

## Supplementary information

### Occurrence and functional significance of the transcriptome in bovine (*Bos taurus*) spermatozoa

Sellappan Selvaraju<sup>1,\*</sup>, Sivashanmugam Parthipan<sup>1,#</sup>, Lakshminarayana Somashekar<sup>1,#</sup>, Atul P Kolte<sup>2</sup>, Binsila B Krishnan<sup>1</sup>, Arunachalam Arangasamy<sup>1</sup> and Janivara Parameshwaraiah Ravindra<sup>1</sup>

<sup>1</sup>Reproductive Physiology laboratory, Animal Physiology Division, <sup>2</sup>Omics laboratory, Animal Nutrition Division, ICAR-National Institute of Animal Nutrition and Physiology, Adugodi, Bengaluru-560030

<sup>#</sup>Department of Biochemistry, Jain University, Bengaluru-560001

**Supplemental Table S1: The RNA integrity of top 100 abundant spermatozoal transcripts. The transcripts integrity was analyzed manually by visualizing the read coverage.**

Gene ID	TPM	UGR	Intact*	5' coverage	3' coverage
<i>PRM1</i>	8659	120	Yes	Yes	Yes
<i>YWHAZ</i>	3050	84	Yes	Yes	Yes
<i>SCP2D1</i>	2726	182	Yes	Yes	Yes
<i>CHMP5</i>	1693	260	Yes	No	Yes
<i>MGC137055</i>	1434	74	Yes	No	Yes
<i>GTSF1L</i>	1416	155	Yes	No	Yes
<i>PAG5</i>	971	962	Yes	No	Yes
<i>CCDC181</i>	919	144	Yes	Yes	Yes
<i>TNP1</i>	875	308	Yes	Yes	Yes
<i>CCDC85A</i>	829	3481	Yes	No	Yes
<i>C23H6orf62</i>	788	207	Yes	Yes	Yes
<i>BCL2L11</i>	783	74	Yes	Yes	Yes
<i>CCDC168</i>	635	179	Yes	Yes	Yes
<i>SSRP1</i>	615	219	Yes	No	Yes
<i>FAM209B</i>	612	77	Yes	Yes	Yes
<i>CAPZA3</i>	558	183	Yes	Yes	Yes
<i>NEK3</i>	493	636	Yes	Yes	No
<i>RAD21</i>	474	242	Yes	Yes	Yes
<i>SP2</i>	461	390	Yes	No	No
<i>SF3B1</i>	456	361	Yes	Yes	Yes
<i>TMED6</i>	455	10	Yes	Yes	Yes
<i>PPWD1</i>	445	37	Yes	Yes	Yes
<i>PAG10</i>	440	416	Yes	No	No
<i>INVS</i>	439	2384	Yes	No	No
<i>RPL23</i>	424	22	Yes	No	Yes
<i>KIAA1731</i>	422	126	Yes	Yes	Yes
<i>ODZ3</i>	419	5074	Yes	No	No
<i>RPL27A</i>	407	536	Yes	No	Yes
<i>LRRD1</i>	407	27	Yes	Yes	Yes
<i>RPS18</i>	403	60	Yes	Yes	Yes
<i>SFXN3</i>	403	557	Yes	No	No
<i>IFN-TAU</i>	389	118	Yes	Yes	Yes

---

<i>IGF1R</i>	384	5979	Yes	No	Yes
<i>SH3GL2</i>	375	2332	Yes	No	No
<i>PAG7</i>	375	119	Yes	Yes	Yes
<i>VPS4A</i>	368	472	Yes	No	Yes
<i>FABP1</i>	2923	1074	No	Yes	No
<i>THSD4</i>	1961	2506	No	No	No
<i>NR2E3</i>	1610	1241	No	No	No
<i>SV2C</i>	1518	2592	No	No	No
<i>TOE1</i>	1359	1743	No	No	Yes
<i>SLC16A7</i>	1284	2831	No	No	No
<i>MCOLN2</i>	1231	1756	No	No	No
<i>UNC119</i>	1136	790	No	Yes	No
<i>CXCR4</i>	1095	975	No	No	Yes
<i>MMP2</i>	933	1417	No	No	Yes
<i>ITPA</i>	919	458	No	Yes	No
<i>DNAJB12</i>	914	2315	No	No	No
<i>GPS2</i>	907	229	No	No	No
<i>BPIFA2A</i>	900	833	No	No	Yes
<i>KCNU1</i>	896	2544	No	No	Yes
<i>TRPC3</i>	892	1303	No	Yes	No
<i>ANKRD34C</i>	867	3053	No	Yes	Yes
<i>ILF2</i>	866	126	No	No	Yes
<i>CYP8B1</i>	804	529	No	Yes	No
<i>ASTE1</i>	785	1292	No	Yes	No
<i>YSK4</i>	756	1742	No	No	No
<i>PRSS21</i>	719	37	No	Yes	No
<i>RDH5</i>	707	496	No	No	Yes
<i>LSM10</i>	649	349	No	No	Yes
<i>DBC1</i>	646	3393	No	No	No
<i>SFXN1</i>	643	1721	No	No	No
<i>PTGER2</i>	620	860	No	No	Yes
<i>GPR89</i>	615	1048	No	Yes	No
<i>VIT</i>	606	2722	No	No	No
<i>FBXO6</i>	586	499	No	No	No
<i>B7H6</i>	550	586	No	No	No

