DOI: 10.21451/1984-3143-AR816

Anim. Reprod., v.14, n.2, p.418-428, Apr./Jun. 2017

Isolation and identification of proteins from swine sperm chromatin and nuclear matrix

Guilherme Arantes Mendonça[1,3], Romualdo Morandi Filho[2], Elisson Terêncio Souza[2], Thais Schwarz Gaggini[1], Marina Cruvinel Assunção Silva-Mendonça[1], Robson Carlos Antunes[1], Marcelo Emílio Beletti[1,2]

[1]Post-graduation Program in Veterinary Science, Federal University of Uberlandia, Uberlandia, MG, Brazil.
[2]Post-graduation Program in Cellular and Molecular Biology, Federal University of Uberlandia, Uberlandia, MG, Brazil.

Abstract

The aim of this study was to perform a proteomic analysis to isolate and identify proteins from the swine sperm nuclear matrix to contribute to a database of swine sperm nuclear proteins. We used pre-chilled diluted semen from seven boars (19 to 24 week-old) from the commercial line Landrace x Large White x Pietran. The semen was processed to separate the sperm heads and extract the chromatin and nuclear matrix for protein quantification and analysis by mass spectrometry, by LTQ Orbitrap ELITE mass spectrometer (Thermo-Finnigan) coupled to a nanoflow chromatography system (LC-MS/MS). We identified 222 different proteins in the sample; a total of 159 (71.6%) were previously described as present in the somatic or sperm nuclei of other species, 41 (18.5%) did not have a previously reported nuclear presence and 22 (9.9%) had not been characterized. The most abundant family of proteins corresponded to ribosomal (13.1%), followed by cytoskeleton (12.2%), uncharacterized (9.9%), histones (5.4%), proteasome subunits (3.6%) and heat shock (1.8%). The other proteins clustered in other families accounted for 54% of the total proteins. The protein isolation of the nuclear matrix of the swine spermatozoa was satisfactory, thus demonstrating that the protocol used was efficient. Several proteins were identified and described. However, it was not possible to identify some protein structures. Therefore, this study helps to establish a starting point for future proteomic studies comparing fertile and sub-fertile animals.

Keywords: epigenetic, mass spectometry, proteome, *sus scrofa.*

Introduction

It has long been thought that the only function of sperm cells is to transmit the paternal genomic DNA to the next generation. This idea was challenged by the discovery of the imprinting of sex-specific genes mediated by DNA methylation differences (set during gametogenesis) that were epigenetically transmitted to the next generation (Oliva and Ballescá, 2012).

DNA condensation by sperm protamines leaves only a small fraction of the sperm genome accessible for DNA binding proteins, which are necessary to enable DNA replication and genes transcription. These sites may be the most important sites for the initiation of paternal genome functions in the early embryo (Yamauchi *et al*., 2011). According to the same authors, these active sperm chromatin sites in protamine toroids may contain important epigenetic information for the developing embryo.

The isolated use of genomic and transcriptomic information may be insufficient to fully understand a complex organism because proteomics and transcriptomics can be discordant and DNA-RNA relationships cannot be fully correlated. Thus, measurements of other metabolic levels should also be obtained, such as the study of proteins (Wright *et al*., 2012). According to these same authors, large-scale protein research in organisms (i.e., the proteome-protein complement expressed by a genome) is equally important because it provides information about the real factors (i.e., enzymes) involved in the metabolic process. However, unlike other areas (i.e., genomics and transcriptomics), proteomics and its present techniques and strategies are still under development.

Proteomics projects related to studies of nuclear proteins in sperm have enabled the creation of catalogs. However, DeMateo *et al*. (2011) related that only small subsets of the identified proteins are nuclear proteins. The aim of this study was to perform a proteomic analysis to isolate and identify proteins from the swine sperm nuclear matrix to contribute to a database of swine sperm nuclear proteins.

Materials and Methods

Semen and sperm processing

Prediluted (diluent BTS) and cooled boar semen was used in this study, with concentration of 2.5 x 10^9 (80 ml/dose). The doses were stored in thermal chamber between 15 and 18°C. The semen was provided by an artificial insemination center located in Uberlandia, Minas Gerais, Brazil. We used semen from seven boars (19 to 24 week-old) from the commercial line Landrace x Large White x Pietran. The period between the collection and analysis ranges from 24 to 48 h. The breeders chosen were normally used by insemination center, where the resulting litters, borned to the date of collection, were within zootechnical levels proposed for the line in question.

Based on routine testing using the CASA system (computerized analysis of sperm), the boars presented semen with 84.63 to 93.54% of motility and 58.78 to 84.11% of progressive motility. We use the standard method for evaluation of sperm morphology,

[3]Corresponding author: guilherme.mendonca@outlook.com
Received: March 2, 2016
Accepted: March 23, 2017

which consisted of 200 cell count at 1000X magnification in phase contrast microscopy with immersion oil. It was observed defects of head, tail and curled tail, acrosome, midpiece, colon and proximal and distal citoplasmic drop. The rate of morphological defects (apart from the distal cytoplasmic drop) was 2.5 to 10.5%.

Analysis of chromatin alterations using semen smears stained with toluidine blue

All semen was evaluated by the toluidine blue technique (Beletti *et al*., 2005) to determine whether the boars in question had normal sperm with respect to condensation of chromatin and the morphology of the sperm head.

The samples of boar semen were fixed using formalin citrate (2 drops of semen to 1 ml of formaldehyde citrate). A drop of sample was fixed in two smears and subsequently air dried at room temperature. These smears were subjected to acid hydrolysis in 4N hydrochloric acid for 20 min and washed in distilled water. After drying, the smears were stained with a drop of 0.025% toluidine blue (pH 4.0) in phosphate-citric acid buffer (McIlvaine buffer) on the slide, followed by placement of the coverslip. After three minutes, 50 digital images in grayscale were captured from each slide using a Leica DM500 optical microscope coupled with a Leica ICC 50 camera (Wetzlar, Germany) with an oil immersion lens at 100X magnification. Digital images were used to segment by thresholding 100 sperm heads from each slide.

