739. The relevance of spermatozoal RNA and quality in livestock production

M.J. Hodge[1,2], S.J. Rindfleish[2], C.P.Stephen[1], S. de las Heras-Saldana[3*] and S.D. Pant[1*]

[1]*School of Agricultural, Environmental and Veterinary Sciences, Charles Sturt University, Wagga Wagga, NSW 2678, Australia;* [2]*Apiam Animal Health, Apiam Genetic Services, Dubbo, NSW 2830, Australia;* [3]*School of Environmental and Rural Science, University of New England, Armidale, NSW 2351, Australia;* *sdelash2@une.edu.au; spant@csu.edu.au*

Abstract

Spermatozoal cells are known to contain a wide variety of both coding and non-coding transcripts, but the physiological role of these transcripts is not fully understood. Therefore, the overall aim of this study was to characterize both coding and non-coding transcripts contained in ovine spermatozoa, and identify differences in expression patterns between three different sheep breeds (Merino, Dohne and Poll Dorset), and ejaculates that vary in quality. A variety of coding (e.g. *MVK, PAMR1, ATP2A3, KLK12,* etc.) and non-coding transcripts (e.g. XLOC_025216, XLOC_048757, XLOC_001809, XLOC_025093, etc.) were found to be differentially expressed in different sheep breeds and ejaculates of contrasting quality. Some of these transcripts have been reported to be associated with spermatozoal function or physiology in the literature. Overall, these results provide novel insights into spermatozoal transcriptomes in sheep, and also potentially indicate their functional relevance.

Introduction

Good reproductive performance is a crucial factor that influences the economics of livestock production, and successful conception is a key determinant of the reproductive performance of livestock. Consequently, artificial insemination is widely used in a variety of livestock species to maximise conception rates, while simultaneously exerting greater control over breeding, which in turn, facilitates increased genetic gain.

While the quality of semen used in artificial insemination is widely believed to influence conception outcomes during artificial breeding, poor conception outcomes can occur even when good quality semen is used (Hall 1981). This indicates that semen quality is not the sole determinant of successful conception, and additional molecular mechanisms functioning on the ram side may exist. The role of spermatozoal transcripts is interesting in this regard, as spermatozoa are known to contain thousands of transcripts, which are transferred to ova during fertilisation. It is generally understood that activation of embryonic genome is preceded by degradation of maternal transcripts post fertilisation. Therefore, the transcripts contained in spermatozoa that are transferred to the ova upon fertilization, could have an important role to play in these events.

Therefore, characterising spermatozoal transcripts, both mRNA and long non-coding RNA (lncRNAs) can be useful in increasing the likelihood of successful conception, which in turn has the potential to increase the efficiency and profitability of livestock farming.

Materials & methods

The methodology involved in semen collection, assessment, RNA isolation, library preparation, sequencing, and differential gene expression analysis to identify coding transcripts has already been published (Hodge *et al.* 2021). RNA samples extracted from spermatozoa were subjected to next-generation sequencing using the Illumina Hiseq2000 platform following manufacturer's instructions. Following sequencing, the quality of reads was assessed with FastQC v0.11.5. Poor quality bases (Phred score Q<30), adaptors, and overrepresented sequences were filtered out with trimmomatic v0.36 (Bolger *et al.* 2014).

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_739,

Identification of Long non-coding RNAs (lncRNAs). Analyses to identify lncRNAs involved aligning cleaned reads to the reference genome (Ovis aries, Oar v.4) via STAR v2.7.7 (Dobin *et al.* 2013). The assemblies were merged with cuffmerge from Cufflink v2.2.1 and the GTF file produced was used in FeatureCounts function from Subread v2.0.3 software to obtain the transcript counts in each sample. A differential expression analysis was performed with DESeq2 v1.30.1 using edgeR v3.32.1 (Robinson *et al.* 2010) to filter the lowly expressed genes based on the logarithmic counts per million (lcpm) and TMM normalised gene counts (trimmed mean of M values). Ejaculate quality was used as a fixed effect in the statistical model to identify statistically significant changes in transcript abundances assessed between breeds (FDR<0.05 and logFC>1). Subsequently, a total of 9 Merino, 9 Dohne and 9 Poll Dorset rams, which passed quality control in edgeR, were included in the differential gene expression analysis. Four contrasts were performed, comparing transcript expression between breeds (Poll Dorset vs Merino, Dohne vs Merino, and Dohne vs Poll Dorset), and ejaculates of relatively high- and low-quality.