---

<i>ADCY8</i>	548	3841	No	No	No
<i>PTPLAD2</i>	535	885	No	No	Yes
<i>BRPF3</i>	517	1527	No	Yes	No
<i>DHRS7C</i>	515	406	No	No	Yes
<i>RPL37</i>	493	27	No	No	No
<i>HTR1A</i>	478	394	No	No	No
<i>MLH1</i>	472	1919	No	Yes	No
<i>TMEM125</i>	458	359	No	Yes	No
<i>GPR110</i>	446	1048	No	No	No
<i>DHFR</i>	438	490	No	Yes	No
<i>RPL36AL</i>	434	105	No	No	Yes
<i>TLR5</i>	423	513	No	Yes	No
<i>BSPRY</i>	421	550	No	No	Yes
<i>BRB</i>	418	206	No	No	Yes
<i>EVI2B</i>	417	440	No	No	Yes
<i>ZNF648</i>	416	361	No	No	No
<i>NHLRC3</i>	413	580	No	No	Yes
<i>APPL2</i>	411	812	No	Yes	No
<i>COMMD1</i>	406	1499	No	Yes	No
<i>PRR5L</i>	401	1260	No	No	Yes
<i>KRT75</i>	398	714	No	No	No
<i>GPA33</i>	392	960	No	No	No
<i>CIH21orf62</i>	390	515	No	No	Yes
<i>PCDHB11</i>	389	581	No	No	Yes
<i>RASL10B</i>	386	209	No	No	Yes
<i>CHTF8</i>	385	677	No	Yes	No
<i>ATPBD4</i>	382	950	No	Yes	No
<i>CABS1</i>	378	48	No	No	No
<i>MAPK1</i>	378	193	No	No	No
<i>VGLL3</i>	370	664	No	No	No
<i>TOMM40L</i>	368	626	No	No	No
<i>CATSPER3</i>	367	427	No	No	No
<i>RPL6</i>	363	6	No	No	No

\*The intactness was considered if the reads were spread covering most of the exons

**Supplemental Table S2: The tissue level expression (RNAs and proteins) and the reproductive functions of top 20 abundant transcripts in bovine spermatozoa.**

Gene ID	Gene Name	Tissue Expression (RNA level)	Protein Localization	Reproductive Functions	References
<i>PRMI</i>	Protamine 1	Testis (enriched) and prostate	Nucleus and chromosome	Histone replacement and chromatin remodeling	1
<i>CHMP5</i>	Charged Multivesicular Body Protein 5	Testis and prostate	Cytoplasm, Cytoskeleton-Microtubules and Cytoskeleton-Cytokinetic bridge (Seminiferous duct)	Regulation of late endosome function during embryogenesis	2
<i>YWHAZ</i>	Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein, Zeta	Testis and Prostate	Intracellular (Cells in Seminiferous duct, Leydig cells and Glandular cells)	Spermatozoa protein phosphorylation	3
<i>GTSF1L</i>	Gametocyte Specific Factor 1-Like	Testis (enriched)	Intracellular (Cells in Seminiferous duct and Leydig cells)	Retrotransposon suppression of male germ cells during spermatogenesis	4
<i>SCP2D1</i>	SCP2 Sterol-Binding Domain Containing 1	Testis (enriched)	Intra cellular (Cells in Seminiferous duct and Leydig cells)	Positive association with semen quality	5
<i>MGC137055</i>	DET1 And DDB1 Associated 1 (uncharacterized protein)	Testis	Not available	Reported as full length transcript in bovine spermatozoa	6
<i>CCDC181</i>	Coiled-Coil Domain Containing 181	Testis (enriched)	Nucleoplasm (Seminiferous duct)	Involved in the microtubular manchette development of spermatozoa.	7
<i>FABP1</i>	Fatty Acid Binding Protein1	Not available	Not available	Spermatogenesis	8