The samples were analyzed using routines developed in the SCILAB environment to obtain the mean and standard deviation of the pixel values within the head of each image. To obtain a reference for the normal color of the sperm head, six sperm heads that were the most homogeneous and had a lighter color (i.e., the sperm was more homogeneous and intensely compacted) were automatically selected in each smear. The average pixel values of these heads were used as the reference value for the normal staining of the sperm (standard head). Then, the differences between the average values of the standard heads and the average values of each head examined were determined for each image. This difference was transformed into a percentage (% unpacking) based on the average value of the standard heads. The coefficient of variation (heterogeneity %) of the gray levels was also calculated (Beletti *et al*., 2005).

Sperm head segregation

The methodology used to segregate the sperm heads was modified from a previous study (Morandi-Filho, 2013). Each semen sample (4 ml) cooled to 2 to 8 degrees Celsius (C) was placed in 15 ml conical bottom tubes containing 8 ml of buffer (50 mM Tris-HCl, pH 7.5, and 1 mM EDTA). The flask was homogenized and centrifuged at 750 x g for 15 min at 4°C, followed by removal of the supernatant. The pellet was resuspended in 8 ml of the same buffer, homogenized and centrifuged again. This procedure was repeated three times.

After the third centrifugation, the pellet was resuspended in 1.5 ml of the same buffer. The material was sonicated on ice for 10 min with 30 s pulses and intervals of 5 s. Subsequently, the material was centrifuged at 1000 x g for 15 min at 4°C, the supernatant was removed, and 2 ml of buffer (50 mM Tris-HCl and 1.1 M saccharose, pH 7.5) was added.

Part of this material was diluted 1:100 in distilled water for counting in a Neubauer chamber to measure the concentration of the heads in the sample. Then, the concentration was adjusted to 1 x 10^7 head/ml using 50 mM Tris-HCl buffer with 1.1 M saccharose (pH 7.5).

The heads were isolated from the tails using ultracentrifugation at 75,600 x g for 45 min at 4°C in a gradient consisting of 2 ml of cesium chloride (2.82 M cesium, 25 mM Tris-EDTA, 5 mM MgCl2 and 0.5% Triton X-100) at the bottom of a 12 ml ultracentrifuge tube that was overlaid with 4 ml of 2.2 M saccharose and covered with 2 ml of the sample in 50 mM Tris-HCl and 1.1 M saccharose (pH 7.5). After centrifugation, the supernatant containing the tails was carefully removed by pipetting. The bottom sediment was resuspended in 25 mM Tris buffer and washed three times by centrifugation at 1000 x g for 30 min at 4°C in 25 mM Tris buffer to remove the excess cesium chloride. After this process, a smear was made of a drop of the sample. The smear was dried in an oven for 15 min, stained with xylidine for 15 min and washed with distilled water to evaluate the purity of the sample relative to the absence of tails. The purity was approximately 95% according to visual evaluation by light microscopy based on counting 100 cells in a field.

Extraction of chromatin and the nuclear matrix

The extraction of chromatin and the nuclear matrix followed the methodology adapted from Codrington *et al*. (2007). The isolated heads were resuspended in 500 µl of a solution containing 1% Triton X-100, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 5 µl of protease inhibitor cocktail (Sigma Aldrich P8340/011M4000) and vortexed for 10 min at room temperature. This treatment removed the acrosome and all membranes while leaving the nucleus condensed and connected to the nuclear envelope and some vestige of the perinuclear material.

The samples were washed three times by centrifugation at 1100 x g for 30 min with 1.5 ml of 50 mM Tris-HCl (pH 7.5). After the last wash, the material was resuspended in 500 µl of decondensation buffer consisting of 40 mM 1,4-dithiothreitol (DTT), 0.25 M $(NH_4)_2S_4$, 25 mM Tris-HCl (pH 7.5), and 5 µl of a protease inhibitor cocktail and incubated for 40 min at room temperature. Then, 4000U of RNAse-free deoxyribonuclease I was added and the samples were homogenized for 60 min under a vortex at room temperature. Finally, the samples was frozen, lyophilized and stored in a freezer prior to processing for mass spectrometry.

Protein quantification

At this point the lyophilized samples were mixed, and from here they were processed as single sample. Initially, the lyophilized sample was resuspended in 100 µl of 0.1 M Tris-HCl buffer (pH 8.8) containing 8 M urea. The Bradford method (Bradford, 1976) with the Protein Assay Dye Concentrate Reagent (Bio-Rad, Hercules, California, USA) was used for protein quantification. The standard curve was performed using different dilutions of bovine serum albumin prepared from a commercially acquired stock (200 mg/ml protein standard, Sigma, St. Louis, Missouri, USA). The sample was distributed in triplicate in microplates. The absorbance at 595 nm was read in a spectrophotometer (Molecular Devices, SpectraMax Plus 384). The quantification of protein by the Bradford method indicated concentration of 3.4 mg/ml.

Sample preparation

The sample preparation for mass spectrometry consisted of three main steps: i) reduction and alkylation of proteins, ii) enzymatic digestion of the proteins with trypsin and iii) clean up/desalting of the samples. We used 38 µl of each sample (50 µg). Briefly, the sample was subjected to the reduction of disulfide bonds of the protein by the addition of DTT (dithiothreitol) in a proportion of 1 mg DTT/mg protein and incubated for 2 h at room temperature. Then, alkylating IA (iodoacetamide) was added in a proportion of 3 mg IA/mg of protein and incubated for 1 h at room temperature in the dark. The volume of the sample was diluted 5-fold in a 0.1 M solution of ammonium bicarbonate (pH≥ 8,0) to obtain a final volume of 500 µl. The sample was incubated with 1 µg of trypsin (Promega, Madison, Wiscousin, USA) at 37°C overnight. Prior to application of the sample into the mass spectrometer, clean-up/desalting of the sample was performed using the OASIS HLB cartridge 1 cc column according to the manufacturer's instructions. The column was equilibrated with a 5% acetonitrile solution containing 0.1% formic acid, and elution of the material of interest was performed with 80% acetonitrile. The sample was dried in a speed vac and applied to a mass spectrometer.