Transcripts longer or equal to 200 nt were identified as lncRNAs, and Cuffcompare v2.2.1 was used to classify transcripts. Only transcripts classified as 'u', 'o', and 'x' were retained for further analysis. Putative lncRNA candidates were then identified if their coding potential score was calculated as <0 via CNCI v2 and their coding probability was calculated as <0.5 via CPC2 v1.0.1. A miRNA database (http://rumimir.sigenae.org) on sheep was then used to remove lncRNAs that had a potential miRNA precursor (E-value 0.001) in BLAST v2.12.0. Transcripts with ORF larger than 300nt were identified using getorf function of EMBOSS (Rice *et al.* 2000) and removed. Transcript sequences were cross-matched against: (1) a protein sequence database (Pfam library) using hmmscan software; (2) the UniProt/SwissProt database; and (3) the annotated proteins for *O. aries* v4 using blastx. In all steps, any transcripts with coding potential were eliminated and remaining transcripts with an average FRKM>0.1 per breed, containing more than one exon were considered as lncRNAs, and retained for further analysis. The identified lncRNAs were used to find target genes (cis-target) within a 10 kb window of the lncRNA. Trans-target genes were identified based on a co-expression analysis using Pearson's correlation (r>±0.95 with *P*-value<0.05) between the mRNA genes and lncRNA.

Results

Differential gene expression analysis identified a number of Differentially Expressed Genes (DEGs) in each contrast that was performed. A summary of the total number of DEGs, as well as the gene symbols of the top 10 DEGs identified in each contrast, is presented in Table 1.

Analyses involving the identification of long non-coding RNAs identified a total of 1,205 lncRNAs in sheep spermatozoa, of which 150 transcripts passed the quality filtering for low expression and only 6 lncRNA were found to be differentially expressed. A summary of key differentially expressed lncRNAs along with their target genes is presented in Table 2.

Table 1. Number of differentially expressed genes in breed specific contrasts, and between ejaculates of varying quality (Hodge *et al.* 2021).

Contrast	No. of DEGs	Top DEGs
Dohne vs Merino	72	*MVK, TNS3, ALAS1, CCN1, MSRA, RRP15, PPARD, MED6, CARMIL1, FLRT2*
Dohne vs Poll Dorset	73	*PAMR1, LY6E, ARGLU1, CDO1, VARS2, TP53I11, POLK, FBXL14, MAN1A1, CHI3L1*
Merino vs Poll Dorset	570	*ATP2A3, SLC35A5, BORCS5, MSRA, KRT4, FAM210B, NKPD1, ITM2C, ANAX2, SLC2A3*
High vs Low quality	39	*KLK12, SPEM2, LYPLA1, TNNC1, LYRM4, OXCT2, SMKR1, FAM57A, SRGN*

Table 2. Differentially expressed target genes and their co-expressed lncRNAs across the comparisons of breeds sampled and ejaculate quality.

