Journal of Animal Science, 2023, **101**, 1–8
https://doi.org/10.1093/jas/skad098
Advance access publication 30 March 2023
Animal Genetics and Genomics

Blood transcriptome of Rasa Aragonesa rams with different sexual behavior phenotype reveals *CRYL1* and *SORCS2* as genes associated with this trait

Kenza Lakhssassi[†,‡], María Pilar Sarto[†], Belén Lahoz[†], José Luis Alabart[†], José Folch[†], Malena Serrano[‖], and Jorge Hugo Calvo[†,$,1]

[†]Agrifood Research and Technology Centre of Aragon-IA2, 50059 Zaragoza, Spain
[‡]INRA Instituts, 6356 Rabat, Morocco
[‖]Department of Animal Breeding and Genetics, INIA-CSIC, 28040 Madrid, Spain
[$]ARAID, 50018 Zaragoza, Spain

[1]Corresponding author: jhcalvo@cita-aragon.es

Abstract

Reproductive fitness of rams is seasonal, showing the highest *libido* during short days coinciding with the ovarian cyclicity resumption in the ewe. However, the remarkable variation in sexual behavior between rams impair farm efficiency and profitability. Intending to identify in vivo sexual behavior biomarkers that may aid farmers to select active rams, transcriptome profiling of blood was carried out by analyzing samples from 6 sexually active (A) and 6 nonactive (NA) Rasa Aragonesa rams using RNA-Seq technique. A total of 14,078 genes were expressed in blood but only four genes were differentially expressed (FDR < 0.10) in the A vs. NA rams comparison. The genes, *acrosin inhibitor 1* (*ENSOARG00020023278*) and *SORCS2*, were upregulated (log2FC > 1) in active rams, whereas the *CRYL1* and *immunoglobulin lambda-1 light chain isoform X47* (*ENSOARG00020025518*) genes were downregulated (log2FC < -1) in this same group. Gene set Enrichment Analysis (GSEA) identified 428 signaling pathways, predominantly related to biological processes. The lysosome pathway (GO:0005764) was the most enriched, and may affect fertility and sexual behavior, given the crucial role played by lysosomes in steroidogenesis, being the *SORCS2* gene related to this signaling pathway. Furthermore, the enriched positive regulation of ERK1 and ERK2 cascade (GO:0070374) pathway is associated with reproductive phenotypes such as fertility via modulation of hypothalamic regulation and GnRH-mediated production of pituitary gonadotropins. Furthermore, external side of plasma membrane (GO:0009897), fibrillar center (GO:0001650), focal adhesion (GO:0005925), and lamellipodium (GO:0030027) pathways were also enriched, suggesting that some molecules of these pathways might also be involved in rams' sexual behavior. These results provide new clues for understanding the molecular regulation of sexual behavior in rams. Further investigations will be needed to confirm the functions of *SORCS2* and *CRYL1* in relation to sexual behavior.

Lay Summary

Analyzing ram sexual behavior via blood transcriptome profiling can help to identify in vivo sexual behavior biomarkers as an innovative alternative to invasive and time-consuming methods in farms. Using RNA-sequencing technique, we compared 12 Rasa Aragonesa rams with different sexual behavior (6 sexually active and 6 nonactive) to identify differentially expressed genes (DEGs) in peripheral blood putatively responsible of *libido* differences between rams. Comparative analysis revealed four candidate genes and several signaling pathways related to sexual behavior such as lysosome, and positive regulation of the extracellular signal-regulated kinase 1/2 (ERK1 and ERK2) cascade. This data will be helpful for further investigations to understand the differences of sheep sexual behavior.

Keywords: sexual behavior, transcriptome, rams, blood, *CRYL1*, *SORCS2*.

