperm-egg adhesion and results in the fusion of the membranes. Among the molecules involved in this essential role are the sperm adhesion molecule 1 (SPAM1) and fertilin β (*FTN*-β). Considerable evidence suggests that the sperm protein SPAM1, (or PH-20), located on both the plasma membrane and acrosomal membrane, is a sperm adhesion protein. SPAM1 has been shown to have a hyaluronidase activity, which is necessary for penetration of the cumulus cells (Myles and Primakoff, 1997) and zona pellucida (Nishimura et al., 2007). Furthermore, immunisation of the female guinea pig against purified PH-20 protein completely inhibits *in vivo* fertilisation (Dube, 2005). The human and mouse versions of the gene have been identified, human *SPAM1* maps to chromosome 7q31 (Day et al., 2002) and the mouse *Spam1* maps to the proximal region of chromosome 6 (Zhang, 2005).

Fertilin is a heterodimeric (subunits β and β) sperm plasma membrane protein. Fertilin is a member of the surface protein family called ADAMs, characterised by A Disintegrin and A Metalloprotease domain (Wolfsberg, 1995), of which approximately 30 are known. Several members of the ADAMs family are specifically, and some exclusively, expressed in the testis, including fertilin □ (*FTN*-β) [ADAM2], cyritestin [ADAM3, tMDCI], ADAM5 [tMDCII], ADAM6, ADAM16 [xMDC16], ADAM18 [tMDCIII], ADAM20, ADAM21, ADAM24 [testase 1], ADAM25 [testase 2], ADAM26 [testase 3], ADAM29 and ADAM30. Other ADAMs, such as fertilin □[ADAM1], are expressed on sperm and other tissues. Five of the testis-specific ADAMs, fertilin β, cyritestin, ADAM5, ADAM16 and ADAM18 are expressed on male germ cells and mature sperm (Wong, 2001). Fertilin β (ADAM2) functions in sperm-egg fusion by binding the sperm to the egg plasma membrane using the disintegrin domain (Rubinstein et al., 2006; Ziyyat et al., 2006). Knockout male mice

lacking either fertilin β (Nishimura et al., 2001) or cyritestin (Shamsadin et al., 1999) are severely infertile. Particularly, fertilin $\beta^{-/-}$ sperm showed reduced levels of binding to the egg plasma membrane. Human fertilin β has been mapped to chro-mosomal region 8p11.2 (Burkin, 1997), while the mouse orthologue is mapped to chromosome 14 (Nishimura et al., 2001).

Twenty percent of human couples are sub-fertile, with a male infertility factor responsible in half of the cases. Human male infertility may not be due to obvious sperm defects, as it is estimated that greater than one-third of male factor infertility has no clear aetiology (Dokemci, 2006). A significant percentage of patients, for whom *in vitro* fertilisation procedures had failed, are fertile when a single sperm is injected into the ooplasm, using a technique known as ICSI (intra-cytoplasmic sperm injection). Higher fertilisation rates are observed in patients treated with ICSI, compared to those who have sperm injected into the perivitelline space (|Harari, 1995). This suggests that sperm-egg interaction and penetration of the perivitelline space is critical for successful fertilisation, hence the possible relevance of fertilin β and *SPAM1* genes in male factor infertility. Both *SPAM1* and *FTN*-β are expressed only in post-meiotic round to elongating spermatid stages during spermatogenesis. Spermatogenesis is a cyclic, exceedingly complex and very efficient developmental process. It begins with the mitotic division of the spermatogonial stem cells, with the daughter cells remaining attached by a cytoplasmic bridge. These then give rise, by subsequent mitotic divisions, to a cohort of interconnected cells that undergo two meiotic divisions, resulting in haploid spermatids. A complex differentiation and transformation process called spermiogenesis is then begun by these postmeiotic cells to produce mature spermatozoa. These very complex processes during meiosis and spermiogenesis require a large variety of unique proteins not found in somatic cells or the mitotically dividing germ cells. Many of such proteins are expressed only in a highly regulated temporal and spatial pattern. Given the critical role of SPAM1 and fertilin β in fertilisation, analysis of the promoters of the genes encoding them and the *cis*-elements in the promoters that control their expression would be very important.

Gene expression is primarily regulated at the transcriptional level and the core promoter architecture plays the most crucial role. Nevertheless, two different segments of promoter structure and function have been described for RNA polymerase II transcription in eukaryotes: a core promoter region of around 50-400 nucleotides adjacent to the transcriptional start site and more distant promoter elements (Malik et al., 2007; Baek et al., 2006). In RNA polymerase II transcription the core promoter elements are defined as 'minimal DNA elements that are necessary and sufficient for accurate transcrip-tion in a reconstituted cell-free systems'. The two key genetic elements, which can function indepen-dently or synergistically, are: 1) the TATA element (consensus $TATA^{A}/_{T}A^{A}/_{T}$) located upstream of the transcription start site near -30 to -25- and/or 2) the initiator (Inr) element, a pyrimidine-rich sequence (consensus $PyPyA_{+1}N^{T}/_{A}PyPy$), located at the transcription start site (Smale, 1989). The transcription start site (TSS) and transcriptional factor binding motifs identify the 5' flanking region of human genes. Although the transcription start site marks the end of a promoter, critically important motifs are sometimes located further downstream in the 5' UTR. All the fragments used in this study include the 5' UTR as part of the constructs. Following sequence analysis and *in silico* prediction of potential transcriptional motifs, the functionality of the promoters and their differential cell-specific expression was determined by insertion into a GFP expression cassette and transfection into various cells, including germ cells, for *in situ* transcription assays using GFP as their reporter.

## MATERIALS AND METHODS

### PCR

PCR was done using 20 pmol of primer pair, 1.5 mM $Mg^{++}$, 2.5 mM dNTPs, 1 x Taq polymerase buffer, 1-2 units of Taq polymerase enzyme, 10-50 ng template DNA in 50 µl reaction volume made up with water. The Expand™ Long Template PCR System (Boehringer Mannheim) was used according to the manufacturer's protocols. Amplification was run for 40 cycles (94°C, 1 min; 55-72°C, 10 s-1 min [depending on the primers]; 72°C, 2 min).

