A molecular analysis of the population of mRNA in bovine spermatozoa

Isabelle Gilbert, Nathalie Bissonnette¹, Guylain Boissonneault², Maud Vallée and Claude Robert

Centre de Recherche en Biologie de la Reproduction, Département des Sciences Animales, Université Laval, Québec G1K 7P4, Canada, ¹Dairy and Swine Research and Development Centre, AAFC, PO Box 90, Lennoxville, Québec J1M 1Z3, Canada and ²Département de Biochimie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada

Correspondence should be addressed to C Robert; Email: claire.robert@crbr.ulaval.ca

Abstract

Spermiogenesis represents the transition from haploid spermatids to spermatozoa. This process entails an extreme condensation of the nucleus and a loss of nearly all cytoplasmic content. The presence of messenger RNAs in the spermatozoa has previously been shown. Generally, these transcripts are considered to be remnants of spermiogenesis. However, it has recently been proposed that there may exist a function for these sperm-associated RNAs. To address the possibility of a functional role for these transcripts, we sought to investigate and characterize the RNA pool found in bovine spermatozoa. The main goals of this study were to examine RNA integrity and survey the mRNA found in spermatids and spermatozoa. Assessment of mRNAs integrity was performed by three approaches: microelectrophoresis, comparative smearing after global amplification, and PCR amplification of target sequences located either in the 5' or the 3' ends, while mRNAs survey was performed by microarray hybridizations. RNA integrity studies in the spermatozoa showed a majority of low molecular size fragments indicating a natural segmentation of the mRNA population. The mRNA survey indicated that the sperm transcriptome harbors a complex mixture of messengers implicated in a wide array of cell functions and representing a large subset of transcripts found in spermatids. Subsequently, such sperm RNA profiling could allow the molecular diagnosis of male gamete quality.

Reproduction (2007) **133** 1073–1086

Introduction

During spermiogenesis, the differentiation of spermatids into spermatozoa involves a change in chromatin structure leading to a greater level of DNA compaction. This is accomplished through the replacement of most histones with protamines and a major loss of its cytoplasm. There is a high level of transcriptional activity in spermatocytes, prior to the first meiotic division, then a gradual decrease in the rate of transcription, before a short surge at the stage of the round spermatid (Dadoue *et al.* 2004). Transcriptional activity shuts down when spermiogenesis reaches the elongating spermatids stage, when a replacement of histone proteins by protamines occurs (Dadoue *et al.* 2004). This model is supported by a recent study of human sperm in which an *in vitro* assay using incorporation of radio-labeled UTPs showed no detectable signs of transcription in the spermatozoa (Grunewald *et al.* 2005). In addition, the translational activity in the spermatozoa is also compromised as it has been shown that the spermatozoa harbor no or few

ribosomal RNAs (Ostermeier *et al.* 2002, Grunewald *et al.* 2005).

The presence of mRNA in spermatozoa is well established; yet little is known regarding its function and purpose (Dadoue *et al.* 2005). Given the inability of the spermatozoon to synthesize RNA, it is assumed that its RNA content originates from the trapped cytoplasmic content remaining after spermiogenesis, and consequently these RNAs may be not more than remnants. Recent interest in spermatid RNA has been motivated by the potential it may offer as a diagnostic tool for infertility. It has been proposed that, because of the delay between transcription and translation during spermatogenesis, the mRNA population present in spermatozoa should be representative of the past events of spermatogenesis (Ostermeier *et al.* 2002, Dadoue 2003, Steger 2003, Lambard *et al.* 2004).

In order to identify specific transcripts important for the production of functional spermatozoa, gene expression analyses have been conducted at specific stages of spermatogenesis as well as in different testicular somatic

cells. In rodents, microarray analysis has revealed a large array of candidates that are differentially and specifically regulated during spermatogenesis (Wrobel & Primig 2005). This observation was confirmed by another large-scale study performed in the rat, showing different transcripts in spermatogonia, spermatocytes, and spermatids when compared with Sertoli cells and other testicular somatic cells (Schlecht *et al.* 2004). It has been reported that the transcriptional regulators, whose absence is often associated with infertility, are mostly transcribed in the meiotic or post-meiotic germinal cells but not in the somatic tissues (Eddy 2002).

Currently, the literature contains only a few reports regarding sperm-associated mRNAs. Some studies highlighted the subcellular localization of the RNAs in the male gamete. For example, the RNA of c-MYC has been localized in the midpiece and the tail (Kumar *et al.* 1993), while other reports showed the presence of RNA within the nuclear compartment of the spermatozoa (Pessot *et al.* 1989, Wykes *et al.* 1997). In addition, mRNAs for the transcription factors nuclear factor-kappa B (NF κ B), homeobox 2A (HOX2A), interferon consensus sequence binding protein (ICSB), and the protein kinase c-jun n-terminal kinase 2 (JNK2), the growth factor heparin-binding EGF-like growth factor (HBEGF) as well as the retinoid X receptor β (RXR β) and epidermal growth factor receptor (ErbB3) receptors were all localized within the sperm head (Dadoune *et al.* 2005).

Recent studies reported that a quality assay for semen could be derived from RNA profiling. More specifically, the RNA stability of ejaculates of varying quality was examined by applying a freeze-thaw stress to the samples followed by a comparative analysis of RNA profiles using microarrays (Ostermeier *et al.* 2005b). Semen quality was found to be associated with transcript abundance as some transcripts were absent in the stress-sensitive spermatozoa. The potential involvement of spermatid transcripts for embryonic development was also proposed. Ostermeier *et al.* (2004) also detected two sperm-associated transcripts, protamine 2 and clusterin, in the fertilized egg therefore suggesting the contribution of spermatid transcripts to postfertilization development. Their finding has sparked considerable interest in the field since such a potential contribution of the male gamete to embryonic development challenges the well-accepted dogma, which restricted the male gamete to a DNA shuttling vector.

The objective of our present study was to characterize the bovine spermatid RNA pool first by assessing the integrity of the RNAs, then by a comparison of the transcripts found in the bovine spermatids and spermatozoa. This approach should provide important clues regarding whether the integrity and complexity of the mRNAs are compatible with a role in early embryonic development.

Material and Methods

Spermatozoa preparation

To eliminate variations between individual, semen samples from ten different Holstein bulls were pooled. These samples were a courtesy of the Centre d'Insemination Artificielle du Quebec (CIAQ, Sainte-Madeleine, QC, Canada). On an average, each ejaculate contained 7.04×10^9 spermatozoa. To eliminate damaged spermatozoa and contaminating somatic cells, the pooled semen samples were purified on a discontinuous gradient of Percoll 40:80 (Sigma-Aldrich). The Percoll solution (described by Parrish *et al.* 1988) was diluted in Tyrode's sperm medium (Sp-TALP; 100 mM NaCl, 3.1 mM KCl, 25 mM NaHCO₃, 0.3 mM NaH₂PO₄, 21.6 mM Na lactate, 2.0 mM CaCl₂, 0.4 mM MgCl₂, 10 mM Hepes, 1 mM pyruvate, and 50 μ g/ml gentamycin). The Percoll gradient purification was carried out as described by Parrish *et al.* (1988). Briefly, samples were deposited on the Percoll gradient and centrifuged at 700 *g* for 30 min. The pellets were washed and centrifuged twice at 250 *g* for 5 min in Sp-TALP solution. The pellets containing the motile spermatozoa were kept at -80°C in RNAlater (Ambion, Austin, TX, USA) until RNA extraction.

