

# Characterization of bovine transcripts preferentially expressed in testis and with a putative role in spermatogenesis

B.C.A. Alves<sup>a</sup>, P.R. Tobo<sup>a</sup>, R. Rodrigues<sup>a</sup>, J.C. Ruiz<sup>b</sup>, V.F.M. Hossepian de Lima<sup>c</sup>,  
C.A. Moreira-Filho<sup>d,\*</sup>

<sup>a</sup> Centro de Pesquisas em Biotecnologia, Universidade de São Paulo, 05508-900, São Paulo, SP, Brazil

<sup>b</sup> Centro de Pesquisas René Rachou, CPqRR, FIOCRUZ, 30190-002 Belo Horizonte, MG, Brazil

<sup>c</sup> Departamento de Reprodução Animal, Faculdade de Ciências Agrárias e Veterinárias, UNESP, 14884-900, Jaboticabal, SP, Brazil

<sup>d</sup> Departamento de Pediatria, Faculdade de Medicina da Universidade de São Paulo, 05403-900, São Paulo, SP, Brazil

Received 30 November 2010; received in revised form 27 April 2011; accepted 28 April 2011

## Abstract

Although the number of genes known to be associated with bovine spermatogenesis has increased in the past few years, regulation of this biological process remains poorly understood. Therefore, discovery of new male fertility genetic markers is of great value for assisted selection in commercially important cattle breeds, e.g., Nelore, that have delayed reproductive maturation and low fertility rates. The objective of the present study was to identify sequences associated with spermatogenesis that could be used as fertility markers. With RT-PCR, the following five transcripts preferentially expressed in adult testis were detected: TET<sub>656</sub> detected only in adult testis; TET<sub>868</sub> and TET<sub>515</sub> expressed preferentially in adult testis but also detected in fetal gonads of both sexes; and TET<sub>456</sub> and TET<sub>262</sub>, expressed primarily in the testis, but also present in very low amounts in somatic tissues. Based on their homologies and expression profiles, we inferred that they had putative roles in spermatogenesis. Detection of sequences differentially expressed in testis, ovary, or both, was a useful approach for identifying new genes related to bovine spermatogenesis. The data reported here contributed to discovery of gene pathways involved in bovine spermatogenesis, with potential for prediction of fertility.

© 2011 Elsevier Inc. All rights reserved.

**Keywords:** Spermatogenesis; Testis gene expression; Fertility marker; Differential gene expression; qPCR

## 1. Introduction

With the completion of the Bovine Genome Project, at least 22,000 protein-coding genes and more than 14,000 orthologous groups with correspondence in humans and other species were identified. Although many sequences were associated with placental function, fetal growth regulation, maternal adaptations to pregnancy and coordination of parturition, few genes were associated with male fertility [1]. There were apparently no associations be-

tween mutations in bovine testis-specific sequences so far identified and bull subfertility or infertility [2,3].

Spermatogenesis involves many strictly regulated testis-specific gene products. More than 50% of the mouse and rat genome are expressed during testicular development. However, among transcripts present in germ cells, 40–60% remain uncharacterized [4]. Moreover, genes are differentially expressed at various stages of murine and rat spermatogenesis, some of them in an alternative splicing form [5–7]. Despite the increasing number of genes associated with spermatogenesis, its regulation is still poorly understood. Therefore, a better knowledge of this process, and of testicular

\* Corresponding author. Tel.: 55 11 30698606; fax: 55 11 30618449.  
E-mail address: [carlos.moreira@icr.usp.br](mailto:carlos.moreira@icr.usp.br) (C.A. Moreira-Filho).

physiology, is crucial to identify mechanisms resulting in infertility and subfertility, and to understand processes regulating sperm production.

Male fertility markers would be of great value for assisted selection in commercially important cattle breeds like Nelore, which have delayed reproductive maturation [8] and low fertility [9]. In Brazil, Zebu cattle have a high rate (40%) of infertile or subfertile bulls in natural service [10] and bull culling represents a major economic loss. As a consequence, the search for bovine genes associated with male fertility—mainly with reproductive maturation and spermatogenesis—is of great relevance. Therefore, the current study focused on the search for bovine testis-associated transcripts through the detection of sequences differentially expressed between gonads and somatic tissues.