### Cloning the 5' flanking regions of fertilin β (*FTN*-β) and *SPAM1* genes

Gene specific primers, nested within a 200 bp interval of the 5' end of the coding regions, were paired with *Alu* consensus primers to amplify the 5' flanking regions from appropriate genomic DNA templates, using the Expand™ Long Template PCR System according to the manufacturer's protocols (Boehringer Mannheim). The first PCR products were diluted 1/100 in distilled water and 5 µl used for nested PCR in the same conditions. Both the first and nested PCR products were examined on 1% agarose gel stained with ethidium bromide and visualised by UV in the MultiImage™ Light Cabinet (Alpha Innotech Corporation) and analysed with Chemi-Imager software and camera (Alpha Innotech Corporation). The nested PCR products were directly sub-cloned into pGEM-E vectors (Promega). The FTN-B 5' flanking region was isolated using the *Alu1*-FT-2R primer pair or *Alu1*-FT-40 primer pair, followed by a nested PCR using *Alu1*-FT-9 primer pair. The SPAM1 5' flanking region was isolated using *Alu1*-Sp1-5 and Spp1a and Spp2 (see Table 1).

### DNA Sequencing

Sequencing was performed using the dyedeoxy termination method with the Prism ready reaction terminator cycle sequencing kit (Applied Biosystems) on a Perkin Elmer Applied Biosystems Automated Sequencer, ABI 373A, and analyses were done with the manufacturer's installed software. Briefly, 500 ng of DNA template, 2.0 pmol of primer, and 9.5 µl of premix in 20 µl total volume were incubated for 30 cycles at 96°C for 30 s, 50°C 15 s and 60°C for 4 min. The product was purified and precipitated in ethanol and run on a 6% denaturing polyacrylamide gel for at least 16 h.

**Table 1.** The oligonucleotide primers sequence.

| Primer | Sequence | Annealing temp | Experiment |
|---|---|---|---|
| SP1.7-F | gat tca caa aga ggc agc att ttc atc aac | 72 | PCR |
| SP1.7-R | tag aag caa gat gac acc aca aag caa aca | 72 | PCR |
| Spp1a | agc tgc caa agc cag att gca gtt ctc cta | 72 | PCR |
| Spp1 | tct gct tct aca aac tga cag agg cac g | 72 | PCR |
| SPAM-1B | gag tgt tca aat aaa tgg atg ggt | 62 | PCR |
| SPAM-2B | atg ttc tgg tag acc tgt ttc g | 62 | PCR |
| SP-p3-5' | ttc ctc cat tca cag tta ctc ctg | 72 | 5'-RACE |
| SP-p2-5'-N1 | tcc agc cac cca atg tct ctt cag tag | 72 | 5'-RACE+Seq |
| SP-p1-5'-N2 | tgg ctg gtc tca atg tct tcc tag agc | 72 | 5'-RACE+Seq |
| FT-2F | gcg gat gga cag taa ttt tc | 62 | PCR |
| FT-2R | tgc att aaa ttc aca gta tat ggt | 62 | PCR |
| FTN-9 | ccc gct gag cag aaa caa gac | 62 | PCR |
| FTN-40 | gaa aat tac tgt cca tcc gca | 62 | PCR |
| FT-p2-5' | gta gga tgc ctg cga ttc aat tcc ttc c | 72 | 5'-RACE |
| FT-p1-5'-N | aaa att act gtc cat ccg cag ccc gcc | 72 | 5'-RACE+Seq |
| *Alu*-1 | tcc caa agt gct ggg att aca g | 55 | PCR |
| *Alu*-2 | ctg cac tcc agc ctg gg | 55 | PCR |
| *Alu*-3 | gat cgc gcc act gca ctc c | 55 | PCR |
| *Alu*-4 | gga tta cag gcg tga gcc ac | 55 | PCR |

Sequences of primers used for the molecular cloning of the 5' flanking regions of *FTN*-□and *SPMA1* genes, including those used for the RACE amplifications.

**Determination of transcription start site**

The transcription start site was determined with Rapid Amplification of cDNA Ends (RACE), using Clontech's human testis cDNA kit (Marathon-Ready™ cDNA, cat. No. 7414-1) according to the manufacturer's instructions. Briefly, internal gene-specific primers were paired with an adaptor primer (AP1) to amplify the 5' flanking region of the cDNA. The first PCR products were diluted 1/100 in distilled water and 5 µl used for nested PCR in the same conditions, using nested GSP and N-AP2 primers. The products were sub-cloned into pGEM-E vectors. Primers FT-p2-5' and AP1 were used for the first round FTN-B RACE PCR, followed by a nested PCR using FT-1-5'-N and N-AP2 primers. For the first round SPAM1 RACE PCR, primer pairs SP-p3-5' and AP1 were used, followed by a nested PCR using primers SP-p2-5'-N1 or SP-p1-5'-N2 with N-AP2.

**Tissue culture**

A human male liver adenocarcinoma (Hep-G2) cell line was obtained from American Type Culture Collection (ATCC). Joanne Davis and Dr. Paul Edwards (Department of Pathology, University of Cambridge) donated a human mammary cell adenocarcinoma (MCF-7) cell line. Hep G2 cells were maintained in Minimum Essential Medium (MEM) supplemented with 1% Penn-Strep, 1% glutamine, 1% Non-essential amino acids (NEAA), and 10% Foetal Calf Serum (FCS), at 37°C, 5% $CO_2$. MCF-7 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 1% Penn-Strep, 1% glutamine, 1% insulin and 10% FCS, at 37°C, 5% $CO_2$.