Isolation of spermatids

Testes from ten bulls were collected at a local slaughter house and transported to the laboratory on ice immediately after excision. These tissues were used for spermatids isolation as well as for controls in several experiments. The purification of spermatids was carried out as described by Morin *et al.* (2005). Briefly, the haploid germ cells were isolated from a piece of fresh testis under the tunica albuginea. The tissue pieces were washed in D-PBS (137 mM NaCl, 2.7 mM KCl, 0.9 mM CaCl₂, 0.5 mM MgCl₂, 1.5 mM KH₂PO₄, 8.1 mM Na₂HPO₄, pH 7.4), digested with trypsin 0.1% (Sigma-Aldrich), treated with Turbo DNase I (Ambion) then filtered on a nylon membrane with a mesh size of 70 μ m. The cells were incubated at 37 $^\circ\text{C}$, for half an hour in Hoeschst 33342 (Sigma-Aldrich). They were then sorted according to their ploidy using a fluorescence-activated cell sorter (Epics Elite ESP, Beckman Coulter, Miami, FL, USA), equipped with a laser Helium Cadmium (HeCd, Omnichrom Model 100 Chino, CA, USA) at a wavelength of 325 nm. The collected haploid cells were preserved in RNAlater (Ambion) and kept at -80°C until RNA extraction.

RNA extraction

Heated TRIzol method

Extraction of total RNA from the spermatozoa, the spermatids and testis was carried out using TRIzol

Reagent (Invitrogen). The protocol was followed according to the manufacturer's recommendation with a minor modification: for extraction, the TRIzol reagent was heated at 65 °C and the samples were incubated for half an hour to completely dissociate the membranes (alternative extraction protocols were tested in Dr Bissonnette's laboratory. In order to prevent any bias potentially arising from the RNA extraction procedure, the same method was used with the spermatid samples. The subsequent steps of the protocol were performed as recommended by the manufacturer. Total RNA sample pellets were dissolved in water and then washed on a RNA extraction column (RNeasy Mini Kit, Qiagen) upon which an RNase-free, DNase I treatment (Qiagen) was performed in order to eliminate contaminating genomic DNA from the samples. The integrity and the concentration of the total RNA samples were evaluated using a 2100-Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) with the RNA PicoLab Chip (Agilent Technologies).

Guanidium thiocyanate–phenol–chloroform extraction

The sperm and testis sample extractions were carried out as described in Chomczynski & Sacchi (1987). Briefly, for the extraction step, total RNA was isolated from 300 mg of a pool of testicular tissues homogenized in 3 ml of solution D (4 M guanidium thiocyanate, 25 mM sodium citrate, pH 7, 0.5% sarcosyl, 0.1 M β -mercaptoethanol; Sigma–Aldrich), while 500 μ l of ejaculate was put in 5 ml of solution D. Then sequentially, 0.1 volume of 2 M sodium acetate, pH 4, 0.1 volume of phenol and 0.2 volume of chloroform–isoamyl alcohol mixture were added to the homogenates. Following centrifugation, the supernatants were transferred and a precipitation step was done by adding one volume of isopropanol and centrifugation. The pellets were dissolved in solution D and a second precipitation was performed by adding one volume of isopropanol and centrifugation. Finally, the RNA pellets were washed in 75% ethanol and resuspended in RNase-free water. In addition, a variation to the protocol was tested by replacing the 0.1 M β -mercaptoethanol with 0.1 M dithiothreitol (DTT; Sigma–Aldrich). Total RNA samples were further purified by performing DNase I treatment on RNeasy Mini Kit columns (Qiagen). The quantity and the quality of the total RNA extracts were measured using a 2100-Bioanalyzer (Agilent Technologies) with the RNA PicoLab Chip for the sperm samples and the RNA NanoLab chip for the testis samples (Agilent Technologies).

5' versus 3' target sequence comparisons

For each of the five candidates, two primer sets were designed. A first primer set was designed to overlap the start codon in 5' and the second primer set to overlap the stop codon in 3'. For both cell types, total RNA extracts

(25 ng) from spermatids and spermatozoa samples were submitted to reverse transcription and global amplification using the BD SuperSMART PCR cDNA synthesis kit (BD Biosciences, Mississauga, ON, Canada) according to the manufacturer's instructions. The PCR amplifications were performed using 1 μ l of the reverse transcription reaction in a final volume of 25 μ l containing 1.5 mM $MgCl_2$, 0.4 mM of each dNTPs (Amersham), and 10 pM of each primers, in the provided reaction buffer at 1X concentration and 0.2 μ l AmpliTaq Gold LD (Applied Biosystems, Foster City, CA, USA). The specific primers for the targeted gene transcripts were designed using the Primer3 Website (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) based on the appropriate sequence reported in Genbank. The primer sequences used are described in Table 1. The PCR conditions were 95 °C for 5 min, followed by 35 cycles of 30 s at 95 °C, 30 s at 57 °C, 45 s at 72 °C, followed by a final extension of 10 min at 72 °C. The PCR products were then resolved on a 2% agarose gel stained with ethidium bromide. Amplicons were detected under u.v. light using a Chemigenius Bioimaging.

Test to detect genomic DNA contamination of samples

Total RNA extracts (25 ng) from spermatids and spermatozoa samples were submitted to reverse transcription using the BD SuperSMART PCR cDNA synthesis kit (BD Biosciences) according to the manufacturer's instructions. Genomic DNA contamination of the samples was tested by PCR under the conditions described above using a set of primers specific to protamine 1 (PRM1) that spans an intron. Leukocyte contamination was evaluated by PCR using two sets of primers respectively targeting the CD4 and CD45 antigens sequences. The positive controls for genomic DNA and leukocyte contamination were as follows: a purified bovine genomic DNA extract isolated from bovine hair follicles using an in-house phenol/chloroform-based protocol routinely used in Dr Bissonnette's laboratory for single nucleotide polymorphism (SNP) analysis; and total RNA extracted from a blood sample with TRIzol (Invitrogen) in accordance with standard procedures respectively. The PRM1, CD4, and CD45 primer sequence used are listed in Table 1.

Microarray probe preparation and hybridizations

Total RNA extracts (25 ng) from spermatids and spermatozoa samples were submitted to reverse transcription and global amplification using the BD SuperSMART PCR cDNA synthesis kit (BD Biosciences) according to the manufacturer's instructions. The samples were confirmed free of genomic and white blood cell contaminations using the procedure described above. The amplified cDNA samples were labeled using an

Table 1 Bovine primer sequences of the genes used.