Mass spectrometry analysis

The digested sample was dried and analyzed in the LTQ Orbitrap ELITE mass spectrometer (Thermo-Finnigan) coupled to a nanoflow chromatography system (LC-MS/MS). The acquired data were automatically processed by the Computational Proteomics Analysis System (CPAS; Rauch, 2006). The identified peptides were grouped into proteins using the algorithm Protein Prophet, and a list of identifications with error rates less than 2.0% was created. A general database of all species was used (Uniprot, 2016).

Statistical analysis

Descriptive statistical analyses were performed on the presented data.

Results

Toluidine blue method

The toluidine blue method was used to analyze the chromatin in 195 sperm heads of each boar. The averages were: unpacking chromatin (%) 2.46 ± 1.73 and heterogeneity of chromatin (%) 4.49 ± 0.94.

Proteins found

In the mass spectrometry analysis, 222 different proteins were identified in the sample (Table 1); a total of 159 of these (71.6%) were previously described as being present in the somatic or sperm nuclei of other species (Uniprot, 2016), 41 (18.5 %) had no previously described nuclear presence and 22 (9.9%) were uncharacterized.

Proteins families

The most abundant family of proteins corresponded to ribosomal (13.1%, 29 of 222), followed by cytoskeleton (12.2%, 27 of 222), uncharacterized (9.9%, 22 of 222), histones (5.4%, 12 of 222), proteasome subunits (3.6%, 8 of 222 and heat shock (1.8%, 4 of 222). The other proteins clustered in other families accounted for 54% of the total proteins (120 of 222).

Table 1. Proteins identified in isolated swine sperm nuclear chromatin containing the molecular mass (MM), number of peptides (Nº Peps), associated gene name (Gene name), associated family (Family- OF: other families; R: ribosomal; UC: uncharacterized; C: cytoskeleton; H: histone; PS: proteasome subunits and HS: heat-shock), description and possible nuclear presence (Nuclear presence).

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
132204	218	LOC100626209	UC	Uncharacterized	NO
62621	140	FAM71B	OF	Protein FAM71B	YES
11807	105	PRM2	OF	Protamine-2	YES
24236	82	RAB2B	OF	Ras-related protein Rab-2B	YES
87427	75	ODF2	C	Outer dense fiber protein 2	NO
29526	67	ODF1	C	Outer dense fiber protein 1	NO
66818	66	CCIN	C	Calicin	YES

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
73173	64	CYLC1	C	Cilicin 1	YES
16837	61	LOC100522926	UC	Uncharacterized	NO
1934	57	GPX4	OF	Phospholipid hydroperoxide Glutathione peroxidase	YES
35571	54	FNDC8	OF	Fibronectin Type-3 domain-containing protein 8	NO
27420	52	GSTM3	OF	Glutathione S-transferase Mu 3	YES
30628	50	Capzb	C	F-actin binding protein, beta subunit	YES
49202	48	ACTL7A	C	Actin-like protein 7A	NO
36938	48	C1orf56	OF	Cromossome 1 open reading frame 56	YES
41728	43	ACTRT2	C	Actin related protein T2	NO
28350	39	SSC.25138	OF	Uncharacterized	NO
63925	36	FAM71A	OF	Protein FAM71A	YES
89288	35	VCP	OF	Transition endoplasmatic reticulum ATPase	YES
26688	35	RPS3	R	Protein-S3 40S ribosomal	YES
89681	35	DPY19L2	OF	C-mannosiltransferase	YES
89289	34	VCP	OF	Transition endoplasmatic reticulum ATPase	YES
50141	32	EEF1A	OF	Elongation factor 1-alpha	YES
23312	31	C7orf61	OF	Cromossome 7 open reading frame 61	YES
149547	30	SPATA31D1	OF	Spermatogenesis-associated protein 31D1	NO
26668	30	TPI1	OF	Triosephosphate isomerase	YES
39359	29	GAPDHS	OF	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	YES
40509	28	CYLC2	C	Cilicin 2	YES
29801	27	LUZP2	OF	Leucine zipper protein 2	NO
44260	26	HOXB3	OF	Homeobox protein Hox-B3	YES
198824	26	C2orf16	OF	Cromossome 2 open reading frame 16	YES
56664	26	LOC100620428	UC	Uncharacterized	NO
13988	25	HIST2H2AC	H	Histone H2A	YES
34878	24	CAPZA3	C	F-actin binding protein, alpha-3 subunit	YES
49960	24	LOC100510930	UC	Uncharacterized	NO
14135	22	HIST1H2AE	H	Histone H2A	YES
30596	21	LOC100511361	UC	Uncharacterized	NO
45511	20	ACTL7B	C	Actin-like protein 7B	NO
41107	20	ACTRT3	C	Actin related protein T3	YES
35515	20	PPP1CB	OF	Serine/threonine-protein phosphatase - beta catalytic subunit	YES
17462	19	TMEM89	OF	Protein TMEM89	YES
28726	19	TFAM	OF	Transcription factor A, mitochondrial	YES
70344	18	HSPA1L	HS	Heat shock 70 kDa type-1	YES
49831	18	TUBB4B	C	Tubulin 4B	YES
31998	17	FHL1C	OF	Four and a half LIM domains 1 protein, isoform C	YES
32404	15	ASRGL1	OF	Isoaspartyl peptidase/L-asparaginase	NO
10960	15	HSPE1	HS	Heat Shock protein 10 KDa	YES
11829	15	TXN	OF	Thioredoxin	NO
16987	14	HSPB9	HS	Heat shock protein beta-9	YES
67705	14	FUBP1	OF	Far upstream element-binding protein 1	YES
42644	14	PRR30	OF	Proline-rich protein 30	NO
10350	14	DYNLL2	OF	Dynein light chain 2, cytoplasmic	YES
11367	13	HIST1H4A	H	Histone H4	YES
53000	13	VRK3	OF	Inactive serine/threonine-protein kinase	YES
45693	13	ACTL9	C	ACTL9	YES
20169	13	RANGRF	OF	Ran guanine nucleotide release factor	YES
22187	13	GPX4	OF	Phospholipid hydroperoxide Glutathione peroxidase	YES