Mouse spermatogonial/spermatocytes (GC1) (CRL-2053) and spermatocytes/round spermatids (GC2) (CRL-2196) cells were also obtained from ATCC. The cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 1% Penn-Strep, 1% sodium pyruvate, and 10% Foetal Calf Serum (FCS), at 37°C, 5% $CO_2$. Subsequent molecular biological analysis by us, using RT-PCR, to test the differentiation stage of the cells indicated that the GC2 cells contain Pgk-2 and protamine transcripts (data not shown), which suggests that the cells contain post-meiotic to round spermatid stage cells. Furthermore, similar analysis of GC1 cells indicated that the GC1 cells are not equivalent to round spermatid stage cells, but that they are post-meiotic stage cells (data not shown).

**Deletion constructs**

The 1.4 kb 5' flanking region fertilin (*FTN*β) fragment was generated by *Alu*-GSP PCR as described above. The 1.4 kb fragment was subcloned in the GFP expression cassette, which was later used to generate the 1 kb fragment following digestion with *Bgl* II. The *SPAM1* 1.7 kb 5' flanking region fragment was generated by primers Spp1a and Spp2 and subcloned into a GFP expression cassette. The subcloned 1.7 kb plus reporter cassette was linearised with *Bgl* II restriction enzymes and divided into two aliquouts. One was digested with *Mbo* I to generate the 150 bp *SPAM1*-GFP cassette, and the other digested with *Hae* III to generate the 900 bp *SPAM1*-GFP cassette.

**Transfection and GFP assay**

DNA was prepared for transfection by purification through a Qiagen column, according to the manufacturer's protocols. A day before transfection, 1 x $10^5$ cells were seeded onto a sterile glass cover slip in each well of a 6-well plate. 1 µg of each construct was co-transfected with an equal amount of pGL-3 luciferase vector as the internal control, with Effectene™ as transfecting agent, according to the manufacturer's protocols. All experiments were performed in triplicate. Cells were generally prepared for reporter assays 48 h

post-transfection. Culture media were removed from each well and the cover-slips transferred into fresh labelled 6-well plates and washed three times with 1 x PBS, after which they were layered with 2 ml PBS/4% Paraformaldehyde for 30 min at room temperature. The PBS/4% Paraformaldehyde was then removed and the cells washed twice in 1 x PBS. The glass cover slips were removed and mounted on a glass slide for the analysis of the cells on a fluorescence microscope. The remaining cells in the culturing plates were used for the luciferase assays. Briefly, after the removal of culture media, the cells are washed once in 1 x PBS. The PBS was aspirated, 250 µl 1 x lysis buffer added per well, and the plates left at room temperature for 10 min. Cells were dislodged by scraping after the 10 min incubation, transferred into 1.5 ml Eppendorf tubes, and centrifuged for 10 s, at 12 000 x g. The supernatant was transferred into fresh Eppendorf tubes for analysis. To normalise for the reporter gene activity and the transfection efficiency, equal amounts of each GFP-construct cassette and luciferase pGL3 internal control were transfected into each cell type in triplicates as stated above. 48 h after transfections the cells were lysed and the supernatant assayed in triplicate using a luminometer. Means of the triplicate luciferase activity for each construct were calculated and the background luciferase activity subtracted from each experiment. The mean net luciferase activity from each transfection was then normalised to the net mean value of the GFP fluorescence activity.

## RESULTS

### Cloning the 5' flanking regions of *FTN*-β and *SPAM1* genes

To isolate the fertilin β and *SPAM1* 5' flanking regions, previously identified YAC genomic clones containing the *FTN*-β and *SPAM1* genes were amplified, using *Alu*-GSP primers (see Materials and Methods). A 1.4 kb fragment was obtained from the *FTN*-β clones (data not shown), which was then sequenced. The sequenced data indicated that the fragment contained the 5' end of the published coding region and the sequences upstream of the identified translation start codon (see Figure 1). Amplification of the *SPAM1* clones, using *Alu*-GSP primers gave a product of 2 kb fragment (data not shown), which was also sequenced. Similarly, the sequenced data indicated that it contained the 5' end of the coding sequence and sequences upstream of the identified translation start codon.

### Determination of transcription start site

To characterise the 5' flanking regions, the transcription start site was determined by RACE. The 5' RACE assays provided extended single band products (data not shown) that indicated that there was only one major transcription initiation site for the *FTN*-β gene. The products were sequenced and the data matched the 5' end of the *FTN*-β coding region (not shown), including the 5' UTR and the transcription start site. Following determination of the transcription start site for *FTN*-β, fragments immediately upstream of the translation start site were cloned by nested PCR from the genomic clones and sub-cloned into GFP expression cassettes.

Subsequent sequence analysis of the products indicated that there was a new first exon, and that the additional sequence indicated that the published initiation codon was in fact part of the second exon, not the first exon. Furthermore, alignment of cDNA products from RACE and genomic sequences in GenBank (gbAC004690.2) indicated that there is about 14 kb of intronic sequences between the actual first exon and the exon (data not shown) containing the transcriptional start site.

### Functional analysis of the *FTN*-β and *SPAM1* promoters

To determine the functionality of the promoter sequence, the *FTN*β 1.4 and 1 kb promoter fragments, and the *SPAM1* 1.7, 1 kb and 150 bp promoter fragments were cloned into a GFP reporter cassette. Each construct was separately transiently cotransfected with a luciferase pGL3 as transfection control in triplicates into GC1, GC2, HepG2 and MCF-7 cells. Furthermore, a transcription control promoterless GFP cassette was similarly transfected into the four cell types. Transcriptional activity driven by each of the promoter fragments were measured by intensity of the GFP fluorescence, relative to that measured from a promoterless GFP cassette transfected into similar cells. The cell lines supported transcriptional activities from the individual promoter constructs to varying degrees, while there was no expression of GFP in all the cells from control promoter-less construct. Furthermore, variations in relative transcriptional activities between fragments from the same promoter per cell type will indicate the presence or absence of important transcriptional motifs.