<i>THSD4</i>	Thrombospondin Type 1 Domain Containing 4	Testis and Prostate (enriched)	Cells in Seminiferous duct, Leydig cells and Glandular cells	Not available	Not available
<i>NR2E3</i>	Nuclear Receptor Subfamily 2 Group E Member 3	Testis and Prostate	Nuclear receptor	mRNA expressed in testis	9
<i>SV2C</i>	Synaptic Vesicle Glycoprotein 2C	Testis and Prostate	Cells in Seminiferous duct, Leydig cells and Glandular cells	Not available	Not available
<i>TOE1</i>	Target of EGR1, Membrane 1 (Nuclear)	Testis and prostate	Cells in Seminiferous duct, Leydig cells and Glandular cells	Not available	Not available
<i>SLC16A7</i>	Solute Carrier Family 16 Member 7	Testis and prostate	Cells in Seminiferous ducts	Involved in sperm metabolism	10
<i>MCOLN2</i>	Mucolipin 2	Testis and prostate	Cells in Seminiferous duct, Leydig cells and Glandular cells	Expressed in epididymis	11
<i>UNC119</i>	Unc-119 Lipid Binding Chaperone	Testis and prostate	Glandular cells	Not available	Not available
<i>CXCR4</i>	C-X-C Motif Chemokine Receptor 4	Testis and prostate	Cells in seminiferous ducts and Glandular cells	Chemokine signaling molecule in spermatozoa	12
<i>PAG5</i>	Pregnancy associated glycoprotein 5	Testis and prostate	Cells in seminiferous ducts, Leydig cells and Glandular cells	Reported in bovine placenta	13
<i>MMP2</i>	Matrix Metalloproteinase 2 (Gelatinase A, 72Kda Gelatinase, 72kDa type IV Collagenase)	Testis and prostate	Leydig cells and Glandular cells	Penetration of zona pellucida	14
<i>ITPA</i>	Inosine	Testis and	Cells in	mRNA	15

	Triphosphatase (Nucleoside Triphosphate Pyrophosphatase)	prostate	seminiferous ducts, Leydig cells and Glandular cells	expression was high in testis.	
<i>DNAJB12</i>	DnaJ Heat Shock Protein Family (Hsp40) Member B12	Testis and prostate	Cells in seminiferous ducts and Glandular cells	Chaperone (Similar to HSP40)	16

Note: The above information's are obtained from [www.ncbi.nlm.nih.gov/pubmed](http://www.ncbi.nlm.nih.gov/pubmed), [www.gpubmed.org](http://www.gpubmed.org); [www.ebi.ac.uk](http://www.ebi.ac.uk), [www.proteinatlas.org](http://www.proteinatlas.org), and [www.uniprot.org](http://www.uniprot.org)

**Supplemental Table S3: The abundantly expressed (Top 10) LOCs in the bovine spermatozoa.**

Gene ID	Gene Name/Description	TPM	UGR
<i>LOC100335847</i>	Protamine 1-like	7817	231
<i>LOC786921</i>	Deubiquitinating enzyme 3-like	6418	1610
<i>LOC101906132</i>	--	4391	711
<i>LOC404073</i>	Histone H2B variant PT15	4174	175
<i>LOC786597</i>	Chromosome 20 open reading frame 106 homolog	3887	244
<i>LOC100335236</i>	PREDICTED: Bos taurus hypothetical protein LOC100335236 transcript	3742	2374
<i>LOC786327</i>	T-cell activation Rho GTPase-activating protein-like	3571	207
<i>LOC786899</i>	T-cell activation Rho GTPase-activating protein-like	3379	1253
<i>LOC781191</i>	Uncharacterized protein, found in RNA derived from captured sperm of mixed chromosome content. The gene mapping by the National Center for Biotechnology Information (NCBI) located DNA for GX1 to the bovine X chromosome (NCBI, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, Md. 20894; Gene ID 781191).	3283	27
<i>LOC101905684</i>	--	2495	149



**Supplemental Table S4: The abundantly expressed (Top 20) non-coding RNAs in the bovine spermatozoa with unique gene reads.**

Transcript ID	TPM	Unique Gene Reads
NONBTAT002138.2	52526	361
NONBTAT029740.1	46868	3315
NONBTAT026069.2	39900	2036
NONBTAT026075.2	26794	2320
NONBTAT015718.2	14800	259
NONBTAT015717.2	13829	206
NONBTAT017665.2	13513	205
NONBTAT027794.1	7933	2971
NONBTAT007060.2	7921	29
NONBTAT029554.1	7846	6267
NONBTAT008960.2	7804	6206
NONBTAT022624.2	6306	3723
NONBTAT030080.1	6070	1454
NONBTAT030770.1	5672	8489
NONBTAT030342.1	4966	5823
NONBTAT030589.1	4619	249
NONBTAT031209.1	4369	4505
NONBTAT022014.2	4324	6851
NONBTAT025041.2	4239	205
NONBTAT001981.2	4089	557

**Supplemental Table S5: The pathways associated with spermatozoal RNAs. The analysis was carried out using transcripts with more than 100 TPM with *Bos taurus* as background using DAVID bioinformatics software.**