In our previous work, we identified (using RAPD analysis) two Y chromosome-specific genomic markers, OPA.06.3216 [11,12] and OPF10<sub>1168</sub> [13], which were usable for embryo sexing. Based on these two sequences, we designed several pairs of primers to determine whether they corresponded to functional genes (autosomal or sex-linked) by testing their ability to detect expressed sequences in testicular cDNA, using Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) and quantitative Polymerase Chain Reaction (qPCR). Consequently, in the present study, comparative developmental expression patterns of five bovine testis-derived sequences were identified and their putative role in spermatogenesis was discussed.

## 2. Materials and methods

### 2.1. Animals and tissue sampling

Testis, ovary, lung, heart, and liver were obtained after slaughter from 10 Nelore animals [three adult females,

three adult males, two fetal female (7 and 8 wk old, respectively) and two fetal males (7 and 9 wk old)]. Shortly after they were recovered, these tissue samples were placed in TRIzol solution (Invitrogen, Carlsbad, CA, USA).

### 2.2. Total RNA isolation and cDNA synthesis

Total RNA was isolated from tissues samples using TRIzol Reagent, according to the manufacturer's instruction. Concentration and quality assessment of total RNA was determined by spectrophotometry. To prevent genomic DNA contamination, total RNA was treated for 1 h at 37 °C with 10 U of deoxyribonuclease I (Invitrogen) per  $\mu\text{g}$  of RNA. Then, 3  $\mu\text{g}$  of total RNA was converted into cDNA using SuperScript® III First-Strand Synthesis System for RT-PCR (Invitrogen), according to the manufacturer's instructions. To assure that amplifications wasn't derived from any genomic DNA molecule, a negative control was synthesized for each cDNA sample under study (included all RNA-cDNA conversion reagents except reverse transcriptase).

### 2.3. Quantitative real-time PCR (qPCR)

Comparative expression analyses were performed by quantitative real-time polymerase chain reaction (qPCR). The qPCR primers and sequences are shown (Table 1). All samples were amplified in triplicate. Amplification reactions were performed in a 25  $\mu\text{L}$  final volume containing 1X SYBR Green mix (Quantitec SYBR Green PCR kit, QIAGEN, Hilden, Germany), 10 pmol of each primer, and 2  $\mu\text{L}$  cDNA (1/10 dilution, synthesized from 1  $\mu\text{g}$  of total RNA). Real time PCR amplifications were performed in Applied Biosystems 7300 Real Time PCR System with SDS v1.4 Software (Applied Biosystems, Foster City, CA, USA) with the following parameters: an initial hot start of 95 °C for 15

Table 1  
Primer sequences used in the current study.

Primer	Sequence (5-3)	Primer	Sequence (5-3)
MD1	GGCTGTCAGGGGTTGCTCCCTCAGGCA	MD1/2a	ATATCCTAGGCTGGGCTGGT
MD2	GGGATCCGGACGTACCGAACGA	MD1/2b	CGTACCGAACGAGAAGGAAG
MD3	CTTCCCATGGCCGTCTTTC	MD3/4a	TCTGAGCCCTCAAAGAAGGA
MD4	GTTTGAAGGCCGGAGAAGG	MD3/4b	ATTTGGCAAAGAAGGCACTG
MD5	TACCTGGTAGAGCTGTC	MD5/6a	CCAGGTCCCACATTCATTCT
MD6	GAGGGTTCAGGATGGG	MD5/6b	GCATGTGCAAGTGACTGTCT
MD7	CAGCCAGTTGTCTTGAG	MD7/6a	TGAGGGTGTGGAGGGTTG
MD8	GTTTATCCTAGTTGCCCTT	MD7/6b	TTCTGATGTGTTCCGCTCTT
MD8/6a	TGGAATGAGCATATGGAGCA	MD8/6b	TGGTGGATTATGTTGGTGT
$\beta$ -actin For	ACCGTGAGAAGATGACCCAG	B-actin Rev	AGGAAGGAAGGCTGGAAGAG
GAPDH For	GAAGACTGTGGATGGCCCTCC	GAPDH Rev	GTTGAGGGCAATGCCAGCCCC

min, followed by 55 cycles of 95 °C for 15 s, 57 °C for 25 s, and 72 °C for 30 s.

To normalize qPCR reactions, two reference genes were included (GAPDH and B-Actin). Although both reference genes gave similar results in normalization, the variance of the estimation of their quantification cycle (Cq) values was reduced by using the mean of both genes [14]. Relative expression was normalized using the median of two endogenous reference genes, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and  $\beta$ -actin (primers described in Table 1), determined by the relative standard curve method and expressed as fold change compared to its expression in adult testis (calibrator).