The GC-1spg (GC1 cells) represents an immortalised pachytene stage primary spermatocytes. The cells were isolated from primary murine pachytene spermatocytes transfected with SV40 Large T antigen (Peluso, 2006). Co-transfection of the SV40 Large T antigen and a temperature-sensitive variant of the p53 gene, under a constitutive promoter control, resulted in the establishment of a second immortalised germ cell line (GC2 cells). Identification of structures resembling proacrosomal granules and observation of haploid peaks, following immunohistochemical and flow cytometric studies respectively, led the researchers to conclude that the GC2 cells are equivalent to secondary spermatocytes to round spermatid stage cells (Kessler et al., 2000). Subsequent molecular biological analysis by us, using RT-PCR, to test the differentiation stage of the cells also indicated that the GC2 cells contain Pgk-2 and protamine transcripts (data not shown), this suggests that the cells contain post-meiotic to round spermatid stage cells. Furthermore, similar analysis of GC1 cells indicated that the GC1 cells do not contain round spermatid stage cells, but that they are post-meiotic stage cells (data not

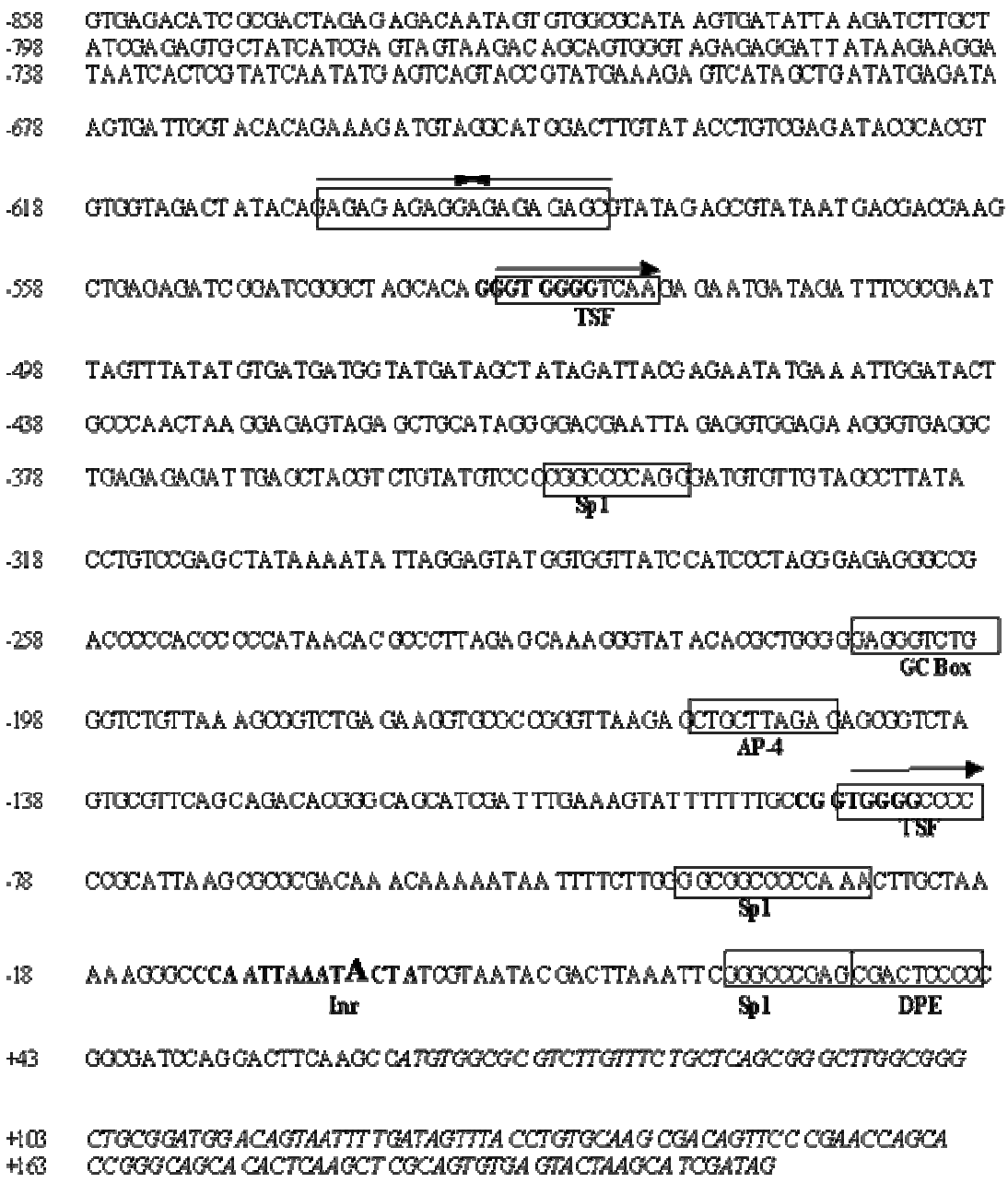


**Figure 1a.** *FTN-* core promoter. Human fertilin β promoter region, showing several transcription factors binding motifs. These include the Sp1, GC box, DPE (downstream promoter element) motifs, and a pair of TSF (testis-specific factor) motifs.

shown). HepG2 is a human male liver adenocarcinoma cell line and MCF-7 cells are human female mammary adenocarcinoma cells.

### The fertilin β promoter

The *FTN*-β promoter fragments were transcriptionally activated in GC1 and GC2 cells. This suggests that the *cis*-elements in the cloned promoter fragments are sufficient to accurately initiate transcription *in situ*. However, the promoter fragments were more active in the GC2 cells by order of a magnitude compared to activity in GC1 cells. There is very little difference in the levels of transcription between the two promoter fragments. This suggests that there is no transcriptionally important motif beyond the 1 kb fragment, at least in these *in situ* assays. Although fertilin β is not normally expressed in somatic