KEEG Pathway	P-value	Genes Count	Observed Genes ID
Ribosome	4.96E-09	23	<i>RPL27A, RPL27, RPL37, RPS2, RPS5, RPL28, RPS3, RPS18, RPS27, RPL32, RPL23, RPL18A, RPS29, RPL31, RPL6, RPL13A, RPLP0, RPS14, FAU, RPL10A, RPS11, UBA52, RPS23</i>
Axon guidance	0.01	12	<i>PAK6, MAPK1, EPHA7, EPHB6, PLXNA4, CXCR4, SEMA4F, SEMA3E, CFL1, SEMA3D, UNC5D, ARHGEF12</i>
Salmonella infection	0.01	9	<i>ACTB, FOS, PFN1, MAPK1, CCL3, IL18, IL1B, TLR5, FLNB</i>
Viral myocarditis	0.01	8	<i>ACTB, EIF4G2, PRF1, EIF4G3, BOLA-DOA, BLA-DQB, DAG1, SGCD</i>
Inflammatory bowel disease (IBD)	0.02	8	<i>IL12RB2, STAT4, IL18, BOLA-DOA, BLA-DQB, IL1B, TLR5, STAT1</i>
PPAR signaling pathway	0.02	8	<i>SLC27A1, GK2, UBC, FABP3, ACSBG2, RXRG, FABP1, CYP8B1</i>
Galactose metabolism	0.03	5	<i>PFKL, AKR1B1, PFKP, G6PC3, G6PC2</i>
Glutathione metabolism	0.05	6	<i>GPX2, GSTM3, SRM, GSTO1, GCLM</i>
Oocyte meiosis	0.07	9	<i>PLCZ1, IGF1R, MAPK1, YWHAZ, ADCY8, ANAPC13, SKP1, CDC25C, CALM2</i>
RNA degradation	0.07	7	<i>PFKL, BTG1, PFKP, RQCD1, DHX36, LSM2, CNOT6</i>

**Supplemental Table S6: Cluster analysis of commonly expressed genes in all animals detected by Ion Proton and Illumina platforms. These genes considered for this analysis were not differentially expressed between platforms.**

Category	Term	Genes Count	P-value
<b>Annotation Cluster 1</b>	<b>Enrichment Score: 2.60</b>		
UP_KEYWORDS	ATP-binding	37	0.00
GOTERM_MF_DIRECT	GO:0005524~ATP binding	22	0.00
UP_KEYWORDS	Nucleotide-binding	43	0.01
<b>Annotation Cluster 2</b>	<b>Enrichment Score: 1.60</b>		
INTERPRO	IPR019844:Cold-shock conserved site	3	0.01
INTERPRO	IPR002059:Cold-shock protein, DNA-binding	3	0.02
INTERPRO	IPR011129:Cold shock protein	3	0.02
SMART	SM00357:SM00357	3	0.02
INTERPRO	IPR012340:Nucleic acid-binding, OB-fold	5	0.11
<b>Annotation Cluster 3</b>	<b>Enrichment Score: 1.58</b>		
UP_KEYWORDS	Zinc-finger	31	0.01
UP_KEYWORDS	Metal-binding	60	0.03
UP_KEYWORDS	Zinc	38	0.05
<b>Annotation Cluster 4</b>	<b>Enrichment Score: 1.46</b>		
UP_KEYWORDS	TPR repeat	8	0.00
INTERPRO	IPR011990:Tetratricopeptide-like helical	11	0.01
UP_SEQ_FEATURE	repeat:TPR 2	4	0.07
UP_SEQ_FEATURE	repeat:TPR 1	4	0.07
UP_SEQ_FEATURE	repeat:TPR 3	3	0.19
<b>Annotation Cluster 5</b>	<b>Enrichment Score: 1.46</b>		
UP_KEYWORDS	TPR repeat	8	0.00
INTERPRO	IPR011990:Tetratricopeptide-like helical	11	0.01
INTERPRO	IPR013026:Tetratricopeptide repeat-containing domain	8	0.02
INTERPRO	IPR019734:Tetratricopeptide repeat	7	0.05
INTERPRO	IPR013105:Tetratricopeptide TPR2	3	0.20
SMART	SM00028:SM00028	5	0.20
<b>Annotation Cluster 6</b>	<b>Enrichment Score: 1.37</b>		
UP_KEYWORDS	Microtubule	11	0.00
GOTERM_BP_DIRECT	GO:0007017~microtubule-based process	3	0.06