Gene	Primer sequences		Unigen and Genbank number	Product size (bp)
	Forward	Reverse		
ACTB_5'	5'-CGCCATGGATGATGATATTG-3'	5'-GGTCATCTTCTACGGTTGG-3'	Bt.14186; AY141970	364
ACTB_3'	5'-CCTAAGTTCGCGAGAAAACG-3'	5'-TACAGGAAAGCCCTGACTGC-3'		529
AKAP4_5'	5'-GCCCATTCCATCTCAAGG-3'	5'-TTTGGCTGCTGTCATTTCC-3'	Bt.218; NM_174235	603
AKAP4_3'	5'-AACAAACATCCGAGATCAGG-3'	5'-TTGGGACACCCGTGATTTGC-3'		784
CLU_5'	5'-GTGGAGAAGGAGGCAAGATG-3'	5'-CATTCTCCATCAGGGAATCG-3'	Bt.12504; NM_173902	495
CLU_3'	5'-TCCTGGAAAGTGGACTGTTTCG-3'	5'-GCATTAGTGCCTTTCAGAGG-3'		597
P80_5'	5'-AATTGATTGCTGTGCTCATCC-3'	5'-TGTCCGGTAGCATTCTTGG-3'	Bt.22967; AY781776	474
P80_3'	5'-GACCTTGTGAATTCGGTTGG-3'	5'-GGGATGACTTCCGAACACC-3'		605
SPAG4_5'	5'-GAGTTACAAAGGAGAGGGTTCG-3'	5'-CAGAGCTGTTGCTGTTGTCG-3'	Bt. 24857; BC109514	238
SPAG4_3'	5'-CCTTTCCCAAGGTGAAGATCC-3'	5'-CCCTCTGAGGTTTCAATACC-3'		100
PRM1	5'-AGATACCCGATGCTGCCTCAC-3'	5'-GTGGCATGTTCAAGATGTGG-3'	Bt. 20926; NM_174156	234 (315 with intron)
CD4	5'-CAATGGCAAAGTCTGTTGG-3'	5'-GATCTGAGACATCCGTTCTGC-3'	Bt. 30898; AJ535319	184
CD45	5'-GACATCGCAGTGTGTTGTC-3'	5'-GGAGGTTACATTCCTCTCG-3'	Bt. 29715; AJ400864	239
MEA	5'-GAGCAGTGAGGAACCAGAGG-3'	5'-CTGGAATGCAGGGAGTTAGG-3'	Bt. 49043; D17340	451
SSP411	5'-GGAGCAGTTCCTGAGTTTCG-3'	5'-GGATGGACAAGTGTGTGTGC-3'	Bt. 5589; CB422327	360
SPAG4	5'-ACTGCTGAGCCTCTTTTGG-3'	5'-AGTAGGCACTGTCGATCC-3'	Bt. 24857; BC109514	338
KLHDC3	5'-TTCCAATGACATCCACAAGC-3'	5'-TACAGCTCCCCATTGTAGCC-3'	Bt. 4489; BT020656	297
TEGT	5'-TCACAGCCATGAAACTGAGC-3'	5'-ATGGCCGACATCAAGATACC-3'	Bt.53719; BU917243	228
NSEP1	5'-GTACTCAACCTCCTGTGC-3'	5'-TGACCTTGGGTCTCATCTCC-3'	Bt.5332; NM_174815	201
PPIH	5'-ATACAGGGGAGACCTTCC-3'	5'-ACGTGCTTCCCCTAGCC-3'	Bt.6641; XM_864506	240
H2AFZ	5'-CTTGAATTGGCAGGAAATGC-3'	5'-GCAGAAATTGGTTGGTTGG-3'	Bt.2515; NM_174809	358
STRBP	5'-TTTTTAAAGCGAGTGCATGG-3'	5'-GCAGCAAAGGTTAGCAAAGG-3'	Bt.5039; CK972240	406
EIF2B2	5'-GAGACTCTGGGACTGCTTCG-3'	5'-CGGATGTCAAGAGTTTGTGC-3'	Bt.3050; BT021006	238
FLOT1	5'-AAGTGCTGGACATCCTGAGC-3'	5'-CAATAACGGGACCATTCAGG-3'	Bt.26845; XM_879742	159

ACTB, actin beta; AKAP4, a kinase anchor protein 4; CLU, clusterin; P80, 80 kDa protein; SPAG4, sperm associated antigen 4; PRM1, protamine 1; CD4, CD4 antigen; CD45, CD45 antigen; MEA, male-enhanced antigen; SSP411, sperm protein SSP411; KLHDC3, kelch domain containing 3; TEGT, testis enhanced gene transcript; NSEP1, nuclease sensitive element binding protein 1; PPIH, peptidylprolyl isomerase H; H2AFZ, H2A histone family member Z; STRBP, spermatid perinuclear RNA binding protein; EIF2B2, eukaryotic translation initiation factor 2B subunit 2 beta; FLOT1, flotillin 1.

indirect procedure incorporating amino allyl-dUTPs (Ambion). The incorporation was done using the Klenow enzyme (New England Biolabs, Pickering, ON, Canada) in a DNA polymerization reaction primed with random decamers (Ambion). The incubation was carried out at 37 °C for 2 h. The fluorophores (Alexa Fluor 555 and 647, Invitrogen) were added chemically to the amino allyl groupings according to the vendor's instructions. The spermatozoa and spermatids sample probes were hybridized for 16 h at 55 °C on the SS-Human19Kv7 slides containing 19 200 single-spotted human cDNA (University Health Network Microarray Center, Toronto, ON, Canada) in the SlideHyb1 buffer (Ambion). Hybridizations were performed in an ArrayBooster using the AC3C Advacard (The Gel Company, San Francisco, CA, USA). Following the overnight incubation, the slides were washed twice in a solution of low stringency (2X SSC–0.5% SDS) at 55 °C during 15 min followed by two washes in a high stringency solution (0.5X SSC–0.5% SDS) at 55 °C for 15 min and three final washes in a solution of 1X SSC. The experimental design included two biological replicates each with a dye-swap technical replicates (sub-replication) in order to account for any differences associated with the fluorophores for a total of four microarray hybridizations. The slides were scanned using a VersArray ChipReader System (Bio-Rad) and analyzed with the ArrayPro Analyzer software (Media Cybernetics, San Diego, CA, USA).

Microarrays analysis

Analysis of the microarray data sets was carried out to obtain a gene lists for each of the two cellular types used. The procedure was inspired by the article published by Vallée *et al.* (2005). The data were transformed into a base 2 log in order to obtain a curve of standard distribution. Uninformative data were removed from the analysis by determining a significant threshold of cut-off based on a degree of confidence associated with the variability associated with the negative controls. This cut-off threshold was calculated as follows: $T = m + 2 \times ST$, where T is the calculated threshold for cut-off, m is the average of negative controls present on the slides and ST is the standard deviation of these negative controls. Moreover, all the data equal or lower to the cut-off threshold determined previously were not considered in the analysis. Finally, a gene is regarded as positive for the analysis and included in the gene list if the signal was higher than the background noise determined and present in both replicates of hybridizations and both of their sub-replications.

Validation of microarray results

For spermatozoa samples, the starting material was generated by SuperSMART as described above. For spermatid samples, total RNA (15–20 ng), treated with

DNase I as described above was reverse transcribed using the Sensiscript reverse transcriptase (Qiagen) according to the manufacturer's recommendations in a volume of 20 μ l. The reaction was primed using an oligo(dT) primer (1 μ M; Ambion) and incubation was allowed to proceed for 1 h at 37 °C in a ThermoHybaid Hybrid Multiblock System (Bio-Rad). The PCR amplifications were performed and visualized on agarose gel as described above.