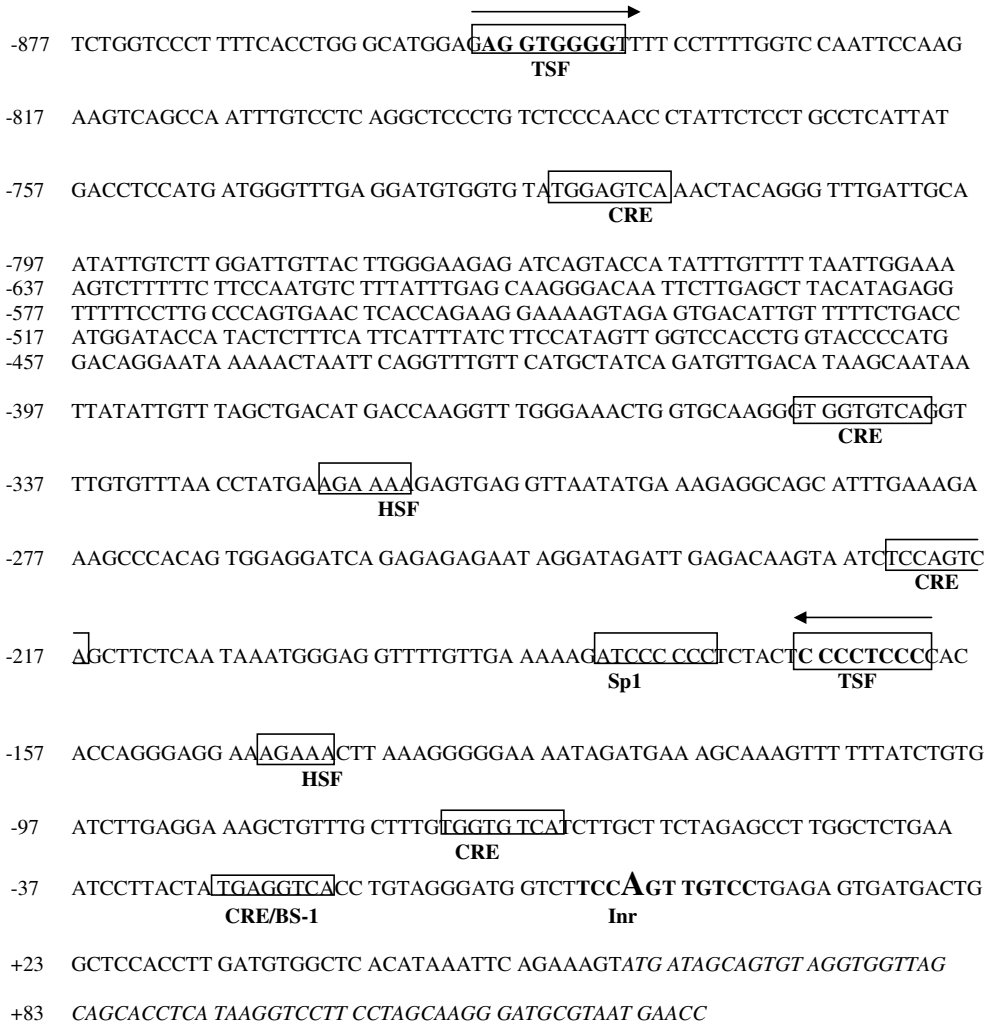


**Figure 1b.** *SPAM1* core promoter. Human *SPAM1* promoter region, showing various transcription factors binding elements. Prominent amongst these is the CRE (cyclic-AMP response element), HSF (heat shock factor), Sp1 and TSF (testis-specific factors) motifs.

tissues, considerable transcriptional activity was detected with the *FTN*β promoter fragments in HepG2 cells (see Figure 2 and Table 2). Interestingly, the relative activity of the *FTN*-β promoter fragments in HepG2 compared to that in GC1 cells is 1:4, thus the *FTN*β promoter had stronger transcriptional activity in the germ cells than in the somatic HepG2 cells. However, other fertilin genes are expressed in somatic tissues and germ cells, such as the fertilin β sub-unit, and many others. The *FTN*-β promoter fragments had a slight transcriptional activity in MCF-7 cells but, the activity did not compare to that seen in HepG2 cells and, compares even far less to that observed in the germ cells. Nevertheless, it does suggest that the *FTN*-β minimal promoter can be activated in somatic tissues. It is still important to note here that the promoter activity in the somatic tissues is not particularly significant, at least relative to that seen in germ cells.

## The *SPAM1* promoter

The *SPAM1* promoter fragments were transcriptionally active in both GC1 and GC2 cells, with a relative GC1: GC2 transcriptional ratio of 1:26 (see Figure 2 and Table 2). There was little difference in transcription among the three promoter fragments in GC1 cells; however, there was a slight increase in transcriptional activity from the 150 bp fragment in the GC2 cells. The *SPAM1* gene is not normally expressed in mitotic germ cells or somatic tissues, including the liver. The *SPAM1* promoter fragments were not transcriptionally activated in the HepG2 cells, which is consistent with what is observed *in vivo*. The levels of *SPAM1* promoter activities in HepG2 cell compares to that of the promoterless GFP control construct, which is background. The MCF-7 cells also, like the liver cells, do not naturally express the *SPAM1* gene,


**Figure 2.** An optimised comparison of the human *SPAM1* and mouse *Spam1* promoter regions. A comparison of the human *SPAM1* and mouse *Spam1* promoter regions. The transcriptional motif binding factors indicated are those for the mouse *Spam1* promoter, aligned with the approximate motif in the human *SPAM1* promoter. The mouse *Spam1* transcription initiation site is indicated as +1.