#### 2.4. Cloning and sequencing

The RT-PCR products amplified from adult testis cDNA were separated in 1.0% low melting point agarose gel. The band was excised and DNA was extracted using the Wizard® SV Gel and PCR Clean-Up System (Promega, Fitchburg, WI, USA; cat. no. A1330) according to the manufacturer's instructions. The PCR product was cloned using the kit TOPO TA Cloning® Kit for Sequencing, Version F (Invitrogen, cat. no. K4575-01) and sequenced using an ALFexpress DNA Sequencer (Pharmacia Biotech). Both strands of five selected clones from each fragment were sequenced using commercially available kits (ALFexpress Auto-Read Sequencing Kit, Pharmacia Biotech 27-2690-02 and Thermo Sequenase Fluorescent Labelled Primer Cycle Sequencing kit with 7-deaza-dGTP, Amersham Life Science, Buckinghamshire, England, RPN 2438) with vector's universal and reverse primers.

#### 2.5. Primer design

Specific primers were designed using Primer3 Input 0.4.0 software (available at <http://frodo.wi.mit.edu/primer3/>).

#### 2.6. Statistical analysis

A Student's *t*-test was used for comparing qPCR data from embryonic, female and male non-gonadal

tissues against adult testicular tissue, used as calibrator sample. Comparisons were made between one experimental group (i.e., embryonic testis) and the calibrator group, using GraphPad Prism, Version 5.03.

### 3. Results

Using RT-PCR, five transcripts preferentially expressed in adult testis were detected. These partial transcripts were herein identified by the acronym TET (standing for Testis-Enriched Transcript) and a subscript number designating their particular nucleotide extension as follows (GenBank accession numbers are shown within parentheses): TET<sub>868</sub> (HQ011401), TET<sub>585</sub> (HQ011404), TET<sub>656</sub> (HQ011405), TET<sub>456</sub> (HQ011403), and TET<sub>262</sub> (HQ011402). Subsequently, nested-RT-PCR assays were performed using internal primers designed according to each TET transcript nucleotide sequence. Afterwards, expression of each TET was quantified using specific internal primers and a panel of cDNA derived from fetal and adult male and female gonadal and somatic tissues (testis, ovary, heart, liver and lung) with quantitative real-time PCR. Primer sequences (Table 1) and external and internal primer pairs and their respective product lengths (Table 2) are shown.

Expression of TET<sub>656</sub> was restricted to adult testis, TET<sub>868</sub> and TET<sub>515</sub> expressions were restricted to gonads and TET<sub>456</sub> and TET<sub>262</sub> were ubiquitously expressed in very low amounts in somatic tissues. The TETs' tissue and developmental expression profiles obtained by qPCR are presented and discussed below. Since total RNA was isolated from whole gonadal tissue, it was not possible to discriminate between somatic and germ cells expression profiles. As described below, none of the TET sequences could be assigned to chromosome Y (or be considered as paralogs) on the basis of homology analysis. This analysis also revealed that each TET was almost certainly derived from a different gene.

Table 2  
External (RT-PCR) and internal (nested-RT-PCR and qPCR) primer pairs.

Transcript	External primer pair	Product length (bp)	Internal primer pair	Product length (bp)
TET <sub>585</sub>	MD1 and MD2	585	MD1/2a and MD1/2b	522
TET <sub>868</sub>	MD3 and MD4	868	MD3/4a and MD3/4b	272
TET <sub>656</sub>	MD5 and MD6	656	MD5/6a and MD5/6b	245
TET <sub>456</sub>	MD7 and MD6	456	MD7/6a and MD7/6b	351
TET <sub>262</sub>	MD8 and MD6	262	MD8/6a and MD8/6b	167

MD, male-derived.



Fig. 1. Representative RT-PCR amplification of  $TET_{656}$ . The L stands for 100 bp Ladder (Invitrogen Life Technologies Brazil). Amplification of  $TET_{656}$  was present in the adult testis (1), but absent in fetal testis (2), adult ovary (3), fetal ovary (4), adult lung (5), adult liver (6) and adult heart (7). The amplification negative control was in lane 8.

### 3.1. Testis-specific transcript $TET_{656}$

Nested RT-PCR revealed that  $TET_{656}$  expression was restricted to the adult testis (Fig. 1). A BLAST search against NCBI database [15] indicated that  $TET_{656}$  had along 84% of its extension 75% similarity with the hypothetical mRNAs LOC786421 (accession no. XM\_001254101), LOC100337398 (accession no. XM\_002693914.1) and LOC783947 (accession no. XM\_001250809.2), all mapped to bovine chromosome 16.