GOTERM_CC_DIRECT	GO:0005874~microtubule	6	0.08
GOTERM_BP_DIRECT	GO:0008152~metabolic process	8	0.26
<b>Annotation Cluster 7</b>	<b>Enrichment Score: 1.36</b>		
KEGG_PATHWAY	bta03040:Spliceosome	12	0.00
INTERPRO	IPR013126:Heat shock protein 70 family	4	0.01
INTERPRO	IPR018181:Heat shock protein 70, conserved site	4	0.01
KEGG_PATHWAY	bta04915:Estrogen signaling pathway	9	0.01
KEGG_PATHWAY	bta05169:Epstein-Barr virus infection	13	0.01
KEGG_PATHWAY	bta05134:Legionellosis	6	0.03
KEGG_PATHWAY	bta04141:Protein processing in endoplasmic reticulum	10	0.06
UP_KEYWORDS	Stress response	5	0.07
KEGG_PATHWAY	bta04010:MAPK signaling pathway	13	0.09
KEGG_PATHWAY	bta05145:Toxoplasmosis	7	0.16
KEGG_PATHWAY	bta04612:Antigen processing and presentation	5	0.19
KEGG_PATHWAY	bta05164:Influenza A	8	0.27
UP_KEYWORDS	Chaperone	6	0.28
KEGG_PATHWAY	bta05162:Measles	6	0.41
<b>Annotation Cluster 8</b>	<b>Enrichment Score: 1.07</b>		
KEGG_PATHWAY	bta05231:Choline metabolism in cancer	8	0.03
KEGG_PATHWAY	bta04666:Fc gamma R-mediated phagocytosis	7	0.04
KEGG_PATHWAY	bta05132:Salmonella infection	4	0.46
<b>Annotation Cluster 9</b>	<b>Enrichment Score: 0.97</b>		
INTERPRO	IPR019821:Kinesin, motor region, conserved site	3	0.09