Results

RNA extraction and integrity assessment

Because of the membrane sturdiness of the bovine spermatozoa, total RNA was extracted by incubating the spermatozoa samples in heated TRIzol reagent for half an hour. Extraction at room temperature resulted in poor RNA yields and was not harsh enough to completely destabilize the cellular membranes as debris were clearly visible at the bottom of the tube following centrifugation (data not shown). To avoid creating a bias, the spermatid samples were extracted with the same method. The impact of this protocol modification to the usual protocol was assessed in order to determine if it affected the integrity of the isolated mRNAs. In order to do so, the integrity of the spermatozoa and spermatid RNA extracted by the heated TRIzol was evaluated by microelectrophoregram using the 2100-bioanalyzer (Fig. 1A and B). The RNA isolated from the spermatid extract shows a normal profile, e.g. the presence of two peaks corresponding to the 18S and 28S rRNA and a lower baseline shift indicative of the mRNA smearing profile. By contrast, the spermatic RNA sample clearly shows an absence of rRNA and the overall RNA profiling displays a majority of short-length fragments. This uncommon profile was further investigated in order to make sure whether the preponderance of short size RNA is a typical spermatic RNA characteristic and not an artifact of the RNA extraction method.

To determine if the observed spermatic RNA profiles were resulting from the extraction procedures or truly representative of the spermatic RNA population, a comparative electrophoresis profiling analysis was performed with sperm and testis total RNA samples extracted under diverse conditions. Electrophoregrams are shown in Fig. 2, which displays the conventional or heated TRIzol methods (Fig. 2 panels A–C) and the single-step RNA isolation method described by Chomczynski & Sacchi (1987) (Fig. 2 panels D and E). Additionally, the Chomczynski & Sacchi (1987) protocol was modified by substituting the β -mercaptoethanol with dithiothreitol for a greater reducing potential to enhance dissolution of the sperm structure (Fig. 2 panels F and G). All of the RNA isolation methods show the same short size RNA profile for the spermatozoa and a normal profile for the testis. Comparison between heated and room temperature

Figure 1 Total RNA microelectrophoretic profiles of spermatids (A) and spermatozoa (B) samples. Total RNA was extracted with the heated TRIzol procedure. M, marker; FU, fluorescence; S, seconds.

TRIzol treatment can be seen in Fig. 2B and C. The electrophoregrams indicate that the peak corresponding to the 28S rRNA is clearly lower in the heated TRIzol samples (Fig. 2 panel B, white arrow) comparatively to the room temperature TRIzol treatment (Fig. 2 panel C). This translates into a higher abundance of midsize fragments in the range of the 18S rRNA (1.8 kb) and smaller than the 28S rRNA (4.5 kb; rRNA product sizes based on reported sequence Genbank DQ222453).

Based on the starting cell concentration used for the RNA extraction and the RNA content measured using the electrophoregram data, each bovine spermatozoon contains $\sim 1.8 \times 10^{-4}$ pg total RNA, whereas the quantity of total RNA was estimated at 0.45 pg per spermatid.

Global amplification of mRNA

Considering the small amount of RNA contained in spermatic samples, a global amplification step was necessary to obtain a sufficient amount of cDNA for the microarray hybridization. To avoid any bias that might be introduced by this amplification step, the spermatic and spermatid samples were amplified in parallel using the same procedure. The global amplification step was carried out using the SuperSMART (Clontech) kit.

Figure 2 Comparison of total RNA microelectrophoresis profiles between the spermatozoa and testis samples extracted using diverse methods. Upper panels: total RNA extracted from sperm (A) and testis (B and C) obtained either by heating the TRIzol (A and B) or using the conventional room temperature TRIzol extraction procedure (C). Middle panels: total RNA from sperm (D) and testis (E) extracted with the guanidium-phenol-chloroform method supplemented with β -mercaptoethanol. Lower panels: total RNA from sperm (F) and testis (G) isolated with guanidium-phenol-chloroform method supplemented with DTT. The white arrow indicates the reduction of the 28S rRNA caused by the heated TRIzol treatment. M, marker; FU, fluorescence; S, seconds.

This PCR-based approach follows a typical sigmoid curve associated with an exponential amplification that reaches a plateau phase. It is necessary to prevent over cycling, which could result in a serious bias in the proportionality of the transcripts, as well as under cycling, which could result in an insufficient amount of material for the microarray probe. Therefore, the manufacturer recommends to perform an optimization reaction in parallel and taken to completion (with aliquots taken periodically), while the experiment reaction is held at 15 cycles. These aliquots are then separated on gel to visualize the amplification profiles. This step is used to determine the optimal number of cycles suitable for the experiment holding reaction. Figure 3 shows the distribution profiles of cDNA populations resulting from aliquots taken at various cycles of the amplification reaction. It is worth noting that the smears generated by the spermatid sample are lower than the one produced by the spermatozoa.

This also confirms the preponderance of RNA species of smaller size present in the spermatozoa when compared with those from the spermatid population.

PCR amplification of target sequences located either in the 5' or the 3' ends

The short size spermatid transcriptome was examined more thoroughly using two sets of primers designed either to cover the initiation codon (5' end) or near the stop codon (3' end; Fig. 4A). The presence of amplicons was compared between the two cell types (Fig. 4B). In all cases, the 3' target sites resulted in the amplification of the respective sequences. However, for the 5' end targets, all amplicons were detected in the spermatids, whereas only four out of five amplicons were detected in the spermatozoa. Since the reverse transcription reaction is primed using an oligo(dT), the 3' end targets are most likely to

Figure 3 Electrophoresis on a 1.2% agarose gel stained with ethidium bromide of the aliquots of cDNA taken at various cycles during the Super-Smart preamplification step. Each lane is numbered according to the number of PCR cycles carried out. M: 1 KB plus DNA marker.

Figure 4 Comparison between the spermatids and the spermatozoa for the presence of PCR amplicons targeting the 5' or 3' ends. (A) Schematic of the localization of the primers sets for each gene candidates. The 5' end primer set overlaps the start codon (ATG), while 3' end primer set is localized near or overlapping the stop codon. (B) Electrophoresis of the PCR amplification products using both primer sets (5' or 3') for both cell types. N, negative control; M, 1 kb plus DNA marker.

be amplified if the transcript population is fragmented. Figure 4B shows that the transcripts for sperm associated antigen 4 (SPAG4), P80, clusterin (CLU), and A kinase anchor protein 4 (AKAP4) are truncated, while bearing at least a short poly (A) tail enabling the annealing of the oligo(dT) primer. The fact that the 5' end target sites were mostly not amplified is indicative that the messenger population in the spermatozoa is segmented contrarily to the one found in the spermatids. However, the sperm transcriptome contains some full-length mRNAs as demonstrated by the 5' and 3' amplicons for actin beta (ACTB; Fig. 4B).

Verification of sample purity

The purity of the samples was checked by RT-PCR. The *PRM1* gene was targeted to detect the presence of genomic DNA (gDNA) contamination, since the intron-spanning primers were used (Fig. 5). Hence, the presence of gDNA in the sample produces an amplicon of 315 bp (Fig. 5, Lane 1). No genomic DNA contamination was detectable in the samples used to prepare the microarray probes. An assessment of contamination by cells of hematopoietic origin was carried out by RT-PCR using the CD4 and the CD45 markers (data not shown). The presence of contamination by hematopoietic cells remained undetected in all of the samples used to prepare the microarray probes.