### 3.2. Gonad-specific transcripts $TET_{585}$ and $TET_{868}$

Based on nested RT-PCR,  $TET_{585}$  and  $TET_{868}$  expression were restricted to gonadal tissue. The qPCR reactions detected  $TET_{585}$  and  $TET_{868}$  expression in both testis and ovary. Significant expression fold changes of  $TET_{585}$  and  $TET_{868}$  in adult testis relative to fetal testis and the fetal and adult ovary are shown (Table 3 and Fig. 2).

A BLAST search revealed that 86% of  $TET_{585}$  sequence was 96% similar to *Bos taurus* Ewing sarcoma breakpoint region 1 (*EWSR1*) mRNA (accession no. NM\_001109800.1), located in chromosome 17. Likewise,  $TET_{585}$  had homology with two *Gallus gallus* cloned sequences CD218853 and BM439438, which are highly similar to human *EWS*. Those avian clones were generated from a cDNA library synthesized from a total RNA pool derived from testis, ovary, and oviduct.

Regarding  $TET_{868}$ , BLAST search revealed that 98% of its sequence was highly similar (99%) to bovine *AFF1*, a member of *AF4/FMR2* family mapped in chromosome 6 (also known as *Af4*, *Rob* or *Mllt2h*).

### 3.3. Preferentially expressed in testis transcripts $TET_{262}$ and $TET_{456}$

Nested RT-PCR and qPCR revealed that  $TET_{262}$  and  $TET_{456}$  expression was ubiquitous, but preferentially occurred in testis (Table 3 and Fig. 3). A BLAST search revealed high similarity (94%) of  $TET_{262}$  with a *Bos taurus* chromosome 27 genomic contig (accession no.

NW\_001494406.2). This contig has no known function annotation. Based on a BLAST search with  $TET_{456}$  using the Cow Sequence database, 98% of this transcript sequence had extensive similarity with E2F transcription factor 5, p130-binding *E2F5* (98%) mapped on chromosome 14.

## 4. Discussion

Detection of sequences differentially expressed in testis and/or ovary is a useful approach for identifying new factors related to bovine spermatogenesis. Studies to characterize germline gene expression in *C. elegans* and *Drosophila* using DNA microarray and computational analyses demonstrated that oogenesis and spermatogenesis had several differences. In that regard, considering all genes expressed in the *C. elegans* germ lineage, 46% are testis-specific, 18% are ovary-specific, and 36% are expressed in both gonadal tissues [16]. Furthermore, in *D. melanogaster*, more than 80% of the

Table 3  
TET sequences significant expression fold changes.

Sequence	Tissue	Fold difference <sup>a,b</sup>
$TET_{585}$	Adult testis	1.00 ± 0.10
	Fetal testis	-6.70 ± 1.08
	Adult ovary	-1.61 ± 0.12
	Fetal ovary	-35.80 ± 2.53
$TET_{868}$	Adult testis	1.00 ± 0.33
	Fetal testis	-4.95 ± 0.91
	Adult ovary	-5.08 ± 0.64
	Fetal ovary	-37.60 ± 2.65
$TET_{262}$	Adult testis	1.00 ± 0.10
	Fetal testis	-3.69 ± 1.21
	Adult ovary	-30.56 ± 2.30
	Fetal ovary	-39.16 ± 3.52
	Adult liver	-26.38 ± 1.50
	Adult lung	-32.13 ± 1.39
	Adult heart	-41.08 ± 1.36
	Fetal liver	-32.55 ± 2.76
	Fetal lung	-34.81 ± 1.42
Fetal heart	-39.78 ± 1.31	
$TET_{456}$	Adult testis	1.00 ± 0.13
	Fetal testis	-11.10 ± 1.90
	Adult ovary	-4.50 ± 0.26
	Fetal ovary	-5.00 ± 0.50
	Adult liver	-5.40 ± 0.21
	Adult lung	-3.50 ± 0.51
	Adult heart	-7.70 ± 0.19
	Fetal liver	-3.60 ± 0.20
	Fetal lung	-3.90 ± 0.13
Fetal heart	-3.50 ± 0.27	

<sup>a</sup> P < 0.0001 for all samples, each one compared to adult testis (calibrator sample).