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
44835	12	SERBP1	OF	Plasminogen activator inhibitor 1 RNA-binding protein	YES
14865	11	RPL23	R	60S ribosomal protein L23	YES
15882	11	EEF1D	OF	Elongation factor 1-delta	YES
22157	11	LOC595122	H	Histone H1.3-like protein	YES
44895	11	PGK2	OF	2-phosphoglycerate kinase	YES
70252	11	HSPA5	OF	78 kDa glucose-regulated protein	YES
30604	11	VDAC1P5	UC	Uncharacterized	NO
70670	11	PABPC1	OF	Polyadenylate-binding protein 1	YES
20252	11	RPL11	R	60S ribosomal protein L11	YES
42030	10	GLUL	OF	Glutamine synthetase	YES
19441	10	C17H20orf106	OF	Orthologue of H. sapiens chromosome 20 open reading frame 106	YES
10045	10	LOC100522848	UC	Uncharacterized	NO
45773	10	FAM71D	OF	Protein FAM71D	YES
11693	10	RPLP2	R	60S ribosomal protein 2	YES
6825	9	LOC100522509	UC	Uncharacterized	NO
28421	9	RPS6	R	60s ribosomal protein 6	YES
21388	9	RPL17	R	60S ribosomal protein L17	YES
15129	9	H2AFX	H	Histone H2A	YES
68379	9	NT5C1B	OF	Cytosolic 5'-nucleotidase 1B	YES
37502	9	PCBP1	OF	Poly(rC)-binding protein 1	YES
37512	9	LOC733611	OF	Serine/threonine-protein phosphatase	NO
43833	8	SPAG4L	OF	Sperm associated antigen 4-like protein	NO
27205	8	LOC100738983	UC	Uncharacterized	NO
14122	8	LOC100519930	UC	Uncharacterized	NO
32694	8	YBX3	OF	Y-box-binding protein 3	YES
12784	8	RPL30	R	60s ribosomal protein 30	YES
33692	8	HNRNPC	R	Heterogeneous nuclear ribonucleoprotein C	YES
24306	8	RPL13	R	60s ribosomal protein 13	YES
13890	8	HIST1H2BN	H	Histone H2B	YES
15816	7	RPS23	R	40S ribosomal protein S23	YES
17266	7	PRDX5	OF	Peroxideroxin 5	NO
17695	7	RPL23A	R	60s ribosomal protein 23A	YES
35460	7	SPEM1	OF	Spermatid maturation protein 1	YES
220718	7	ARHGEF17	OF	Protein Arhgef17	NO
40353	7	SUV39H1	H	Histone-lysine N-methyltransferase SUV39H1	YES
14728	6	UBA52	R	Ubiquitin-60s ribosomal protein L40	YES
16730	6	AWN	OF	Carbohydrate-binding protein AWN	NO
20688	6	PDAP1	OF	28 kDa heat- and acid-stable phosphoprotein	YES
132137	6	NRD1	OF	Protein NRD1	YES
34484	6	C9orf24	OF	Cromossome 9 open reading frame 24	YES
272275	6	FN1	OF	Protein FN1	NO
29945	6	RPS3A	R	40S ribosomal protein S3a	YES
16832	6	LOC100517970	UC	Uncharacterized	NO
53469	6	GC	OF	Vitamin D-binding protein	NO
69823	6	HSPA2	HS	Heat shock-related 70 kDa protein 2	YES
57993	6	KRT10	C	Keratin 10	YES
15860	6	RPL32	R	60S ribosomal protein L32	YES
40941	6	TSSK2	OF	Testis-specific serine/threonine-protein kinase 2	YES
87445	6	PHTF1	OF	Putative homeodomain transcription factor 1	YES
269468	6	TLN1	OF	Talin-1	NO
11528	6	LOC100523874	UC	Uncharacterized	NO