Microarray hybridizations

Due to the lack of well-developed bovine microarray commercially available, a microarray containing 19 200 human cDNAs was used to survey the transcripts present in the spermatids and those from spermatozoa. In our experience and others, it is known that cross-species hybridization is a powerful and robust method that allows the identification of transcripts (Adjaye *et al.* 2004, Shah *et al.* 2004, Vallée *et al.* 2005, 2006). Moreover, comparative analyses of expressed sequence tags (ESTs) between human and bovine show an average identity level higher than 85% (unpublished observations). Using

Figure 5 The intron-spanning protamine 1 (PRM1) primer set was used to detect the presence of genomic DNA contamination. Lane 1: positive control, a spermatozoa RNA sample was spiked with genomic DNA to produce the amplicon of 315 bp corresponding to genomic contamination and an amplicon of 234 bp corresponding to PRM1 cDNA. Lane 2: without contamination.

the stringent cut-off limit described in the Material and Methods section and the requirement that a target must be present in all four replicates in order to be considered, the array analysis revealed 2583 and 1117 positive sequences for the spermatids and the spermatozoa samples respectively (see supplementary data for transcripts identity which can be viewed online at www.reproduction-online.org/supplemental/). Of these positive signals, 996 were found in both cell types (Fig. 6). To evaluate the diversity of function associated with the transcripts, the candidates were grouped according to their known cellular functions (Fig. 7). For each cell type, a list of 45 gene transcripts that produced the strongest signal is reported in Tables 2 and 3.

Validation of the microarray hybridization results

For microarray validation, specific candidate genes were chosen according to their known implication in spermatogenesis or their contribution to early embryonic development. The microarray analysis indicated that male-enhanced antigen (MEA), sperm protein SSP411 (SSP411), SPAG4, kelch domain containing 3 (KLHDC3), testis enhanced gene transcript (TEGT), nuclease sensitive element binding protein 1 (NSEP1), peptidylprolyl isomerase H (PPIH), and flotillin 1 (FLOT1) amplicons were present in both cell types, whereas 2A histone family member Z (H2AFZ), spermatid perinuclear RNA binding protein (STRBP), and eukaryotic translation initiation factor 2B subunit 2 beta (EIF2B2) sequences were exclusive to spermatids (Fig. 8). During the validation process, all candidate genes were undetectable in the unamplified spermatozoa samples. However, cell type specificity for the spermatozoa became clear only when using a preamplified sample indicating the low level of abundance of the targeted transcripts. In contrast, cell-type specificity for the candidate genes

could be confirmed in the spermatids without prior amplification. For all the selected candidates, RT-PCR results confirmed the microarray data. FLOT1 amplicon was selected as a spermatozoa-specific transcript, since it did not meet the selection requirement in the spermatid samples as it was only positive in one of the biological microarray replicates (positive for both sub-replications). The RT-PCR validation detected the presence of FLOT1 messenger in both cell types.

Discussion

The possibility that the role of the male gamete goes beyond the well-established DNA shuttling function has gained recent interest (Ostermeier *et al.* 2002, 2004, 2005a). Although the presence of RNAs trapped within the spermatozoa is well documented (Kumar *et al.* 1993, Ostermeier *et al.* 2002, Dadoune *et al.* 2005, Grunewald *et al.* 2005), both the identity and the role of these transcripts remain largely uncharacterized. The goal of the present study was to characterize and survey the RNA population found in the bovine spermatozoa. Many precautions were taken to obtain cellular samples free of genomic DNA. This ensured that the following data would be representative of the cell type under study. The use of the discontinuous Percoll gradient to purify the semen samples eliminated the damaged spermatozoa, ensuring that the downstream characterization is specific to intact cells. In addition, pooled samples were used to avoid conclusion based on individual specificities. The spermatid samples were isolated by staining the DNA with Hoechst in order to sort the cells according to their ploidy. It is noteworthy that this approach does not discriminate between the round, the elongating, and the elongated spermatids. Thus, for this study, the spermatids samples include all the three types of spermatids.

Figure 6 Venn diagram of the proportion of specific or common transcripts between spermatids and spermatozoa. A total of 2583 sequences were found positive in the spermatid, while spermatozoa contain 1117 positive signals. Of these positive signals, 978 were found in both cell types. None of the 121 candidates considered to be present in the spermatozoa were absent in any of the four hybridizations performed with the spermatids.

Figure 7 Illustration of the various cellular functions associated with the positive candidates detected by microarray. The upper diagram represents the spermatids, while the lower diagram represents the spermatozoa. The results are expressed as a percentage.

Even though the general structure of the spermatozoa is similar between species, very harsh methods must be used to destabilize the membrane structure of bovine spermatozoa. This is not the case for all species, for instance, the porcine spermatozoa are completely destabilized in extraction buffer at room temperature (data not shown). [Ostermeier et al. \(2005a, 2005b\)](#) prepared their human spermatic RNA samples by heating the lysis buffer provided by Qiagen. This approach was clearly ineffective for bovine spermatozoa because substantial cellular debris is deposited at the bottom of the tube following lysis, which resulted in poor RNA yields (data not shown). Despite the harsh conditions in which the bovine spermatozoa had to be

lysed, the impact of this treatment showed no sign of mRNA damage but resulted in the partial degradation of the 28S rRNA. Such observation has been reported for bacterial RNA extraction where heated lysis buffer is used to eliminate the rRNAs without damaging the mRNAs ([Sung et al. 2003](#)). Interestingly, the two major rRNAs were absent from the spermatic RNA samples as shown by the electrophoregram profile. This observation is in agreement with previous reports ([Ostermeier et al. 2002](#), [Grunewald et al. 2005](#)). Indeed, it is generally accepted that mature spermatozoa are not translationally active so that rRNAs essential for ribosome assembly may not be available. Despite such assumption, the presence of 80S cytoplasmic ribosomes containing the

Table 2 List of the 45 microarray probes producing the strongest signals for spermatids cDNA targets.

GeneBank no.	Unigene no.	Clone no.	Name
Known genes			
R85191	Hs.440885	180728	Essential meiotic endonuclease 1 homolog 1
BQ064037	Hs.356572	284659	Ribosomal protein S3A
BQ066818	Hs.356502	239082	Ribosomal protein, large P1
N46720	Hs.511758	277033	Serologically defined colon cancer antigen 8
BM928583	Hs.387208	230519	Finkel-Biskis-Reilly murine sarcoma virus
BM999886	Mm.253	490090	Nuclease sensitive element binding protein
W39594	Hs.74497	322882	Nuclease sensitive element binding protein
BQ072702	Hs.424126	296260	Small EDRK-rich factor 2
H21568	Hs.414042	160008	Cyclin M3
AA058476	Hs.77318	489369	Platelet-activating factor acetylhydrolase
W05242	Hs.279806	298689	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
BQ045003	Hs.425808	486800	Calmodulin 2 (phosphorylase kinase, delta)
BM922198	Hs.433615	301465	Tubulin, beta, 2
BQ021777	Hs.46405	342659	Polymerase (RNA) II (DNA directed)
BQ056602	Hs.150580	115010	Putative translation initiation factor (SUI1)
R17451	Hs.278362	32327	Male-enhanced antigen
H38887	Hs.446567	191907	Basic transcription factor 3
N92087	Hs.84113	293274	Cyclin-dependent kinase inhibitor 3
N36269	Hs.5308	269776	Ubiquitin A-52
BM552637	Hs.279806	306186	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
H61388	Hs.408096	236390	Fragile X mental retardation, autosomal homolog 1
AA151568	Hs.35052	504349	Testis enhanced gene transcript (BAX inhibitor 1)
BQ062070	Hs.406300	191910	Ribosomal protein L23
BG248268	Hs.356572	298773	Ribosomal protein S3A pseudogene 6
BG502582	Hs.425808	66495	Calmodulin 2 (phosphorylase kinase, delta)
Uncharacterized genes			
N28396	Hs.128927	263052	mRNA; cDNA DKFZp666B189
AA131892	Hs.32826	504286	Chromosome 6 open reading frame 74
Unknown genes			
BM844216	N/A	239014	N/A
Unknown	N/A	301431	N/A
BM782284	N/A	176939	N/A
Unknown	N/A	273919	N/A
Unknown	N/A	366637	N/A
Unknown	N/A	682763	N/A
Unknown	N/A	146816	N/A
BM783962	N/A	302402	N/A
Unknown	N/A	117727	N/A
BM837817	N/A	172403	N/A
AL558551	N/A	155222	N/A
Unknown	N/A	366302	N/A
Unknown	N/A	119886	N/A
Unknown	N/A	357679	N/A
Unknown	N/A	290360	N/A
H86672	N/A	223136	N/A
Unknown	N/A	238885	N/A
Unknown	N/A	221846	N/A