Abbreviations: A, active; BCS, body condition score; BDNF, brain-derived neurotrophic factor; BP, biological process; CC, cellular component; CITA, Agrifood Research and Technology Centre of Aragon; CPM, count per million; DEG, differential expressed gene; ES, enrichment score; FDR, false discovery rate; GEO, Gene Expression Omnibus; GnRH, gonadotropin-releasing hormone; GO, gene ontology; GSEA, gene set enrichment analysis; LH, luteinizing hormone; Log2FC, log2 fold change; LW, live weight; MF, molecular function; NA, not active; NES, normalized enrichment score; PCA, principal component analysis; RIN, RNA integrity number; RT-qPCR, reverse transcription quantitative real-time PCR; SNP, single-nucleotide polymorphism; TMM, trimmed mean of M-values; trkB, tyrosine receptor kinase B

Introduction

In livestock production systems, fertility is a key factor of farms efficiency. In Mediterranean sheep, this is particularly important since the animals are seasonal breeding and mate in autumn and winter, hence the appellation "short-day breeders". This seasonality impedes the balanced supply throughout the year (Hazard, 2010) and consequently induces major variation in lamb production as well as disparity in the market price. To improve flock sustainability and profitability, several methods have been developed to avoid the seasonality such as the use of hormonal and/or photoperiodic treatments (Chemineau et al., 1988). For ethical reasons, breeders are forced to reduce the use of these treatments (Martin and

Received January 25, 2023 Accepted March 28, 2023.

© The Author(s) 2023. Published by Oxford University Press on behalf of the American Society of Animal Science. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

Kadokawa, 2006). The male effect is used in sheep as a strategy to improve flock fertility during seasonal anestrus period. The ram emits a considerable range of different sensory stimuli (courtship, odor and vocalizations) that may be responsible for inducing a reproductive neuroendocrine response in ewes by stimulating luteinizing hormone (LH) secretion (Fabre-Nys et al., 2015). However, differences in sexual behavior among individuals have been reported (Lindsay, 1996). Rams expressing high sexual behavior produce a greater stimulus during seasonal anestrus, which lead into a higher percentage of mated ewes and higher fertility (Delgadillo et al., 2008). On the other hand, rams with low sexual behavior extend the lambing season, and decrease the number of lambs born per ewe lambing, leading to an increase of farm costs (Alexander et al., 2012). This is why ram selection is fundamental to increase the efficiency of sheep farming (Simitzis et al., 2006), and sexual behavior is one of the most important traits to be considered during the selection of replacement rams and ewes. Yet, selection processes rarely include an evaluation of sexual behavior, such as sexual behavior pen tests, because of time, labor and facilities constraints (Alexander et al., 2012). Hence, alternative solutions must be sought, such as the identification of sexual behavior biomarkers. In a previous study, aiming to determine the possible molecular mechanisms underlying the sexual behavior of Rasa Aragonesa rams, our research group investigated transcriptional changes in the hypothalamus, *pars tuberalis*, and pineal gland in rams with different sexual behavior, identifying important differential expressed genes related to sexual behavior such as *MTNR1A, CHRNA2, FSHB, LHB, GNRHR, AVP, PRL, PDYN, CGA, GABRD*, and *TSHB* (Lakhssassi et al., 2023). Nevertheless, brain gene expression studies are highly invasive because they are based on *post mortem* tissue samples. Intending to identify in vivo sexual behavior biomarkers that may aid farmers in breeding decisions, this study investigate the differentially expressed genes in peripheral blood between active (A) and nonactive (NA) rams of Rasa Aragonesa sheep breed using RNA-Seq technique.

Material and Methods

All experimental procedures, including care of animals and euthanasia, were performed in accordance with the guidelines of the European Union regulations for the use and care of animals in research (Directive 2010/63/EU) and approved by the Animal Ethics Committee of the Research Centre (protocol number 2017/02).