**Table 2.** The DPE sequence motifs from *Drosophila* to Man.

| Gene | DPE Sequence | Position |
|---|---|---|
| *Drosop Abd-B* | G G A C G A T | +30 |
| *Drosop Antp* | A G A C G T G | +30 |
| hGK-2 | A G A G C T G | +30 |
| hIRF | A G A C G T G | +30 |
| *Drosop joc* | G G A C G T G | +30 |
| mIRF | G G A C G T G | +30 |
| hGK-1 | G G A C G T G | +27 |
| hFTN-□ | G A G C G A C | +30 |

DPE sequence motifs from *Drosophila* to Man, indicating their location relative to the TSS. The *mIRF-1* and *hIRF-1* (Miyamoto et al., 1988; the *Drosophila Abdominal-B* ((DeLorenzi et al., 1988)), *Drosophila Antennapedia* P2 (Hsu et al., 2003) and *Drosophila jockey* (Hsu et al., 2003) as well as *hGK-1*, *hGK-2* (unpublished data) and *hFTN*β, are TATA-less promoters, which almost exclusively contain DPE motifs. Notice the perfect match between the *Drosop. joc*, *mIRF* and *hGK-1* on one hand, and *Drosop Antp*, *hIRF* and *hGK-2* on the other.

thus were not expected to support transcriptional activation of the gene. The *SPAM1* promoter fragments were not, as in HepG2, transcriptionally active in the MCF-7 cells.

## DISCUSSION

Until fairly recently, there were no immortalised germ cells with which to do *in situ* study of germ cell-specific functions. Most promoter studies were previously done *in vitro*, using testis whole cell nuclear extracts in run-off transcription assays. The first immortalised germ cells were made using mouse cells (Peluso, 2006). The GC-1spg (GC1 cells) represents an immortalised pachytene stage primary spermatocytes. The cells were isolated from primary murine pachytene spermatocytes transfected with SV40 Large T antigen (Peluso, 2006). The cells do not undergo differentiation in culture, or even

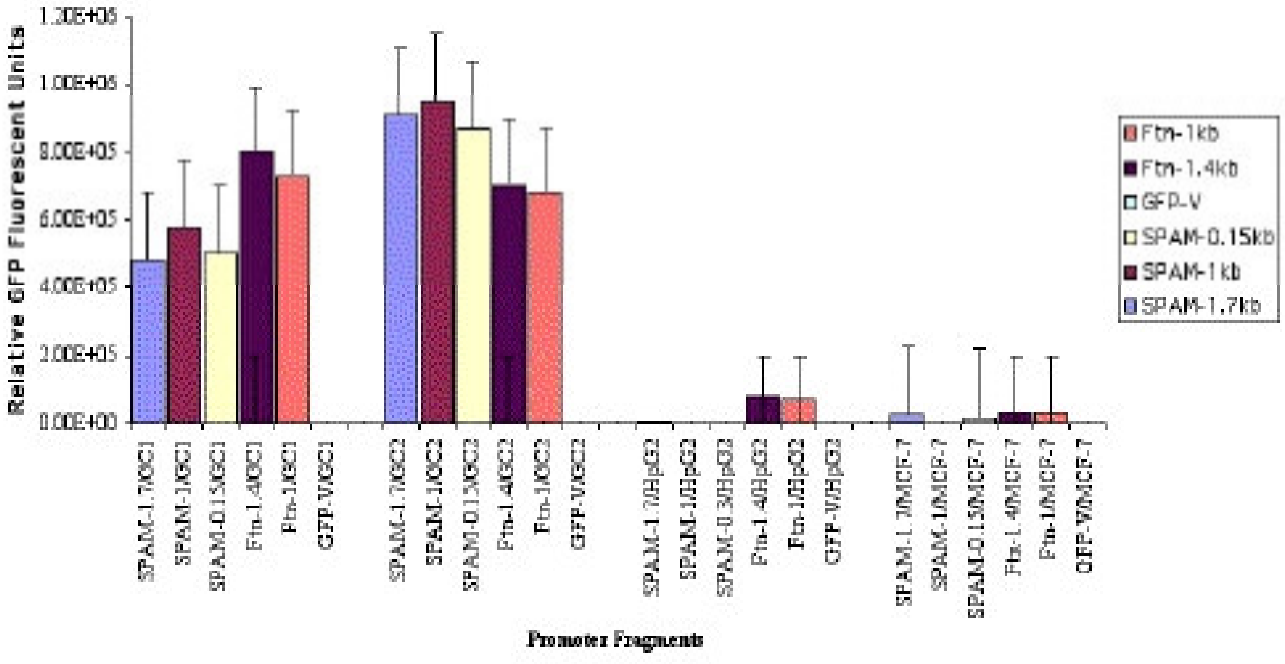


**Figure 3.** *SPAM1* and *FTN*-ß promoters activity in four cell types. GFP expression from *SPAM1* and *Fertilin ß* promoters' activity in four cell types. No significant GFP fluorescence was detected from either promoter in MCF-7 cells, while FTN-ß showed slight activity in Hep G2 cells. Both promoters were preferentially activated in mouse GC1 and GC2 germ cells, with most fluorescence detected in GC2 cells. The error bar is the standard deviation from the averaged fluorescence from the triplicate transfections per construct.

in co-culture with immortalised Sertoli cells. Prior to their utilisation for *in situ* promoter studies, the cells were genetically characterised and were found to express *Pgk-2* and protamine transcripts (data not shown). *Pgk-2* transcription is known to occur only in post-meiotic germ cells (Feig et al., 2007), while protamine is only expressed in round spermatids stage (Ike et al., 2007). Both *FTN*-β and h*SPAM1* are germ cell-specific genes and are not normally expressed elsewhere. This perhaps explains the remarkable difference in the transcriptional activity of these promoters in GC1 and GC2 cells compared to that observed in the HepG2 and MCF-7 cells.

## The *FTN*-β and *SPAM1* core promoters architecture

The *FTN*-βpromoter sequence is characterised by an initiator motif at its transcriptional start site (TSS) (see Figure 3A). The *FTN*-β promoter does not contain either a TATA or CCAAT box consensus sequences, thus it is a TATA-less promoter. Within its core promoter region it has a Sp1 binding site at -30 nucleotide position, where the TATA motif would have been, and a +30 downstream promoter element, as seen in other TATA-less initiator element containing promoters (Burke, 1997). It has several Sp1 and GC box motifs, with one Sp1 site right next to the downstream promoter element (DPE) motif. Significantly, however, it has two copies of the testis-specific factor, TSF, one in either orientation, a feature it shares with other germ cell-specific genes, such as the transition proteins, protamines, and PGK-2 (Eller et al., 2007). It is important to note that the two copies of the TSF motif have so far been observed to be in one orientation or the other in a particular gene, however, they are in the opposite orientation in the *FTN*-β promoter region. Furthermore, unlike most other testis-specific genes, the *FTN*-β core promoter region does not contain a cyclic-AMP response element motif (CRE), however, it has several Sp1 binding motifs, which perhaps suggests that it is controlled by the Sp1 transcription factors. The *SPAM1* core promoter sequence also contains the initiator motif at the transcription start site and, does not contain either the TATA or CCAAT consensus motifs. Although a TATA-less promoter with initiator motif at the TSS, it does not contain a DPE motif. Unlike, the *FTN*-β core promoter region, the *SPAM1* contains several CRE motifs, including at -27 and -70 nucleotide positions. It also contains several heat shock factor binding sites.