N/A, not available.

18S and the 28S rRNAs has been detected in bovine spermatozoa (Gur & Breitbart 2006). In this report, the presence of 18S rRNA was indeed detected by RT-PCR and a translational activity was attributed to the mitochondrial ribosomes while those of the cytoplasm were inactive. The presence of 18S rRNA only detectable by RT-PCR is therefore in agreement with previous reports showing that the rRNAs are present at an extremely low level in the spermatozoa.

Comparative size profiling of the transcript population performed by the microelectrophoresis approach and further confirmed by the preamplification step revealed that when compared with the spermatid and testis

samples, the spermatid RNA population is predominantly composed of small-size RNAs (<1 kb). This observation is divergent from the smear length previously reported in the human where it spanned over the 28S ribosomal RNA size (Ostermeier *et al.* 2002, Grunewald *et al.* 2005). These discrepancies could be species specific as it is known that the human spermatid population is much more heterogeneous than the bovine, which has been submitted to intensive breeding programs. Indeed, human semen is composed of sperm with varying degrees of structural and functional differentiation and normality (Buffone *et al.* 2004), which is not the case for bull semen samples.

Table 3 List of the 45 microarray probes producing the strongest signals for the spermatozoa cDNA targets.

GeneBank no.	Unigene no.	Clone no.	Name
Known genes			
R85191	Hs.440885	180728	Essential meiotic endonuclease 1 homolog 1
BQ064037	Hs.356572	284659	Ribosomal protein S3A
BQ066818	Hs.356502	239082	Ribosomal protein, large, P1
N46720	Hs.511758	277033	Serologically defined colon cancer antigen 8
AA054071	Hs.326392	380366	Son of sevenless homolog 1 (<i>Drosophila</i>)
R90798	Hs.486214	167228	REV3-like, catalytic subunit of DNA pol. zeta
N31674	In multiple clusters	266981	Amyotrophic lateral sclerosis 2 (juvenile)
H70514	Hs.302145	296266	Hemoglobin, gamma G
H50471	Hs.379186	179003	Programmed cell death 6
H88495	Hs.436885	252928	Histidine rich calcium binding protein
N34169	Hs.902	267720	Neurofibromin 2
W24097	Hs.188553	306852	Retinoblastoma binding protein 6
R81899	Hs.416216	147662	Dual specificity phosphatase 12
H30637	Hs.124165	190079	Mitochondrial ribosomal protein S30
H75353	Hs.1227	230467	Aminolevulinate, delta-, dehydratase
N45505	Hs.116237	277859	Vav 1 oncogene
AA043244	Hs.367877	486133	Matrix metalloproteinase 2
AA130221	Hs.41690	504161	Desmocollin 3
AA136761	Hs.395309	490996	Thioredoxin
Uncharacterized genes			
N28396	Hs.128927	263052	mRNA; cDNA DKFZp666B189
AA131892	Hs.32826	504286	Chromosome 6 open reading frame 74
N53883	Hs.343513	281551	KIAA0276 protein
AA044889	Hs.40836	488418	EST, clone image: 488418, 5'end
R61165	Hs.47166	42868	Chromosome 3 open reading frame 14
H47172	Hs.504646	193109	KIAA0329
N94798	Hs.491571	278098	EST, clone image: 278098 5'end
N24364	Hs.408702	261698	Hypothetical protein FLJ13154
Unknown genes			
BM783962	N/A	302402	N/A
Unknown	N/A	357679	N/A
Unknown	N/A	301431	N/A
Unknown	N/A	380794	N/A
BG167084	N/A	286050	N/A
AV690844	N/A	36318	N/A
AL698489	N/A	129473	N/A
Unknown	N/A	290360	N/A
H86672	N/A	223136	N/A
BM837817	N/A	172403	N/A
Unknown	N/A	301503	N/A
H96724	N/A	251637	N/A
AA035434	N/A	471595	N/A
R84724	N/A	180258	N/A
AU135696	N/A	153754	N/A
Unknown	N/A	147732	N/A
AA035019	N/A	471750	N/A
BM844216	N/A	239014	N/A

N/A, not available.

The presence of mainly small size RNA is not a confirmation of the complete absence of full-length transcripts. This question was addressed using the 5' versus the 3' ends PCR amplifications. This test clearly shows that the spermatid RNA contains full-length messenger RNA, whereas for four out of five candidates of the spermatid transcripts were totally truncated. The use of an oligo(dT) to prime the reverse transcription reaction is not informative on whether it is degraded or rather cut-off from the 3' end. In eukaryotes, there are two general mRNA decay pathways that are both initiated with the removal of the poly(A) tail (Meyer *et al.* 2004). Following deadenylation, the mRNA is either degraded

from the 5' end, which involves the removal of the cap structure and the activation of a 5'–3' exonuclease or simply from the 3' end through the activity of a 3'–5' exonuclease, which ends by the hydrolysis of the cap structure (Meyer *et al.* 2004). In the case of the spermatid transcriptome, the integrity assessment described herein supports the presence of short 3' ends bearing at least a short poly(A) tail.

The majority of functional mRNAs must have a poly(A) tail to be translated. It is known that during spermatogenesis, mRNAs found at the elongated spermatid steps undergo deadenylation (Kleene 1993, 1996, Schmidt *et al.* 1999). It would therefore be logical to find

Figure 8 Validation of the microarray results for selected candidates by RT-PCR. For all candidates, tissue specificities found by microarray were confirmed by RT-PCR. It is important to mention that none of the candidates could be detected by RT-PCR in the unamplified spermatozoa samples. Tissue specificity for spermatozoa samples was confirmed in SuperSMART preamplified samples, whereas confirmation of specificity for spermatids was conducted in non-amplified samples. ^aCommon transcripts found in spermatids and spermatozoa. ^bSpermatid-specific transcripts. ^cSpermatozoa-specific transcripts (but positive in half of the spermatid replicates). SPZ, spermatozoa (pre-amplified); SPD, spermatid; N, negative control; M, 1 kb plus DNA marker.

deadenylated mRNAs in the spermatid transcriptome. This processing of the poly(A) tail is typical of a cellular state of translational silencing and generally incomplete, leaving behind a short stretch of poly(A) residues. This is also the case in the oocyte, where the maternal RNAs are stored under a deadenylated form that will be readenylated upon recruitment for translation (Bachvarova *et al.* 1985, Huarte *et al.* 1992). Interestingly, microRNAs are known translational inhibitors, but have also recently been associated with directing the rapid deadenylation of mRNAs (Wu *et al.* 2006) reminiscent of the situation that prevails in the mature male gamete.