<sup>b</sup> SD.

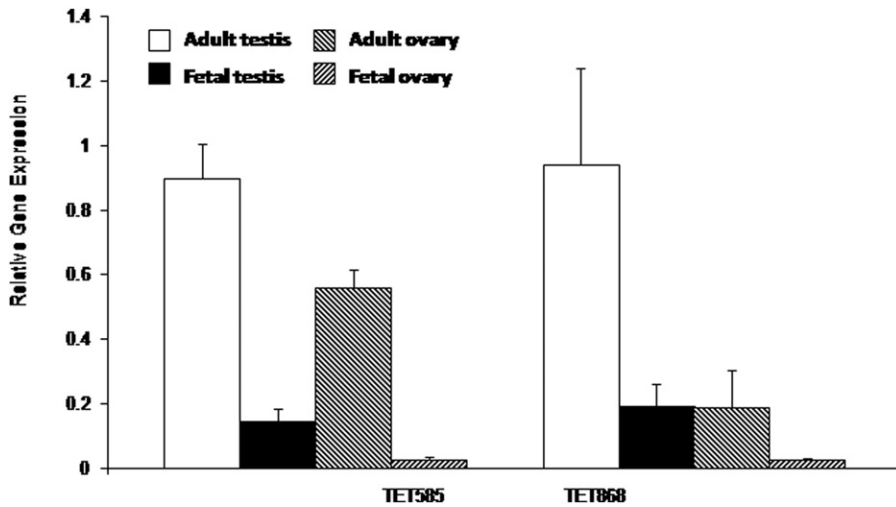


Fig. 2. Relative expression of TET<sub>585</sub> and TET<sub>868</sub>. In all cases, expression fold changes differed ( $P < 0.0001$ ).

genes expressed in the testis had doubled expression levels relative to their expression in ovaries [17]. In mice, genes involved in spermatogenesis had testis-specific, gonadal or ubiquitous expression patterns, depending on various regulatory factors and on particular periods during sexual maturation [18,19]. In the present study, sequences specific or preferentially expressed in testis with a putative role in spermatogenesis were identified.

#### 4.1. Testis-specific sequence

Although TET<sub>656</sub> had similarity with some hypothetical proteins whose functions were not yet described, it was noteworthy that this transcript was expressed only in adult testis. Based on this restricted expression profile, we inferred that it had a role in spermatogenesis, since it was not found in fetal testis.

#### 4.2. Gonad-specific sequences

##### 4.2.1. TET<sub>868</sub>

This transcript had high similarity with the bovine gene *AFF1*, a member of the *AF4/FMR2*. This gene family codes for nuclear proteins known as tissue-specific transcript activators [20]. According to the Human Genome Nomenclature Committee (HGNC) [21], *AFF1* is also known as *AF4* [22]. In mice, the developmental expression pattern of *AF4* suggests a role in the development of hematopoietic, cardiovascular, skeletal and central nervous systems. Indeed, murine *AF4* expression was already found in kidney, brain, lung, liver, spleen, skeletal muscle, and testis [23]. More specific studies in gonadal tissues revealed the presence of this transcript in ovary, epididymis and testis. Testicular expression can be found more specif-

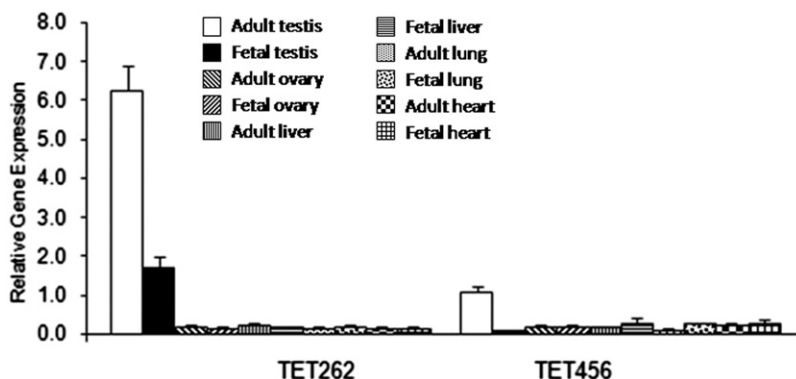


Fig. 3. Relative expression of TET<sub>262</sub> and TET<sub>456</sub>. In all cases, expression fold changes differed ( $P < 0.0001$ ).



ically in type A spermatogonia, thus in the beginning of spermatogenesis [5,24–26].