## Comparison of the h*SPAM1* and m*Spam1* core promoter

The m*Spam1* promoter does not contain either a TATA or CCAAT box. However, it contains several transcription factor binding sites, such as the cyclic-AMP response

**Table 3.** The CRE sequences at -50 to -70 promoter region.

| CRE Sequence (Consensus: TGACGTCA) | Consonant gene | Position |
|---|---|---|
| AAGTGAGG GTAGTTCA GCGGACTCG | hGK-1 | -70 |
| AGCTTCCTCT TGACTTCA TAATTCCT | Protamine 1 | -57 |
| TGGGCCG ACAGGTCA CAGTGGGG | Protamine 2 | -57 |
| TATGTAG TGACGTCA CAAGAGAG | Transition protein 1 | -57 |
| CACTTTG TGATGTCA TCTGTTCC | mSpam1 | -57 |
| TGCTTTG TGGTGTCA TCTTGCTT | hSPAM1 | -70 |
| GTTCTAAG TGACCTCA CAATATGG | m-calspermin | -50 |
| TGAGGGAA TGATGTCA TTTGAGT | h-calspermin | -70 |

The CRE sequence motifs as seen at -50 to -70 regions found in many postmeiotic to haploid cells-expressed genes. In hGK-1 (unpublished data), Prm-1, Prm-2, angiotensin-converting enzyme (Somboonthum et al., 2005), calspermin (Sun et al., 1995), Transition proteins and mouse Spam1 (Zheng, 1999).

**Table 4.**The testis specific factor sequence motif.

| TSF Consensus Sequence: G G G T G G G G | Gene | Reference |
|---|---|---|
| 5'-**G G G T G G G G**-3' | *PGK-2* | (Ella, 2007) |
| 5'-**G G G T G G G G**-3' | | |
| 5'-**G G G T G G G G**-3' | *FTN*-β | |
| 3'-**G G G T G G G G**-5' | | |
| 5'-**A G G T G G G G**-3' | *SPAM1* | |
| 3'-**G G G A G G G G**-5' | | |
| 5'-**G G G T G G G G**-3' | *hGK-2* | unpublished data |
| 5'-**G G G T G G G G**-3' | | |
| 5-**G G G T G G G G**-3' | | |
| 5'-**G G G T G G G G**-3' | *Prm-1* | (Domenjoud 1990) |
| 5'-**G G G T G G G G**-3' | *Prm-2* | (Domenjoud 1990) |
| 5'-**G G G T G G G G**-3' | | |

Testis specific factor (TSF) sequence motif as seen in germ cell-specific promoters. The TSF motif seemed to play a very important role in transcription of many testis-specific genes from mouse to man, including rat proacrosin (Kremling,1991a; Kremling,1995), mouse proacrosin (Kremling ,1991b), human proacrosin (Keime,1990), mouse protamines 1 and human protamine 1 and 2 (Domenjoud,1990), human PGK-2 (Ella, 2007) and human GK-1 and 2 (unpublished data).

element (CRE) motif at -50, HSFs (Heat Shock Factors) at -36, -71 and -228 nucleotide positions. It also contains an ADR1 (alcohol dehydrogenase receptor) and a possible SRY binding site. The most remarkable finding in the m*Spam1* promoter is the CRE motif at -50 significant effect on promoter activity, while the mutated -70 CRE reduced transcription by more than 60% ((Sun et al., 1995)). nucleotide position. A CRE motif at -50 or -70 nucleotide position has been described in several haploid-expressed gene promoters in mouse and Man (see Table 3 and 4). The h*SPAM1* promoter is significantly different than the m*Spam1* promoter. Firstly, the h*SPAM1* promoter has an initiator motif at the transcription start site and, secondly, it has several, not one, CRE motifs. The CRE at -70 nucleotide position in the *hSPAM1* promoter is highly similar to the CRE in *mSpam1* promoter at about the same location (see Table 4) and, it is compatible with the CRE found in several other haploid-expressed genes. Interestingly, however, the calspermin gene contains two CRE elements, one at -50 and the other at -70 nucleotide position. In an experiment to test the role of each of the two CRE motifs with respect to their influence on the calspermin promoter activity, the two CRE motifs were mutated. The mutated -50 CRE had a small but Interestingly, the 150 bp h*SPAM1* promoter fragment containing the -70 CRE motif had as strong or stronger transcriptional activity as the 1.7 and 1 kb fragments containing additional CRE motifs further upstream (see Figure 2 and Table 2). Although the influence of the CRE motif at -28 position cannot be ruled out, this suggests, also, that it is probably the -70 CRE motif that is most relevant in the h*SPAM1* promoter.