Contrast	DE lncRNA (Target Genes)
Dohne vs Poll Dorset	XLOC_048235 *(C1H1orf141, SLC38A4, XPNPEP3, NUP43, ELP4, CHRM1, PTH2R, FBXO7, ABCC9)*
	XLOC_072802 (LOC105607053, LOC105609923)
Merino vs Poll Dorset	XLOC_001101 *(CSF3R, CLEC5A, TRNAC-GCA, GRAMD1A, TREML2)*
	XLOC_001727 *(CLIC6, EX26)*
	XLOC_066444 *(FBXO7, CABS1, PPM1E, MED31)*
High vs Low quality	XLOC_023268 (LOC105611383, *SOBP)*

Discussion

There are a variety of factors on both the ram and ewe sides that influence conception outcomes in sheep, particularly when artificial breeding technologies like artificial in-semination are used. While a significant emphasis is placed on semen quality in artificial breeding programs, it is clear that other determinants influencing conception outcomes remain uncharacterized. The role of spermatozoal transcripts is of significant interest in this regard, because while spermatozoa have been shown to contain a large repertoire of transcripts which include both protein-coding mRNA and non-coding RNAs, very little is known about their influence on ejaculate quality and fertilisation.

This current study provides evidence indicating that a variety of mRNA and lncRNAs are expressed in sheep spermatozoa. A very high number of spermatozoal cells were needed for RNA extraction, because each spermatozoa conceivably carries a small amount of RNA relative to normal somatic cells. This supports the notion that spermatozoal cells are likely transcriptionally inactive for the most part.

However, in our study, we have identified several transcripts in ovine spermatozoa (e.g. *CABS1*, *CD82*, *ATP1B3*, *EFHB,* etc.), that have been previously associated with spermatogenesis, sperm morphology and ejaculate quality parameters. Therefore, it is likely that such spermatozoal transcripts play a physiological role in the development of spermatozoa, and could therefore be used as markers. Our study also identified several transcripts that were differentially expressed between spermatozoa sourced from ejaculates that varied in quality. Therefore, it is also possible that some of these transcripts can be used as markers to identify good quality semen that is likely to yield desirable conception outcomes.

Furthermore, the spermatozoal transcriptomes in different species have been found to contain thousands of transcripts, which is consistent with our findings (Selvaraju *et al.* 2017; Prakash *et al.* 2021). These transcripts are transferred from spermatozoon to the ova during fertilisation, where they could potentially play a crucial role in directing early embryonic development, thereby influencing pre-natal growth that may be correlated with economically important traits like birth weight, etc.

Overall, our findings accord with a variety of studies focused on spermatozoal transcripts in other species that also support a functional role for these transcripts. Therefore, our study offers several novel insights indicating spermatozoal transcripts could have crucial physiological roles and could potentially influence conception outcomes, and even early embryonic development. Therefore, future in depth investigations aiming to characterise the physiological role of spermatozoal transcriptomes are warranted.

References

Bolger A.M., Lohse M. and Usadel B. (2014) Bioinformatics 30(15): 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Dobin A., Davis C.A., Schlesinger F., Drenkow J., Zaleski C. *et al.* (2013) Bioinformatics 29(1): 15-21. https://doi.org/10.1093/bioinformatics/bts635

Hall J.L. (1981) Fertil Steril 35(4): 457-463. https://doi.org/10.1016/s0015-0282(16)45444-0

Hodge M.J., De Las Heras-Saldana S., Rindfleish S.J., Stephen C.P. and Pant S.D. (2021) Genes (Basel) 12(2). https://doi.org/10.3390/genes12020203

Prakash M.A., Kumaresan A., Ebenezer Samuel King J.P., Nag P., Sharma A. *et al.* (2021) Front Cell Dev Biol 9(647717. https://doi.org/10.3389/fcell.2021.647717

Rice P., Longden I. and Bleasby A. (2000) Trends Genet 16(6): 276-277. https://doi.org/10.1016/s0168-9525(00)02024-2

Robinson M.D., Mccarthy D.J. and Smyth G.K. (2010) Bioinformatics 26(1): 139-140. https://doi.org/10.1093/bioinformatics/btp616

Selvaraju S., Parthipan S., Somashekar L., Kolte A.P., Krishnan Binsila B. *et al.* (2017) Sci Rep 7(42392. https://doi.org/10.1038/srep42392