Based on a phylogenetic analysis of *AF4* family members from various species, *AFF1* and *AFF4* are sister clades, with bootstrap value of 100% [22]. It is noteworthy that *AFF4*, also known as *AF5q31* is another member of *AF4/FMR2* gene family that has a specific role during spermatogenesis, since knockout mice for this gene are azoospermic [27]. Mutant *AF5q31*<sup>-/-</sup> mice have embryonic and neonatal lethality phenotype. Surviving *AF5q31*<sup>-/-</sup> male mice are infertile; their seminal fluid was devoid of mature sperm, indicating an arrest of spermatogenesis, but female *AF5q31*<sup>-/-</sup> mice are fertile. These mutant mice have small testes and are azoospermic. So far, *AF5q31* is the only member of its family known to be expressed in testis [27]. Based on these findings, we inferred that *AF5q31* was much more relevant to spermatogenesis than to oogenesis, since it was expressed in a higher rate in testes than in ovary, and mutant males for this gene are infertile.

Although *AF4* family members had already been identified in cows [22], their expression pattern have not been described. Despite the presence of *AFF1* presence in several murine somatic tissues, in cattle, TET<sub>868</sub> expression was detected only in gonadal tissues, with an adult testicular expression five-fold higher than in adult ovary, more than four-fold higher than in fetal testis, and more than 30-fold higher than in the fetal ovary. Based on detection of TET<sub>868</sub> in fetal testis, its expression occurred in interstitial cells, since fetal testis germ cells are not yet functional. Moreover, based on the expression fold change between fetal and adult testis, plus the homologies described, we inferred that this transcript was also expressed in type A spermatogonia and acted as a transcription factor in the beginning of spermatogenesis.

#### 4.2.2. TET<sub>585</sub>

This transcript was similar to human *EWSR1*, a gene that codes for a nuclear transcription regulatory protein [28]. Kawano et al [29] were the first to demonstrate by immunohistochemistry and western blotting that *EWS* expression was more abundant in rat testis than in somatic tissues. The *EWSR1* gene is very conserved among chimpanzee, dog, cow, mouse, rat, chicken and zebrafish; despite the ubiquitous expression pattern of *EWSR1* in human [30], TET<sub>585</sub> was detected only in gonadal tissue, with preferential expression in the adult testis. Based on microarray assays of rat and mouse tissues, the transcript *Ewsr1* was present in testis, seminiferous tubule, germ cells (spermatogonia, spermatocyte and spermatids), in Sertoli cells, and in the epididymis [7], as well as in the mouse ovary [6]. The expression profile of rat and murine *Ewsr1* in gonads is shown (Mammalian Reproductive Genetics website [26]). So far, no mutations in this gene were associated with alterations in spermatogenesis or infertility.

#### 4.3. Sequences highly expressed in adult testis, but detected in both gonadal and somatic tissues

##### 4.3.1. TET<sub>456</sub>

This sequence was homologous to *E2F5*, a member of *E2F* family of transcription factors. Transcripts for *E2F* are ubiquitously expressed in human tissues and are required for activation or repression of differentially expressed gene pathways along the cell cycle [31] through CpG methylation [32]. Members of the *E2F* family are differentially expressed in rodent gonadal cell types. In that regard, *E2F5* was expressed throughout meiotic cycle [31]. In mice, *E2F5* expression was identified in whole epididymis, in testis and also in ovary. In the epididymis, where sperm become fertile, its expression was detected in all segments [5]. In the human testis, *E2F5* was identified in seminiferous tubules, germ cells, types A and B spermatogonia, and in round spermatids, where it was higher expressed [19]. In the human ovary, this transcription factor was expressed in the ovarian cortex, follicle, primordial follicle, primary follicle and secondary follicle [24]. In the rat, *E2F5* was expressed in testis (seminiferous tubule, germ cell, spermatogonia, spermatocyte, spermatid and Sertoli cell) and in epididymis [6]. The murine and rat gonadal *E2F5* expression patterns have been reported (Mammalian Reproductive Genetics database [26]). Since TET<sub>456</sub> was highly expressed in the testis and shared homology with the *E2F5*, perhaps it had a role in the temporal expression control of genes related to bovine spermatogenesis.

##### 4.3.2. TET<sub>262</sub>

This transcript has high similarity with of a *Bos taurus* chromosome 27 genomic contig segment. However, no function has been annotated.