INTERPRO	IPR001752:Kinesin, motor domain	3	0.10
UP_KEYWORDS	Motor protein	5	0.10
SMART	SM00129:SM00129	3	0.13
<b>Annotation Cluster 10</b>	<b>Enrichment Score: 0.96</b>		
GOTERM_MF_DIRECT	GO:0005516~calmodulin binding	5	0.05
UP_KEYWORDS	Calmodulin-binding	5	0.05
KEGG_PATHWAY	bta04020:Calcium signaling pathway	7	0.49

**Supplemental Table S7: The primers used for assessing the contamination of epithelial cells, germ cells and leukocytes RNA and also spermatozoal gDNA in isolated spermatozoal total RNA.**

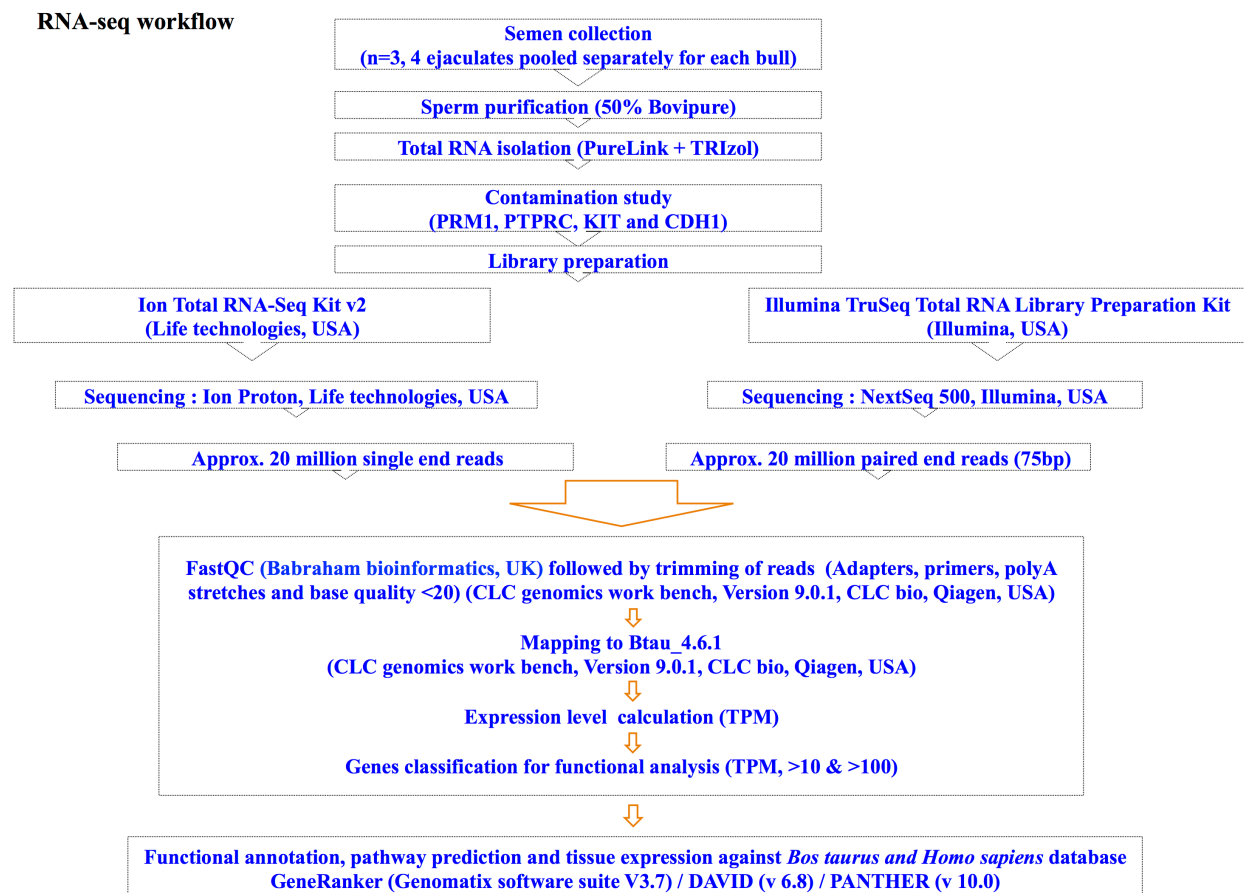
Primer ID	Primer Sequence (5' – 3')	Primer length (bp)	Product size (bp)	Intron Span (Yes/No)	Accession Number	Purpose
<i>CDHI-F</i>	CTGCATTCCTGGCTTTGGTG	20	171	Yes	NM_001002763.1	Epithelial cell
<i>CDHI-R</i>	GTAAGCACGCCATCTGTGTG	20				
<i>KIT-F</i>	GAATAGCTGGCATCAGGGTG	20	224	Yes	AF263827.1	Germ cell
<i>KIT-R</i>	CCAGATCCACATTCTCTCCATC	22				
<i>PTPRC-F</i>	ACCCAACCTTCTACTCAAGATG	22	124	Yes	NM_001206523.1	Leukocytes
<i>PTPRC-R</i>	CGTATTTGTTCTCACATGGTGG	22				
<i>PRM1-F</i>	AAGATGTTCGCAGACGAAGGAG	21	222	Yes	NM_174156.2	Sperm RNA & gDNA
<i>PRM1-R</i>	GTGGCATTGTTTCGTTAGCAGG	21				
<i>PRM1-F*</i>	ATGGCCAGATACCGATGC	18	156	Yes	NM_174156.2	Coding region
<i>PRM1-R*</i>	TTACTGTCTTGTACACCT	18				