Animals' selection

Animal selection, sexual behavior test and classification into sexual active/nonactive rams are described in Lakhssassi et al. (2023). Briefly, 49 previously exposed to ewes during breeding season Rasa Aragonesa rams (1.8 to 8.3 years of age) were submitted to individual sexual behavior pen tests. These rams were exposed to three adult ewes synchronized in estrus, and their sexual behavior was recorded. Sexual activity phenotype was defined as the sum of mounts and services (Lakhssassi et al., 2023). Blood samples for testosterone determination and RNA analyses (validation studies) were taken from all rams. Body condition score (BCS: 3.3 ± 0.3; 1 to 5 scale, according to Russel et al. (1969)), live weight (LW: 81.5 ± 14.44 kg) and mean antero-posterior maximum testicular diameter, subtracting the scrotal skin thickness, using a caliper (Folch and Roca Bernaus, 1982) (6.00 ± 0.58 cm) were also measured. A tree-based regression model for the sexual activity phenotype and a principal component analysis (PCA) were performed, considering all the variables for the sexual activity phenotype (age, BCS, LW, testosterone, and testicular diameter), for choosing a population subgroup with similar values for the significant variables, but different sexual activity phenotype for next analyses (Lakhssassi et al., 2023). Then, we could choose a subgroup of 21 rams with similar age and BCS (significant variables) that were submitted to a second individual sexual behavior pen test fifteen days later than the first one to confirm the results from the first pen test. Rams from this group ($n = 21$) were classified as either active (A) (9.0 ± 7.5, average mounts plus services ± standard deviation (sd); $n = 11$) or nonactive (NA; without mounts or services; $n = 10$). Six rams of each group were selected the day after the second sexual behavior pen test for taking the blood sample for RNA-Seq (6 A and 6 NA rams). Blood samples taken from the total ram population ($n = 59$) the day of the first pen test were used for validation of RNA-Seq data by reverse transcription quantitative polymerase chain reaction (RT-qPCR).

Blood samples, RNA extraction, and sequencing

The day after the second pen test, blood samples were taken in Tempus tubes (Applied Biosystems, Thermofisher, UK) and stored at – 80 °C until total RNA extraction. The Tempus Spin RNA Isolation Reagent Kit was used to extract the total RNA from peripheral blood (Applied Biosystems, Thermofisher, UK) according to the manufacturer´s instructions. The Agilent 2100 Bioanalyzer machine (Agilent Technologies, Palo Alto, CA) was used to check for RNA concentration and RNA Integrity Number (RIN) values. For library construction, only samples with a RIN > 7 were maintained for sequencing. The RNA-Seq library was prepared at the CNAG (Centro Nacional de Análisis Genómico, Spain; https://www.cnag.crg.eu/) as described by Lakhssassi et al. (2023), and sequencing was carried out with Novaseq 6000 (Illumina) generating paired-end reads of 151 bp. The RNA-Seq sequencing data have been deposited in the NCBI's Gene Expression Omnibus repository (Barrett et al., 2013) with the accession number GSE218667 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE218667).

Validation of RNA-Seq data by RT-qPCR

The validation of the RNA-Seq results was done in the 59 rams sexually characterized in the first pen test (15 days before taking blood samples for RNA-Seq study). Blood samples from the total ram population were collected the day of the first pen test in Tempus tubes and stored at – 80 °C until total RNA extraction. RNA extraction was performed as described above. Three genes identified to be differentially expressed in the 12 blood samples by RNA-Seq were chosen for RT-qPCR validation. These genes were: *acrosin inhibitor 1*, *CYRL1* and *SORCS2*. *Acrosin inhibitor 1*, and *CYRL1* genes were DEGs in the two scenarios, while *SORCS2* was differential expressed only in the first scenario, but its function was related to sexual behavior. Furthermore, four housekeeping genes (*B2M*, *GAPDH*, *RPL19*, *RPL32*) commonly used as reference genes in expression studies, were used to normalize the gene expression of *acrosin inhibitor 1*, *CYRL1*, and *SORCS2*. The gene expression stability was studied using NormFinder (Andersen et al., 2004). Initially, we treated total RNA (1 µg) from each sample with DNAse (Invitrogen, Carlsbad, CA, USA). Then,