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
20853	6	TMPO	C	Lamina-associated polypeptide 2, isoform alpha	YES
14123	6	LOC100152135	H	Histone H2B	YES
36832	6	PPP1CC	OF	Serine/threonine-protein phosphatase - gamma catalytic subunit	YES
7379	6	ATOX1	OF	Copper transport protein ATOX1	NO
39623	5	ZPBP	OF	Zona pellucida-binding protein	YES
73702	5	PLCZ	OF	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase zeta	YES
39591	5	CXorf66	OF	Cromossome X open reading frame 66	YES
61477	5	BAG3	OF	BAG family molecular chaperone regulator 3	YES
28658	5	LOC100155139	PS	Proteasome subunit beta type OS	YES
77524	5	KHSRP	OF	Far upstream element-binding protein 2	YES
32388	5	WBP2NL	OF	Protein WBP2NL	NO
28433	5	LOC100154408	PS	Proteasome subunit alpha type	YES
199915	5	MAP2	C	Microtubule-associated protein	YES
54427	5	KRT8	C	Keratin 8	YES
19256	5	RPL18	R	60S ribosomal protein L18	YES
37231	5	PCBP2	OF	Poly(rC)-binding protein 2	YES
20810	5	RPL12	R	60S ribosomal protein L12	YES
15300	4	TNP2	OF	Nuclear transition protein 2	YES
80207	4	ACSL6	OF	Long-chain-fatty-acid--CoA ligase 6	YES
35692	4	NUDT18	OF	8-oxo-dGDP phosphatase NUDT18	YES
62282	4	COIL	OF	Coilin	YES
41868	4	ACTRT1	C	Actin-related protein T1	NO
28268	4	H1FNT	H	Histone H1 testis-specific	YES
36023	4	YBX1	OF	Protein Y-BOX 1	YES
345426	4	DCHS1	OF	Protocadherin-16	NO
23384	4	NACA	OF	Nascent polypeptide-associated complex subunit alpha	YES
30101	4	RPL7A	C	60S ribosomal protein L7a	YES
35453	4	ANTXRL	OF	Anthrax toxin receptor-like	NO
18922	4	CUTA	OF	Protein Cuta	NO
4312	4	LOC100525679	UC	Uncharacterized	NO
32330	4	TMEM38B	OF	Trimeric intracellular cation channel type B	NO
16495	4	UBE2V1	OF	Ubiquitin-conjugating enzyme E2 variant 1	YES
7660	4	WFDC6	OF	WAP four-disulfide core domain protein 6	YES
16377	4	CCDC58	OF	Coiled-coil domain-containing protein 58	YES
29329	4	ERP29	OF	Endoplasmic reticulum resident protein 29	NO
76725	4	POLH	OF	DNA polymerase eta	YES
17791	4	LOC100738931	UC	Uncharacterized	NO
22127	3	SPATA3	OF	Spermatogenesis-associated protein 31D1	NO
16588	3	LOC100517228	UC	Uncharacterized	NO
14227	3	PDCD5	OF	Programmed cell death protein 5	YES
13001	3	LOC100736633	H	Histone H2A	YES
30211	3	TSSK6	OF	Testis-specific serine kinase 6	YES
14719	3	LOC100514544	UC	Uncharacterized	NO
25594	3	AK3	OF	GTP:AMP phosphotransferase	YES
23726	3	SARNP	R	SAP domain-containing ribonucleoprotein	YES
27086	3	C12orf60	OF	Cromossome 12 open reading frame 60	YES
26411	3	PSMA5	PS	Proteasome subunit alpha type-5	YES

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
27399	3	PSMA6	PS	Proteasome subunit alpha type-6	YES
176173	3	EIF4G1	OF	Eukaryotic translation initiation factor 4 gamma 1	NO
13332	3	RPS20	R	40s ribosomal protein 20	YES
49404	3	KRT18	C	Keratin 18	YES
14463	3	RPL31	R	60S ribosomal protein L31	YES
51677	3	MECP2	OF	Methyl-CpG-binding protein 2	YES
81805	3	FTNB	OF	Fertilin beta	NO
7841	3	RPS28	R	40S ribosomal protein S28	YES
14550	3	RPL35	R	60s ribosomal protein 35	YES
27390	3	FTMT	OF	Ferritin	NO
29820	3	LOC100523540	UC	Uncharacterized	NO
15327	3	H3F3A	H	Histone H3.3	YES
58722	3	PKM	OF	Pyruvate kinase	YES
29483	3	PSMA4	PS	Proteasome subunit alpha 4	YES
13739	3	RPS25	R	40s ribosomal protein 25	YES
38915	3	CSNK1A1	C	Casein kinase I isoform alpha	YES
19358	3	LOC100737887	OF	Peptidyl-prolyl cis-trans isomerase	NO
35836	3	GAPDH	OF	Glyceraldehyde-3-phosphate dehydrogenase	YES
27928	3	PSMA8	PS	Proteasome subunit alpha 8	YES
24156	3	PSMA7	PS	Proteasome subunit alpha 7	YES
16744	3	PKD2L2	OF	Polycystic kidney disease 2-like 2 protein	NO
21378	3	IQCF5	OF	IQ domain-containing protein F5	NO
37638	3	GDE1	OF	Glycerophosphodiester phosphodiesterase gde1	NO
16697	3	PFDN2	OF	Prefoldin subunit 2	YES
10231	3	BANF2	OF	Barrier-to-autointegration factor-like protein	YES
87428	3	ODF2	C	Outer dense fiber 2	NO
49257	2	HNRNPK	R	Heterogeneous nuclear ribonucleoprotein K	YES
42009	2	ACTA2	C	Actin alpha 2	NO
14759	2	RPL22	R	60S ribosomal protein 22	YES
151110	2	ACIN1	OF	Acinus	YES
32190	2	RPL6	R	60S ribosomal protein 6	YES
48197	2	LOC100739434	UC	Uncharacterized	NO
135902	2	AHNAK	OF	Neuroblast differentiation-associated protein	NO
17232	2	CENPV	OF	Centromeric protein V	YES
12441	2	RPL36A	R	60S ribosomal protein L36A	YES
26047	2	RPL18A	R	60S ribosomal protein L18A	YES
89391	2	PPP1R9B	OF	Neurabin 2	YES
35431	2	DECR1	OF	2,4-dienoyl-CoA reductase	YES
15097	2	RPS19	R	40S ribosomal protein S19	YES
12423	2	LOC100519900	UC	Uncharacterized	NO
32953	2	PSMG1	PS	Proteasome subunit G1	YES
57160	2	C17orf74	OF	Cromossome 17 open reading frame 74	YES
5970	2	MT2A	OF	Metallothionein-2A	YES
19660	2	LOC100522904	UC	Uncharacterized	NO
31591	2	LDHC	OF	L-lactate dehydrogenase C	YES
24423	2	RAN	OF	GTP-binding nuclear protein	YES
33347	2	RBMX	OF	RNA-binding motif protein, X chromosome	YES
84967	2	SF3A1	OF	Splicing factor 3A subunit 1	YES
20972	2	PEBP1	OF	Phosphatidylethanolamine-binding protein 1	YES
50020	1	KRT28	C	Keratin 28	YES