\* The protamine1 (PRM1) primer covering coding region was used for confirmation of presence of spermatozoal RNA.

**Supplemental Table S8: The primers used for validating transcripts presence in spermatozoa. The transcripts were selected based on the expression levels (TPM) and read coverage.**

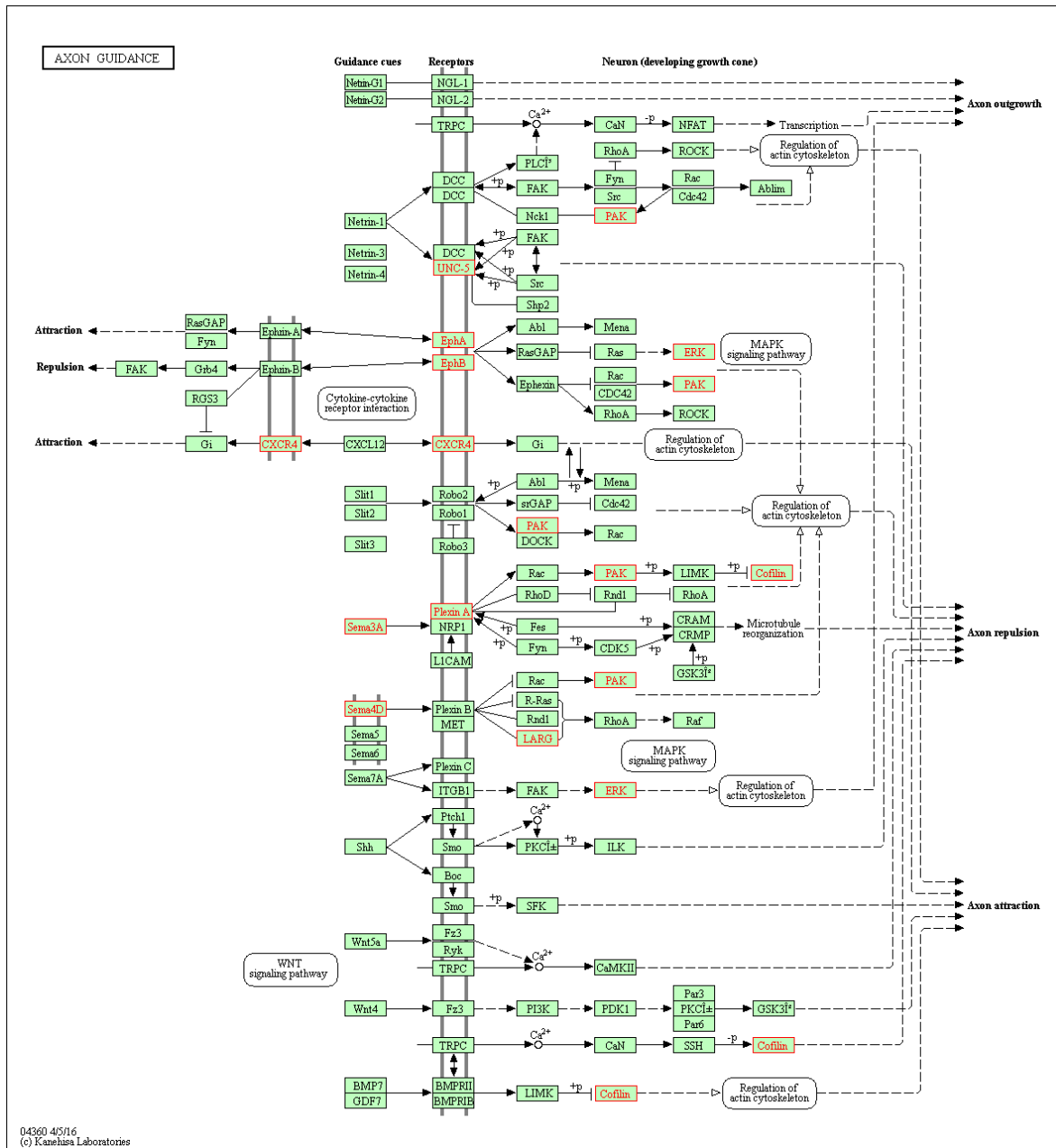
Primer ID	Primer sequence (5' – 3')	Primer length (bp)	Product size (bp)	Intron Span Yes/No	Accession Number
<i>HMGB4-F</i>	AGACGGATGTGGACAAGCAG	20	150	No	NM_001040562.3
<i>HMGB4-R</i>	GCTCGAACTCTCTGCACTGG	20			
<i>CHMP5-F</i>	TGACCTCGAAGCAGAGTTGG	20	120	Yes/E-E	NM_001034682.2
<i>CHMP5-R</i>	TTCGTGTCAGTGGGAACACC	20			
<i>YWHAZ-F</i>	GGAAATGCAACCAACACATCC	21	126	Yes/E-E	NM_174814.2
<i>YWHAZ-R</i>	ATGGCTTCATCAAATGCTGTC	21			
<i>SCP2D1-F</i>	AGTGCAGTGGGATGTGGAAG	20	133	No	NM_001040507.2
<i>SCP2D1-R</i>	TCTGACAGCGAGAGAGGGTG	20			
<i>GTSF1L-F</i>	GGCATCTGCTTTGTGTGGTC	20	137	No	NM_001079601.2
<i>GTSF1L-R</i>	ACAACCTCGGGCTGCCTATTC	20			
<i>C23H6orf62-F</i>	AGACAGCACTTGCAGACTCC	20	134	Yes/E-E	NM_001077999.2
<i>C23H6orf62-R</i>	GGCATGTAGGGATGGAGGTG	20			
<i>SDCBP2-F</i>	TCCAGAAAGTGCCAGTGCAG	20	126	No	XM_003586805.2
<i>SDCBP2-R</i>	CATAGGTGGATCTCGCGCAC	20			
<i>HSFY2-F</i>	AAGCAAGTGTCAACGAAGATG	21	128	No	NM_001077006.1
<i>HSFY2-R</i>	GACAATCTGGATGGTATGAGGC	22			
<i>OR6C2-F</i>	GCTGCTTGTGCCAGTCAAC	19	149	No	XM_001254810.2
<i>OR6C2-R</i>	ACGTTCTTGCAGACTCTAATGC	22			
<i>MGC157405-F</i>	TACACATGGACCGCATCTCC	20	135	Yes/E-E	NM_001083697.2
<i>MGC157405-R</i>	CGCCGATGAGCTTCTGTATG	20			
<i>ZNF280B-F</i>	GGGACCGGATGTGACCTTC	19	103	Yes/E-E	NM_001077935.1
<i>ZNF280B-R</i>	CCACAAGTCCCAGTGCTTCC	20			
<i>PLCZ1-F</i>	TGCATGAACAGAGGTTACCG	20	118	No	NM_001011680.2
<i>PLCZ1-R</i>	TGTCAAGTACCATGAGTGGC	20			

**Supplemental Figure 1: The work flow of spermatozoal RNA sequencing procedure and the softwares used for analyzing the transcriptome data.**



**Supplemental Figure 2: The involvement of abundant (>100 TPM) spermatozoal transcripts in axon guidance (a) and oocyte meiosis (b) based on KEGG pathway (<http://www.kegg.jp/kegg/kegg2.html>). The genes observed in the present study were highlighted in red colour.**