Therefore, these observations strongly suggest that the spermatid RNA population is composed mainly of naturally truncated mRNAs, which are at least partly adenylated. This is not typical of RNA turnover or decay since the general pathways involve exonucleases rather than endonucleases (Meyer *et al.* 2004). Such internal cleavage could be the signature of an unidentified nonspecific endonuclease or perhaps the RNA interference pathway, which is also triggered by microRNAs called small interfering RNAs (siRNA; Preall & Sontheimer 2005).

Aside from the transcriptional activity found in the mitochondria, the spermatozoon is transcriptionally inactive (Alcivar *et al.* 1989, Grunewald *et al.* 2005). Spermatid RNAs must therefore originate from the previous stages of spermatogenesis. Therefore, a comparative survey between spermatids and spermatozoa was undertaken to evaluate the level of similarity between the two populations of mRNAs. The objective of this analysis was clearly not of quantitative nature considering that equal amounts of starting RNA correspond to very different cell numbers between both cell types. The data generated by the microarrays was

analyzed by a statistical method inspired by Vallée *et al.* (2005). To be considered significant, the signal for a candidate had to be above a threshold value determined according to the fluorescence output of the negative controls printed on the microarray. Furthermore, to be included in the cell-type specific transcript list, the gene candidates had to be considered present based on the above threshold value in all four of the microarray hybridizations. The microarray slide used contains 19 200 human ESTs of which 2583 and 1117 positive sequences met the requirements to be qualified as present in the spermatids and the spermatozoa respectively. This detection rate is lower than the one reported earlier by about a twofold difference (Ostermeier *et al.* 2002, Zhao *et al.* 2006). This lower detection rate is attributed to the high stringency of the analysis rather than lack of transcript diversity in these cell types or the cross-species hybridization.

The microarray hybridizations were performed as cross-species between the bovine probes targeting human cDNAs. Previous reports of cross-species microarray hybridizations revealed that they are informative and precise because of their high overall correlation with homologous hybridizations (Adjaye *et al.* 2004, Shah *et al.* 2004, Vallée *et al.* 2005, 2006). In the study of Adjaye *et al.* (2004), only 5.7% of the candidates were shown to yield less reproducible hybridization results across the four replicates when compared with 4.0% for the human homologous hybridization.

The comparison of the transcript inventories between the two cell types revealed that most of the transcripts present in the spermatozoa are present in the spermatids. As Fig. 6 shows, a total of 121 transcripts were classified as specific to the spermatozoa relative to the spermatid. This is attributable to the high stringency of our analysis,

which is expected to produce very few false-positives, however, this level of confidence is counterbalanced by the loss of many true-positives. Indeed, for every one of the 121 transcripts classified as spermatozoa specific, at least one of spermatid hybridization generated a positive signal. When the raw data is analyzed individually, 98 out of 121 (81%) transcripts were present in three of the four spermatid hybridizations. Furthermore, 18 (15%) of the remaining spermatid-specific transcripts showed signals in two of the spermatid replicates and all of the five (4%) remaining candidates were detected in one of the spermatid hybridizations. Interestingly, the identities of these last five transcripts, which seem to be more prevalent in the spermatozoa relatively to the spermatid, remain unknown.

Confirmation of the tissue specificity by RT-PCR validated the results of the microarray analysis. The fact that none of the transcripts could be confirmed in the unamplified spermatid cDNA samples is indicative of a very low abundance of the targeted sequence in the spermatid RNA population. By contrast, the tissue specificity of all the spermatid candidates was directly validated in unamplified samples. However, the RT-PCR validation approach only allows to detect the presence of a targeted section of the mRNA thus, the amplicon represents a fragment of the mRNA sequence.

Recently, it was proposed that spermatid RNA could be an important factor determining the success of early embryonic development (Ostermeier *et al.* 2004). The survey of the bovine spermatid population reveals transcripts involved in a wide array of cellular functions. The absence of genes associated with cell cycle regulation in the spermatozoon is in agreement with the fact that the mature gamete does not require the cell cycle machinery. Similar reports in the human used different platforms to survey the spermatid transcriptome, either cDNA microarrays (Ostermeier *et al.* 2002) or serial analysis of gene expression (Zhao *et al.* 2006). They both also found a list of candidates involved in a very wide spectrum of cellular functions. Considering the limited capacity of the spermatozoon to store RNA relatively to the female gamete, a specific role of the spermatid RNA for early embryonic development could imply the targeting of a particular pathway by the accumulation of specific transcripts during the later stages of spermatogenesis. Otherwise, it is expected that an extremely low content of a heterogeneous mixture of transcripts would simply be diluted relatively to the large pools of stored maternal RNA also containing a similar diversity of transcripts (Robert *et al.* 2000). All the reported survey including the present report does not highlight such specificity in the spermatid RNA population.

The characterization of spermatid RNA indicates a clear fragmentation of the mRNA population that is a subset of the transcriptome found in previous stages of the spermatogenesis. The nature of the spermatid mRNA

population does not seem to target a specific metabolic pathway, but is rather a global representation of a wide array of basic cellular functions. This is by definition an overlap to the maternally stored RNAs pools that are known to support early development at least during the embryonic transcriptional silence period, which lasts until the 8 cell stage in the bovine (Barnes & First 1991, Memili *et al.* 1998). These facts in addition to the lack of RNA integrity do not lead to any clear evidence towards the potential to translate these as an essential paternal contribution to early embryonic development. An interesting recent publication has provided evidence of novel epigenetic influences of spermatid micro-RNAs originating from the fragmentation of cellular mRNA (Rassoulzadegan *et al.* 2006). Therefore, instead of the traditional translational fate of the mRNA, the contribution of the paternal RNA could rather be associated with epigenetic events. These novel roles of the paternal RNAs still remain to be validated. Finally, as the spermatid RNA profiles are associated with the previous steps of spermatogenesis, infertile phenotypes arising from an abnormal spermatogenesis could potentially be detected in the spermatid RNA population. The development of such a tool in species of economical interest such as the bovine could be used for the identification of fertility markers for animal breeding purposes.

Acknowledgements

Our thanks to Dr Pierre Leclerc for giving the spermatid isolation protocol and Dr Maurice Dufour for his help in sorting the spermatid cells. The authors also thank Steve Perrault for critical reading of the manuscript. This project was supported by a grant of the Fond Québécois de la Recherche sur la Nature et les Technologies (#107924) within the new investigator program. Some complementary funding was provided by DairyGen and NSERC to test the RNA extraction protocol. Isabelle Gilbert's wages were paid by an NSERC grant (#155182-05) supplemented by funds from Agriculture and Agri-Food Canada. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References

- Adjaye J, Herwig R, Herrmann D, Wruck W, Benkahl A, Brink TC, Nowak M, Carnwath JW, Hultschig C, Niemann H & Lehrach H 2004 Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics* **5** 83.
- Alcivar AA, Hake LE, Millette CF, Trasler JM & Hecht NB 1989 Mitochondrial gene expression in male germ cells of the mouse. *Developmental Biology* **135** 263–271.
- Bachvarova R, De Leon V, Johnson A, Kaplan G & Paynton BV 1985 Changes in total RNA, polyadenylated RNA, and actin mRNA during meiotic maturation of mouse oocytes. *Developmental Biology* **108** 325–331.
- Barnes FL & First NL 1991 Embryonic transcription in *in vitro* cultured bovine embryos. *Molecular Reproduction Development* **29** 117–123.