MM (Da)	Nº Peps	Gene name	Family	Description	Nuclear presence
49417	1	KRT15	C	Keratin 15	YES
65109	1	KRT75	C	Keratin 75	YES
57113	1	KRT4	C	Keratin 4	YES
40202	1	PGK1	OF	Phosphoglycerate kinase 1	YES
13377	1	H2AFV	H	Histone fragment H2A	YES
48977	1	DLST	OF	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	YES
8550	1	CHTOP	OF	Chromatin target of PRMT1 protein	YES
94733	1	MATR3	OF	Matrin 3	YES
36135	1	ELAVL1	OF	ELAV type 1	YES
27206	1	LOC100519489	UC	Uncharacterized	NO
81382	1	CC2D1B	OF	Coiled-coil and C2 domain-containing protein 1B	YES
24416	1	LOC100525255	UC	Uncharacterized	NO
153433	1	BRD4	OF	Protein BRD4	NO
22022	1	LOC100152612	OF	Peptidyl-proline cis-trans isomerase	YES
29598	1	RPS4	R	40S ribosomal protein 4	YES
48074	1	pdi-p5	OF	Protein disulfide isomerase P5	NO
28677	1	PGAM2	OF	Phosphoglycerate mutase 2	YES

Discussion

The results of toluidine blue analyses suggest that the animals in question presented normal sperm with respect to chromatin compaction (Beletti *et al*., 2005).

According to DeMateo *et al*. (2011) the proteomic analyses of human sperm nuclei revealed that the most abundant proteins were the histone family (9.7%), followed by cytoskeletal proteins (cytokeratins, tubulin and tektinas, 8.6%), ribosomal proteins (6.7%), proteasome subunits (6.2%), uncharacterized proteins (6.2%), spanx proteins (1.7%), and heat shock proteins (1.2%). Notably DeMateo *et al*. (2011) studied the proteins of human sperm nuclei, whereas this study focused on the proteins of the swine sperm nuclear matrix. Similarities between the identified families can be observed, although there is little similarity between the described proportions.

Notably LOC100626209 protein had the highest number of peptides (218) in this study, followed by FAM71B (140 peptides). The third most detected (105 peptides) was protamine 2. The LOC100626209 protein was classified in this study as an uncharacterized protein and had not previously been described in any species, and thus we cannot speculate as to its possible actions. Lemos (2013) described uncharacterized proteins with high molecular weights in the bovine nuclear annulus, which may indicate that the LOC100626209 protein is present in this structure. The second protein FAM71B was identified in the human sperm nucleus and might be involved in RNA biogenesis (Van Koningsbruggen *et al*., 2007). This protein was identified for the first time in swine sperm nuclear matrix.

Interestingly, the third protein in terms of the number of peptides found was protamine 2. Protamine 2 was described as being absent in the swine sperm nucleus, whereas protamine 1 was identified in various mammalian species, including swine (Pirhonen *et al*., 1994; Andrabi, 2007). Although some studies did not identify protamine 2, this protein was reported to be transcribed and translated in swine at low levels. However, an 8 amino acid deletion occurred at the amino acid terminus of the molecule and probably had functional relevance (Maier *et al*., 1990).

The absence of protamine 1 in the sample of swine sperm chromatin also occurred in the evaluation of the human sperm nucleus (DeMateo *et al*., 2011). The same authors reported that protamine 2 was present among the basic nuclear proteins, as was demonstrated in this study. These authors also suggested a possible explanation for the lack of protamine 1 based on its particular amino acid composition. Protamine is a basic protein that is highly enriched in lysine and arginine (more than 50% in the mature form). Trypsin cleaves peptide chains primarily on the carboxyl side of lysine or arginine, which results in very small peptide fragments that cannot be detected under the conditions used in mass spectrometry. This limitation is particularly important for protamine 1 because it is more enriched in arginine than protamine 2.

Genes for two protamines (PRM1 and PRM2) and two transition proteins (TNP1 and TNP2) have been characterized in several mammalian species (Engel *et al*., 1992). According to the same authors, the human, swine and bull genes for PRM1, PRM2 and TNP2 are closely connected along a specific stretch of DNA, whereas the gene for TNP1 in all of the species studied is located on another chromosome. In this study we detected just genes for PRM2 and TNP2, but we also detected a peptide of the CHTOP protein, which is responsible for the chromatin marking of protamine 1. This protein is also involved in transcriptional regulation. It contains a region rich in glycine and arginine that interacts with the RNA or DNA directly or

in combination with other nucleotide binding proteins (Takai *et al*., 2014).

Proteins related to the ribonucleosome were identified in the sample. The 40S ribosomal proteins RPS25, RPS3A and RPS6 were also reported in the human sperm nuclear proteome by DeMateo *et al*. (2011). Some 60S ribosomal proteins were also found in the human sperm nucleus, as in this study. The presence of RPL9 protein in human sperm nuclei was confirmed using an immunofluorescence technique (DeMateo *et al*., 2011).

The proteomic analysis of isolated sperm nuclei indicated the presence of only the cytoplasmic ribosomal proteins (80S (60S + 40S)) and not the mitochondrial ribosomal proteins (55S). Thus, the detection of cytoplasmic ribosomal proteins in this study was consistent with the cytoplasmic translation proposed by some authors (Lambard *et al*., 2004; Galeraud-Denis *et al*., 2007; DeMateo *et al*., 2011The ribosomal proteins are recognized as cytoplasmic proteins. Thus, their detection in swine sperm nuclei is an important finding and may clarify possible paternal epigenetic actions in swine.