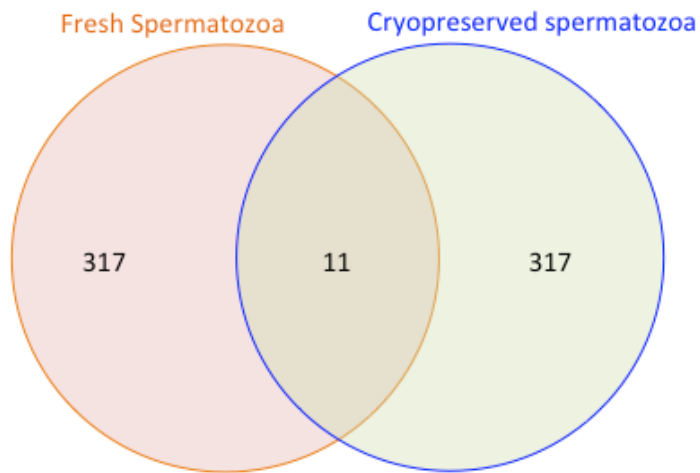
**a) Axon guidance**







**Supplemental Figure 3: The common number of genes (top 328 transcripts) observed in this current study from fresh bull spermatozoa were compared with the earlier study in cryopreserved sperm<sup>6</sup>.**



## References

1. Dogan, S. *et al.* Sperm protamine-status correlates to the fertility of breeding bulls. *Biol. Reprod.* **92**, 1-9 (2015).
2. Shim, J. H. *et al.* CHMP5 is essential for late endosome function and down-regulation of receptor signaling during mouse embryogenesis. *J. Cell Biol.* **172**,1045-1056 (2006).
3. Puri, P., Myers, K., Kline, D. & Vijayaraghavan, S. Proteomic analysis of bovine sperm YWHA binding partners identify proteins involved in signaling and metabolism. *Biol. Reprod.* **79**, 1183-1191(2008).
4. Yoshimura, T. *et al.* Gtsf1/Cue110, a gene encoding a protein with two copies of a CHHC Zn-finger motif, is involved in spermatogenesis and retrotransposon suppression in murine testes. *Develop. Biol.* **335**, 216-227 (2009).
5. Ing, N. H. *et al.* Sequencing of RNA from stallion sperm identifies potential markers of fertility. *J. Equine Vet. Sci.* **43**, S64 (2016).
6. Card, C. J. *et al.* Cryopreserved bovine spermatozoal transcript profile as revealed by high-throughput ribonucleic acid sequencing. *Biol. Reprod.* **88**, 49 (2013).
7. Schwarz, T., Prieler, B., Schmid, J. & Neesen, J. Molecular characterization of Ccdc181, a Hook1 and microtubule interacting protein. 11th YSA PhD Symposium, Abstract Book. Abstract no. 112. 72-73 (2015).
8. Liu, R. Z., Li, X. & Godbout, R. A novel fatty acid-binding protein (FABP) gene resulting from tandem gene duplication in mammals: transcription in rat retina and testis. *Genomics*, **92**, 436-445 (2008).

9. Nishimura, M., Naito, S. & Yokoi, T. Tissue-specific mRNA expression profiles of human nuclear receptor subfamilies. *Drug Metab. Pharmacokinet.* **19**, 135-149 (2004).
10. Lee, J., Lee, D. R. & Lee, S. The genetic variation in Monocarboxylic acid transporter 2 (MCT2) has functional and clinical relevance with male infertility. *Asian J. Androl.* **16**, 694 (2014).
11. Zhang, J. S., *et al.* Genome-wide profiling of segmental-regulated transcriptomes in human epididymis using oligo microarray. *Mol. Cell. Endocrinol.* **250**, 169-177 (2006).
12. Zuccarello, D. *et al.* How the human spermatozoa sense the oocyte: a new role of SDF1-CXCR4 signalling. *Int. J. Androl.* **34**, e554-e565 (2011).
13. Ushizawa, K. *et al.* Global gene expression analysis and regulation of the principal genes expressed in bovine placenta in relation to the transcription factor AP-2 family. *Reprod. Biol. Endocrinol.* **5**, 17 (2007).
14. Ferrer, M. *et al.* MMP2 and acrosin are major proteinases associated with the inner acrosomal membrane and may cooperate in sperm penetration of the zona pellucida during fertilization. *Cell Tissue Res.* **349**, 881-895 (2012).
15. Behmanesh, M. *et al.* Characterization of the structure and expression of mouse Itpa gene and its related sequences in the mouse genome. *DNA Res.* **12**, 39-51(2005).
16. Wenhao, T. *et al.* miRNA-mediated regulation of heat shock proteins in human ejaculated spermatozoa. *Turkish J. Med. Sci.* **45**, 1285-1291(2015).