## 5. Conclusion

Independent of their expression profiles, the five transcripts described had clear preferential expression in adult testis, consistent with a putative role in spermatogenesis. That TET<sub>456</sub>, TET<sub>585</sub> and TET<sub>868</sub> had homologies with known genes yielded clues regarding their functions, but corroboration will only be achieved with characterization of the entire transcript. Con-

versely, TET<sub>262</sub> and TET<sub>656</sub> had similarity with hypothetical proteins or with nonannotated sequences, yielding no clues regarding their functions. Perhaps they represent novel genes with an unexpected role in spermatogenesis. If so, they could provide new avenues for investigating molecular mechanisms underlying delayed maturation and low fertility in Nelore bulls.

In order to use the identified sequences as markers of fertility, it will be first necessary to characterize the full transcripts corresponding to our TET sequences and, furthermore, their complete genomic sequences. This step is utterly important for confirming alternative transcription and/or sequence polymorphisms determining alterations in spermatogenesis. Considering the high evolutionary conservation herein for TET sequences, the present study could facilitate discovery of potential genomic markers useful not only for Nelore fertility management, but also for other domestic breeds.

## Acknowledgments

This work was partially supported by FAPESP grant no. 04/14738-2 (BCA Alves post-doctoral fellowship), CNPq grants 305635/2009-3 and 300996/2006-7 (CAMF) and by FAPEMIG/CNPq (JC Ruiz). BCAA participated in the design of the study, carried out the identification and molecular characterization of all TET sequences and drafted the manuscript; PRT and RR participated in the isolation of genetic material and qPCR experiments; JCR carried out bioinformatics analysis and sequence annotation; VFMHL provided all tissue samples and participated in the design of the study; CAMF participated in the design of the study, in the analysis of the results, and helped to draft the manuscript.

## References

- [1] The Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 2009;324:522–8.
- [2] Jakubiczka S, Schnieders F, Schmidtke J. A bovine homologue of the human TSPY gene. *Genomics* 1993;17:732–5.
- [3] Delbridge ML, Harry JL, Toder R, O'Neill RJW, Ma K, Chandley AC, Graves JAM. A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome. *Nat Genet* 1997;15:131–6.
- [4] Lee TL, Pang AL, Rennert OM, Chan WY. Genomic landscape of developing male germ cells. *Birth Defects Res C Embryo Today* 2009;87:43–63.
- [5] Johnston DS, Jelinsky SA, Bang HJ, Di Candeloro P, Wilson E, Kopf GS, Turner TT. The mouse epididymal transcriptome: transcriptional profiling of segmental gene expression in the epididymis. *Biol Reprod* 2005;73:404–13.
- [6] Jelinsky SA, Turner TT, Bang HJ, Finger JN, Solarz MK, Wilson E, Brown EL, Kopf GS, Johnston DS. The rat epididymal transcriptome: comparison of segmental gene expression in the rat and mouse epididymides. *Biol Reprod* 2007;76:561–70.
- [7] Johnston DS, Jelinsky AS, Zhi Y, Finger JN, Kopf GS, Wright WW. Identification of testis-specific male contraceptive targets: insights from transcriptional profiling of the cycle of the rat seminiferous epithelium and purified testicular cells. *Ann NY Acad Sci* 2007;1120:36–46.
- [8] Alves BCA, Unanian MM, Silva E, Oliveira M, Moreira-Filho CA. Use of RAPD markers for identifying Nelore bulls with early reproductive maturation onset. *Anim Reprod Sci* 2005;85:183–91.
- [9] Pineda NR. Motivar para a real precocidade sexual. *Revista ABCZ* 1(3), 2001. Available at: [www.abcz.org.br/revista/2001](http://www.abcz.org.br/revista/2001). Accessed 20 February 2010.
- [10] Sereno JRB. Utilização Racional de Touros em Monta Natural, 2004. Available at: <http://www.agronline.com.br/artigos/artigo.php?id=139>. Accessed 18 February 2011.
- [11] Alves BCA, Hossepian de Lima VFM, Teixeira CM, Moreira-Filho CA. Use of primers derived from a new sequence of the bovine Y chromosome for sexing *Bos taurus* and *Bos indicus* embryos. *Theriogenology* 2003;59:1415–9.
- [12] Alves BCA, Mayer MG, Taber AP, Egito AA, Fagundes V, McElreavey K, Moreira-Filho CA. Molecular characterization of a bovine Y-specific DNA sequence conserved in taurine and zebu breeds. *DNA Seq* 2006;17:199–202.
- [13] Alves BCA, Hossepian de Lima VFM, Moreira-Filho CA. Development of Y-chromosome-specific SCAR markers conserved in taurine, zebu and bubaline cattle. *Reprod Domest Anim* 2010;45:1047–51.
- [14] Toussaint J, Sieuwerts AM, Haibe-Kains B, Desmedt C, Rouas G, Harris AL, Larsimont D, Piccart M, Foekens JA, Durbecq V, Sotiriou C. Improvement of the clinical applicability of the Genomic Grade Index through a qRT-PCR test performed on frozen and formalin-fixed paraffin-embedded tissues. *BMC Genomics* 2009;10:424–36.
- [15] BLAST: Basic Local Alignment Search Tool. Available at: <http://blast.ncbi.nlm.nih.gov/>. Accessed 18 February 2011.
- [16] Reinke V, Smith H, Nance J, Wang J, Van Doren C, Begley R, Jones S, Davis E, Scherer S, Ward S, Kim S. A global profile of germline gene expression in *C. elegans*. *Mol Cell* 2000;6:605–16.
- [17] Andrews J, Bouffard G, Cheadle C, Lu J, Becker K, Oliver B. Gene discovery using computation and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res* 2000;10:2030–43.
- [18] Eddy EM. Male germ cell gene expression. *Recent Prog Horm Res* 2002;57:103–28.
- [19] Shima JE, McLean DJ, McCarrey JR, Griswold MD. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod* 2004;71:319–30.
- [20] Gu Y, Shen Y, Gibbs RA, Nelson DL. Identification of FRM2, a novel gene associated with FRAXE CCG repeat and CpG island. *Nat Genet* 1996;13:109–13.
- [21] HUGO Gene Nomenclature Committee. Available at <http://www.genenames.org>. Accessed 18 February 2011.