The family of cytoskeletal proteins (i.e., keratin, tubulin and actin) that was identified in this study has also been identified in the human sperm nucleus (DeMateo *et al*., 2011). The same authors reported that some cytoskeletal molecules were demonstrated to participate in the formation of the sperm tail and sperm nucleus format. Specifically, tubulin was also detected in the heads of sperm, suggesting a possible role related to the acrosome reaction. Cytokeratin and actin were also associated with the sperm nuclear matrix in guinea pigs (Ocampo *et al*., 2005), and now in swine sperm nuclear matrix.

This study identified several histone variants in mature and ejaculated sperm. In this respect, we demonstrated the transport and incorporation of the sperm histones to the zygote, indicating another potential effect of parental epigenetic reprogramming on the zygote after fertilization that is independent of the imprinting state (Miller *et al*., 2010). Histones were only a small part of the proteins identified in this sperm chromatin sample, indicating that many other proteins can transmit epigenetic information to the zygote.

Epigenetic control of gene expression exists for the activation of DNA methylation because methylation, acetylation and phosphorylation of histones and histone entry into the oocyte leaves a space for DNA and epigenetic signaling based on histones, which can be important for subsequent embryonic development (Miller *et al*., 2010). The same authors described results from analyses of the composition of soluble (linked to histones) and insoluble (linked to protamines) domains in human and murine sperm and indicated that chromatin was actually the most significant contributor to an epigenetic signal in these cells.

Concurrent with visible changes in the organization of the sperm chromatin, the histones would be removed from the DNA of initial spermatids and spermatocytes and replaced by transition proteins. Subsequently, the transition proteins are replaced by protamines that are responsible for the final condensation and stabilization of the sperm chromatin (D'Occhio *et al*., 2007). The same authors reported that histones could persist in ejaculated mature sperm from humans and other mammals, as demonstrated in this study for several types of histones.

The fact that sperm nucleus histones are not necessarily characterized as an error in chromatin compaction has led to a concept change. Beletti (2013) reported that there are regions interspersed between the toroidal structures of sperm chromatin that contain nucleosome sequences, which often contain hypomethylated DNA. These regions may be involved in important functions related to early embryonic development and paternal epigenetic inheritance.

A total of 25 different proteasome subunits were reported in the human sperm nucleus (DeMateo *et al*., 2011); PSMA4, 5, 6 and 8 were also identified in this study. These proteins are the protein machine for ubiquitin-mediated proteolytic degradation, which has been implicated in many cellular processes including cell cycle progression, transcriptional regulation, signal transduction, and determining the fate of the cell (Zhong and Belote, 2007).

Is interesting to note the identification of proteins involved in the regulation of gene expression, such as PHTF1 protein (Table 1). It was thought that the sperm is an inert transcriptionally cell, with the sole objective is to provide its DNA (packed by protamines) to the oocyte (DeMateo *et al*., 2011). Thus, the sperm nuclear proteins identified here may be only a remnant of the sperm cell differentiation process, or may be relevant to fertilization or success of embryo development.

For example, the PPP1CC protein, that according to MacLeod *et al*. (2014) is an essential phosphatase protein in spermatogenesis, was found in the sample, which concurs with a previous result (DeMateo *et al*., 2011). An important nuclear protein (WBP2NL) found in this study, was also identified in rats by Chen *et al*. (2014) and may play a role in meiotic resumption and male pronucleus formation, which are related to early embryonic development. The protein glyceraldehyde 3-phosphate dehydrogenase (GAPDH) found in this sample performed functions in glycolysis and participated in nuclear events, including transcription, RNA transport, DNA replication and apoptosis (Applequist *et al*., 1995).

The heterogeneous nuclear ribonucleoprotein (HNRNPs) is involved in many biological processes, such as cell signaling, DNA repair, and regulation of gene expression and protein (Almeida *et al*., 2014). The HNRNPC presented 8 peptides in the sample and was previously described in the human sperm nucleus (DeMateo *et al*., 2011). The heterogeneous nuclear ribonucleoprotein K (HNRNPK) was also found in this sample, and according to Almeida *et al*. (2014) this protein is predominantly located in the nucleus, where it is involved in multiple steps in gene expression such as transcription, RNA splicing and translation.

Proteins of the nuclear matrix may be involved in different genetic markers that can contribute after fertilization to establish the order of paternal gene

reactivation (Oliva, 2006) and may serve as an indication of protein epigenetic function. Therefore, based on the results obtained in this study and the functions of some described proteins, the nuclear matrix, which is composed in part of chromatin proteins, transmits essential mechanisms for the growth and differentiation of the oocyte. Thus, the proteins become not only structural components of the chromatin/matrix architecture but also essential components of successful reproduction.

In this survey, 9.9% of the proteins were classified as uncharacterized (i.e., there are no reports in the literature regarding the description, identification and isolation of these proteins for any species). Thus, future studies can be designed in an attempt to characterize these proteins and elucidate their functions to clarify possible epigenetic paternal inheritance that is beyond the current scientific knowledge.

The set of proteins present in swine sperm chromatin shows that the nuclear matrix plays important roles in the development and maturation of sperm cells. Therefore, it would be important to determine how these protein structures are linked to the establishment of epigenetic functions and how they can affect the embryo development.

In conclusion the protein isolation from the swine sperm nuclear matrix was satisfactory and demonstrated that the protocol was efficient. Some protein families were identified and described. However, it was not possible to identify some protein structures, such as protamine 1.

Therefore, this study contributes to a catalog of protein structures that may be useful in future proteomics studies. The comparison between fertile and sub-fertile animals can contribute to the search for proteomic variations. These studies can improve reproductive technologies and animal breeding of several species by focusing breeders' choices on epigenetic potential as well as high genetic potential.