- Buffone MG, Doncel GF, Marin Briggiler CI, Vazquez-Levin MH & Calamera JC 2004 Human sperm subpopulations: relationship between functional quality and protein tyrosine phosphorylation. *Human Reproduction* **19** 139–146.
- Chomczynski P & Sacchi N 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry* **162** 156–159.
- Dadoune JP 2003 Expression of mammalian spermatozoal nucleoproteins. *Microscopy Research and Technique* **61** 56–75.
- Dadoune JP, Siffroi JP & Alfonsi MF 2004 Transcription in haploid male germ cells. *International Review of Cytology* **237** 1–56.
- Dadoune JP, Pawlak A, Alfonsi MF & Siffroi JP 2005 Identification of transcripts by macroarrays, RT-PCR and *in situ* hybridization in human ejaculate spermatozoa. *Molecular Human Reproduction* **11** 133–140.
- Eddy EM 2002 Male germ cell gene expression. *Recent Progress in Hormone Research* **57** 103–128.
- Grunewald S, Paasch U & Anderegg U 2005 Mature human spermatozoa do not transcribe novel RNA. *Andrologia* **37** 69–71.
- Gur Y & Breitbart H 2006 Mammalian sperm translate nuclear-encoded proteins by mitochondrial-type ribosomes. *Genes and Development* **20** 411–416.
- Huarte J, Stutz A, O'Connell ML, Gubler P, Belin D, Darrow AL, Strickland S & Vassalli JD 1992 Transient translational silencing by reversible mRNA deadenylation. *Cell* **69** 1021–1030.
- Kleene KC 1993 Multiple controls over the efficiency of translation of the mRNAs encoding transition proteins, protamines, and the mitochondrial capsule selenoprotein in late spermatids in mice. *Developmental Biology* **159** 720–731.
- Kleene KC 1996 Patterns of translational regulation in the mammalian testis. *Molecular Reproduction and Development* **43** 268–281.
- Kumar G, Patel D & Naz RK 1993 c-MYC mRNA is present in human sperm cells. *Cellular and Molecular Biological Research* **39** 111–117.
- Lambard S, Galeraud-Denis I, Martin G, Levy R, Chocat A & Carreau S 2004 Analysis and significance of mRNA in human ejaculated sperm from normozoospermic donors: relationship to sperm motility and capacitation. *Molecular Human Reproduction* **10** 535–541.
- Memili E, Dominko T & First NL 1998 Onset of transcription in bovine oocytes and preimplantation embryos. *Molecular Reproduction Development* **51** 36–41.
- Meyer S, Temme C & Wahle E 2004 Messenger RNA turnover in eukaryotes: pathways and enzymes. *Critical Reviews in Biochemistry and Molecular Biology* **39** 197–216.
- Morin G, Lalancette C, Sullivan R & Leclerc P 2005 Identification of the bull sperm p80 protein as a PH-20 ortholog and its modification during the epididymal transit. *Molecular Reproduction and Development* **71** 523–534.
- Ostermeier GC, Dix DJ, Miller D, Khatri P & Krawetz SA 2002 Spermatozoal RNA profiles of normal fertile men. *Lancet* **360** 772–777.
- Ostermeier GC, Miller D, Huntriss JD, Diamond MP & Krawetz SA 2004 Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature* **429** 154.
- Ostermeier GC, Goodrich RJ, Moldenhauer JS, Diamond MP & Krawetz SA 2005a A suite of novel human spermatozoal RNAs. *Journal of Andrology* **26** 70–74.
- Ostermeier GC, Goodrich RJ, Diamond MP, Dix DJ & Krawetz SA 2005b Toward using stable spermatozoal RNAs for prognostic assessment of male factor fertility. *Fertility and Sterility* **83** 1687–1694.
- Parrish JJ, Susko-Parrish J, Winer MA & First NL 1988 Capacitation of bovine sperm by heparin. *Biology of Reproduction* **38** 1171–1180.
- Pessot CA, Brito M, Figueroa J, Concha II, Yanez A & Burzio LO 1989 Presence of RNA in the sperm nucleus. *Biochemical and Biophysical Research Communications* **158** 272–278.
- Preall JB & Sontheimer EJ 2005 RNAi: RISC gets loaded. *Cell* **123** 543–545.
- Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I & Cuzin F 2006 RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441** 469–474.
- Robert C, Barnes FL, Hue I & Sirard MA 2000 Subtractive hybridization used to identify mRNA associated with the maturation of bovine oocytes. *Molecular Reproduction and Development* **57** 167–175.
- Schlecht U, Demougin P, Koch R, Hermida L, Wiederkehr C, Descombes P, Pineau C, Jegou B & Primig M 2004 Expression profiling of mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. *Molecular Biology of the Cell* **15** 1031–1043.
- Schmidt EE, Hanson ES & Capecchi MR 1999 Sequence-independent assembly of spermatid mRNAs into messenger ribonucleoprotein particles. *Molecular and Cellular Biology* **19** 3904–3915.
- Shah G, Azizian M, Bruch D, Mehta R & Kittur D 2004 Cross-species comparison of gene expression between human and porcine tissue, using single microarray platform. *Clinical Transplantation* **18** 76–80.
- Steger K 2003 Perspectives in the diagnosis of testicular biopsies using molecular biological techniques. *Andrologia* **35** 183.
- Sung K, Khan SA, Nawaz WS & Khan AA 2003 A simple and efficient X-100 boiling and chloroform extraction method of RNA isolation from Gram-positive and Gram-negative bacteria. *FEMS Microbiology Letters* **229** 97–101.
- Vallée M, Gravel C, Palin MF, Reghenas H, Stothard P, Wishart DS & Sirard MA 2005 Identification of novel and known oocyte-specific genes using complementary DNA subtraction and microarray analysis in three different species. *Biology of Reproduction* **73** 63–71.
- Vallée M, Robert C, Methot S, Palin MF & Sirard MA 2006 Cross-species hybridizations on a multi-species cDNA microarray to identify evolutionarily conserved genes expressed in oocytes. *BMC Genomics* **7** 113.
- Wrobel G & Primig M 2005 Mammalian male germ cells are fertile ground for expression profiling of sexual reproduction. *Reproduction* **129** 1–7.
- Wu L, Fan J & Belasco JG 2006 MicroRNAs direct rapid deadenylation of mRNA. *PNAS* **103** 4034–4039.
- Wykes SM, Visscher DW & Krawetz SA 1997 Haploid transcripts persist in mature human spermatozoa. *Molecular Human Reproduction* **3** 15–19.
- Zhao Y, Li Q, Yao C, Wang Z, Zhou Y, Wang Y, Liu L, Wang Y, Wang L & Qiao Z 2006 Characterization and quantification of mRNA transcripts in ejaculated spermatozoa of fertile men by serial analysis of gene expression. *Human Reproduction* **21** 1583–1590.

Received 23 October 2006

First decision 16 November 2006

Revised manuscript received 23 January 2007

Accepted 14 February 2007