- [22] Niedzielski MF, Hopewell R, Ismail Z, Estable MC. MCEF is localized to the nucleus by protein sequences encoded within three distinct exons, where it represses HIV-1 Tat-transactivation of LTR-directed transcription. *Int J Biol Sci* 2007;3:225–36.
- [23] Baskaran K, Erfurth F, TAborn G, Copeland NG, Gilbert DJ, Jenkins NA, Iannaccone PM, Domer PH. Cloning and developmental expression of the murine homolog of the acute leukemia proto-oncogene AF4. *Oncogene* 1997;15:1967–78.
- [24] Pan H, O'Brien MJ, Wigglesworth K, Eppig JJ, Schultz RM. Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. *Dev Biol* 2005;286:493–506.
- [25] McCarrey JR, O'Brien DA, Skinner MK. Construction and preliminary characterization of a series of mouse and rat testis cDNA libraries. *J Androl* 1999;20:635–39.
- [26] Mammalian Reproductive Genetics. Available at: <http://www.mrg.genetics.washington.edu>. Accessed 18 February 2011.
- [27] Urano A, Endoh M, Wada T, Morikawa Y, Itoh M, Kataoka Y, Taki T, Akazawa H, Nakajima H, Komuro I, Yoshida N, Hayashi Y, Handa H, Kitamura T, Nosaka T. Infertility with defective spermiogenesis in mice lacking AF5q31, the target of chromosomal translocation in human infant leukemia. *Mol Cell Biol* 2005;25:6834–45.
- [28] Plougastel B, Zucman J, Peter M, Thomas G, Delattre O. Genomic structure of the EWS gene and its relationship to EWSR1, a site of tumor-associated chromosome translocation. *Genomics* 1993;18:609–15.
- [29] Kawano J, Nakayama T, Takami Y, Kotani T, Sawaguchi A, Nagaike R, Oinuma T, Suganuma T. A monoclonal antibody against insect CALNUC recognizes the prooncprotein EWS specifically in mammalian cells. Immunohistochemical and biochemical studies of the antigen in rat tissues. *Histochem Cell Biol* 2001;115:421–8.
- [30] GeneCards, the Human gene compendium. Available at <http://www.genecards.org>. Accessed 18 February 2011.
- [31] El-Darwish KS, Parvinen M, Toppari J. Differential expression of members of the E2F family of transcription factors in rodent testes. *Reprod Biol Endocrinol* 2006;4:63–84.
- [32] Campanero MR, Armstrong MI, Flemington EK. CpG methylation as a mechanism for the regulation of E2F activity. *Proc Natl Acad Sci U S A* 2000;12:6481–6.