### Testis specific factor motif

*Prm-1* initiates transcription in round spermatids, as opposed to the pachytene cells for the metabolic enzymes (Ike et al., 2007). However, comparison of the regulatory regions of *PGK-2* and *Prm-1* revealed sequence elements, called testis specific factor motif that may be involved in tissue-specific expression. The TSF motif (consensus GGGTGGG) has been found in several postmeiotic and haploid-expressed genes, such as the *PGK-2*, *PDHA-2*, *TP-1*, *Prm-1* and *Prm-2*. Both *Prm-1* and human *PGK-2* 5' regions contain two copies each of the sequence GGGTGGG, present in opposite orientation in the two genes. This sequence is also found in the upstream region of the *Prm-2* gene, in the same orienttation as in *PGK-2* (Ella, 2007). Interestingly, the TSF is also found in both h*SPAM1* and *FTN*-β promoter regions (see Figure 3A and B). Analysis of the *GK-1* and *GK-2* promoter sequences also revealed that *GK-1* does not contain the motif, while *GK-2* does. The *GK-2* promoter contains two copies of the motif, as found in *Prm-1* and *PGK-2*, however, in the same orientation as in *PGK-2* (unpublished data).

There is considerable evidence for similarity between human and mouse gene promoters, including the ones observed in this paper. Although there have been problems with studying human gene promoters in mouse cell lines, many human testis-specific promoters have been activated in mouse germ cells. The results obtained in the analyses above suggest that the *SPAM1* and *FTN*-β promoters could be supported and retain their tissue specificity in mouse. Furthermore, they suggest that expressions of key genes, like *SPAM1,* are spatially and temporally regulated. This possibility will be analysed and confirmed in transgenic models, which is currently ongoing. While many testis-specific promoter motifs have been identified, such as TSF and DPE, their role in modulating transcription in germ cells remain to be determined.

#### REFERENCE


Baek HJ, Kang YK, Roeder RG (2006). Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. J. Biol Chem. 281(22): 15172-15181.

Day AE, Quilter CR, Sargent CA, Mileham AJ (2002). Characterization of the porcine sperm adhesion molecule gene SPAM1- expression analysis, genomic structure, and chromosomal mapping Anim. Genet. 33(3): 211-214.

DeLorenzi M, Ali N, Saari G, Henry C, Wilcox M, Bienz M (1988). Evidence that the *Abdominal-B* r element function is conferred by a trans-regulatory homeoprotein. EMBO 7: 3223-3231.

Dokemci D (2006).Testicular torsion oxidative stress and the role of antioxidant therapy.Folia Med (Ploidv).Rep. 48(3-4): 16-21.

Eller CD, Regelson M, Merriman B, Nelson S, Horvath S, Marahrens Y (2007). Repetitive sequence environment distinguishes housekeeping genes. Gene 390(1-2): 153-165.

Feig C, Kirchhoff C, Ivell R, Naether O, Schulze W, Spiess AN (2007). A new paradigm for profiling testicular gene expression during normal and disturbed human spermatogenesis. Mol. Hum. Reprod. 13(1): 33-43.

Hsu LC, Liu S, Abedinpour F, Beech RD, Lahti JM, Kidd VJ, Greenspan JA, Yeung CY (2003).The murine G+C-rich promoter binding protein mGPBP is required for promoter-specific transcription. Mol. Cell Biol. 23(23): 8773-85.

Ike A, Tokuhiro K, Hirose M, Nozaki M, Nishimune Y, Tanaka H (2007). Comprehensive analysis of gene expression in testes producing haploid germ cells using DNA microarray analysis. Int. J. Androl. 30(5): 462-475.

Kessler SP, Rowe TM, Gomos JB, Kessler PM, Sen GC (2006). N-cadherin mediated cell contact inhibits germinal vesicle breakdown in mouse oocytes maintained *in vitro*. Reproduction 131(3): 429-437.

Malik S, Barrero MJ, Jones T (2007). Identification of a regulator of transcription elongation as an accessory factor for the human Mediator coactivator. Proc Natl. Acad. Sci. U S A. 10, 104(15): 6182-6187.

Miyamoto M, Fujita T, Kimura Y, Maruyama M, Harada H, Sudo Y, Miyata T, Taniguchi T (1988). Regulated expression of agene encoding a nuclear factor IRF-1, that specifically binds to IFN-ß gene regulatory elements. Cell 54: 903-913.

Myles DG, Primakoff P (1997). Why did the sperm cross the cumulus? To get to the oocyte. Functions of the sperm surface proteins PH-20 and fertilin in arriving at and fusing with the egg. Biol. Reprod. 56 (2): 320-327.

Nishimura H, Cho C, Branciforte DR, Myles DG, Primakoff P (2001). Analysis of loss of adhesive function in sperm lacking cyritestin or fertilin beta. Dev. Biol. 233(1): 204-213.

Nishimura H, Myles DG, Primakoff P (2007). Identification of an ADAM2-ADAM3 complex on the surface of mouse testicular germ cells and cauda epididymal sperm. J. Biol. Chem. 282(24): 17900-17907.

Peluso JJ (2006). N-cadherin mediated cell contact inhibits germinal vesicle breakdown in mouse oocytes maintained *in vitro*. Reproduction. 131(3): 429-37.

Rubinstein E, Ziyyat A, Wolf JP, Le Naour F, Boucheix C (2006). The molecular players of sperm-egg fusion in mammals. Semin Cell Dev. Biol. 17(2): 254-263.

Shamsadin R, Adham IM, Neyernia K, Heinlein UAO, Oberwinkler H, Engel W (1999). Male mice deficient for germ-cell cyritestin are infertile. Biol Reprod 61: 1445-1451.

Somboonthum P, Ohta H, Yamada S, Onishi M, Ike A, Nishimune Y, Nozaki M (2005). cAMP-responsive element in TATA-less core promoter is essential for haploid-specific gene expression in mouse testis. Nucleic Acids Res. 33(10): 3401-3411.

Sun Z, Sassone-Corsi P, Means A (1995). Calspermin gene transcription is regulated by two cyclic AMP response elements contained in alternatetive promoter in the calmodulin kinase IV gene. Mol Cell Biol 15: 561-571.

Ziyyat A, Rubinstein E, Monier-Gavelle F, Barraud V, Kulski O, Prenant M, Boucheix C, Bomsel M, Wolf JP (2006). CD9 controls the formation of clusters that contain tetraspanins and the integrin alpha 6 beta 1, which are involved in human and mouse gamete fusion. J. Cell Sci. 119(3): 416-424.