Acknowledgments

The authors acknowledge CNPq and FAPEMIG for financial support. We also thank the Federal University of Uberlandia, particularly the post-graduate program in veterinary science.

References

Almeida LO, Garcia CB, Matos-Silva FA, Curti C, Leopoldino AM. 2014. Accumulated SET protein up-regulates and interacts with hnRNPK, increasing its binding to nucleic acids, the Bcl-xS repression, and cellular proliferation. *Biochem Biophys Res Commun*, 445:196-202.

Andrabi SMH. 2007. Mammalian sperm chromatin structure and assessment of DNA fragmentation. *J Assist Reprod Genet*, 24:561-569.

Applequist SE, Keyna U, Calvin MR, Beck-Engeser GB, Raman C, Jäck HM. 1995. Sequence of the rabbit glyceraldehyde-3-phosphate dehydrogenaseencoding Cdna. *Gene*, 163:325-326.

Beletti ME, Costa LF, Guardieiro MM. 2005. Morphometric features and chromatin condensation abnormalities evaluated by toluidine blue staining in bull spermatozoa. *Braz J Morphol Sci*, 22:85-90.

Beletti ME. 2013. Cromatina espermática: quebrando paradigmas. *Rev Bras Reprod Anim*, 37:92-96.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analyt Biochem*, 72:248-254.

Chen M, Wang H, Li X, Li N, Xu G, Meng Q. 2014. PLIN1 deficiency affects testicular gene expression at the meiotic stage in the first wave of spermatogenesis. *Gene*, 543:212-219.

Codrington AM, Hales BF, Robaire B. 2007. Exposure of male rats to cyclophosphamide alters the chromatin structure and basic proteome in spermatozoa. *Hum Reprod,* 22:1431-1442.

DeMateo S, Castillo J, Estanyol JM, Ballescá JL, Oliva R. 2011. Proteomic characterization of the human sperm nucleus. *Proteomics*, 11:2714-2726.

D'Occhio MJ, Hengstberger KJ, Johnston SD. 2007. Biology of sperm chromatin structure and relationship to male fertility and embryonic survival. *Anim Reprod Sci*, 101:1-17.

Engel W, Keime S, Kremling H, Hameister H, Schluter G. 1992. The genes for protamine 1 and 2 (PRM1 and PRM2) and transition protein 2 (TNP2) are closely linked in the mammalian genome. *Cytogenet Cell Genet*, 61:158-169.

Galeraud-Denis I, Lambard S, Carreau S. Relationship between chromatin organization, mRNAs profile and human male gamete quality. *Asian J Androl*, 9:587-592.

Lambard S, Galeraud-Denis I, Martin G, Levy R. 2004. Analysis and significance of mRNA in human ejaculated sperm from normozoospermic donors: relationship to sperm motility and capacitation. *Mol Hum Reprod*, 10:535-541.

Lemos MS. 2013. Caracterização morfológica e bioquímica do ânulo nuclear de espermatozoide bovino. Uberlandia, Brazil: Federal University of Uberlândia. Thesis.

MacLeod G, Taylor P, Mastropaolo L, Varmuza S. 2014. Comparative phosphoproteomic analysis of the mouse testis reveals changes in phosphopeptide abundance in response to Ppp1cc deletion. *EuPA Open Proteomics*, 2:1-16.

Maier WM, Nussbaum G, Domenjoud L, Klemm U, Engel W. 1990. The lack of protamine 2 (P2) in boar and bull spermatozoa is due to mutations within the P2 gene. *Nucleic Acids Res*, 18:1249-1254.

Miller D, Brinkworth M, Iles D. 2010. Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. *Reproduction,* 139:287-301.

Morandi-Filho R. 2013. Análise da estrutura e identificação de proteínas da cromatina nuclear espermática de bovinos. Uberlandia, Brazil: Federal University of Uberlandia. Thesis.

Ocampo J, Mondragon R, Roa-Espitia AL, Chiquete-Felix N, Salgado ZO, Mújica A. 2005.

Actin, myosin, cytokeratins and spectrin are components of the guinea pig sperm nuclear matrix. *Tissue Cell,* 37:293-308.

Oliva R. 2006. Protamines and male infertility. *Hum Reprod Update*, 12:417-435.

Oliva R, Ballescá JL. 2012. Proteomics of the spermatozoon. *Balcan J Med Genet*, 15:27-30.

Pirhonen A, Linnala-Kankkunen A, Maenpaa PH. 1994. Identification of phosphoseryl residues in protamines from mature mammalian spermatozoa. *Biol Reprod*, 50:981-986.

Rauch A. 2006. Computational Proteomics Analysis System (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. *J Proteome Res*, 5:112-121.

Takai H, Masuda K, Shirahige K, Akiyama T. 2014. 5-Hydroxymethylcytosine plays a critical role in glioblastomagenesis by recruiting the CHTOPMethylosome complex. *Cell Rep*, 9:48-60.

Uniprot.org [homepage on the internet]. 2015. Proteins data bank, Inc. Avaiable on: http://www.uniprot.org/. Accessed on: Jun 1st 2015.

Van Koningsbruggen S, Straasheijm KR, **Sterrenburg E**, **Graaf N**, **Dauwerse HG**, **Frants RR**, **Van der Maarel SM**. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma*, 116:53-64.

Wright PC, Noirel J, Ow SY, Fazeli A. 2012. A review of current proteomics technologies with a survey on their widespread use in reproductive biology investigations. *Theriogenology,* 77:738-765.

Yamauchi Y, Shaman JA, Ward WS. 2011. Non-genetic contributions of the sperm nucleus to embryonic development. *Asian J Androl*, 13:31-35.

Zhong L, Belote JM. 2007. The testis-specific proteasome subunit Pros_6T of *D. melanogaster* is required for individualization and nuclearmaturation during spermatogenesis. *Development*, 134:3517-3525.