the Super-ScriptIII Reverse Transcriptase kit (Invitrogen) was used following the manufacturer's recommendations, to synthesize first-strand cDNA. Primer3 software was used (https://primer3.ut.ee/; last accessed on November 11, 2022), to design the primers in specific exon-spanning regions. Conventional PCR was achieved using standard conditions for all the genes to confirm the specificity of the primers. Gene identity was confirmed by Sanger sequencing of the PCR products by the company Stabvida (https://www.stabvida.com/es). Homology searches were performed with BLAST (National Center for Biotechnology Information: https://blast.ncbi.nlm.nih.gov/Blast.cgi; last accessed on November 11, 2022). The quantitative real-time PCR was carried out in a 10 µl PCR total reaction mixture containing SYBR Green Master Mix: SYBR Premix Ex Taq II (Tli RNase H Plus, Takara, Demlab, Zaragoza, Spain). Reactions were run in triplicate on an ABI Prism 7500 platform (Applied Biosystem, Madrid, Spain) following the manufacturer's cycling parameters. Table 1 shows the annealing temperatures, primer concentrations, and primer sequences. The efficiency (E) and correlation coefficient (R^2) of PCR amplification for each gene were assessed using the standard curve method using the formula $E = 10^{(-1/\text{slope})}$, through 10-fold serial dilutions of pooled cDNA. Two "connector samples" were replicated in all plates to remove technical variation from this source of variability.

Statistical analysis

RNA-Seq analysis

All transcriptome analyses were carried out within OmicsBox platform v2.0.36 from BioBam's (https://www.biobam.com/omicsbox; last accessed on November 11, 2022). The quality of reads was assessed using FastQC tools (Andrews, 2015). The adapters sequences and low-quality reads were trimmed running the Trimmomatic program (Bolger et al., 2014) using the Sliding Window Trimming option with a window size of 40, and a required quality of 20. The filtering configuration of reads was set to 20 and 35 for the average quality and the minimum length, respectively. We mapped the obtained clean reads to the ovine reference genome Oar_rambouillet v1.0 (GCA_002742125.1) using STAR v2.7.8a (Dobin et al., 2013) supported by an annotated gene file in GTF format. Gene-level quantification was estimated by HTSeq (Anders et al., 2015) using the Union model as overlapping mode, gene as feature and strand specific forward. For subsequent analysis, only reads mapping unambiguously to a single genomic feature were considered. EdgeR software (Robinson et al., 2010) was run to detect the differentially expressed genes (DEGs) using a simple design where A rams was defined as the primary contrast condition and NA rams as primary reference condition. A threshold of count per million (CPM) greater than or equal to 0.5 in at least two samples and the trimmed mean of M-values (TMM) with Zero Pairing normalization method were applied to filter the low count genes. The statistical analysis was achieved using the exact test with robust option to avoid potential outlier genes. Multiple-testing corrections were performed using the Benjamini and Hochberg step-up false discovery rate (FDR) procedure to calculate adjusted P-values. To judge the significance of gene expression difference, the absolute log2FC ≥ 1 and FDR ≤ 0.10 were used as the threshold. The choice of relaxed FDR (<10%) was based on that our objective was mainly exploratory, so it is convenient to retain some information that we could lose being more restrictive.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) implemented in OmicsBox (https://www.biobam.com/omicsbox) were applied to identify potential pathway responsible to cause differences in sexual behavior between rams. GSEA includes in the analyses all the expressed genes, regardless of whether they were differentially expressed, and determines both modest and robust, coordinated, biologically relevant changes in molecular signaling pathways, and explores and identifies the most significant and over-represented gene ontology (GO) terms. In this study, GSEA was performed to determine whether an a priori defined set of genes is statistically significant between A and NA rams in blood, classifying the biological functions of the set genes on molecular functions (MFs), biological process (BP), and cellular component (CC). GSEA Enrichment scores (*ES*) were calculated according to the rank-ordered gene list and GO terms containing a minimum of 15 genes. A normalized enrichment

Table 1. Characteristics of reverse transcription quantitative real-time PCR (RT-qPCR) for validation of RNA-Seq for each gene

Genes	Primers	Amplification Size (bp)	AT[1]	E[2]	R^2*
RPL19	F-5′caactcccgccagcagat-3′	76	60°C	1.99	0.99
	R-5′ccgggaatggacatgcaca-3′				
RPL32	F-5′ggcaccagtcagaccgatatg-3′	75	60°C	1.99	0.99
	R-5′cccagacaggagaggttcag-3′				
GAPDH	F-5′tccatgaccactttggcatcgt-3′	80	60°C	2.05	0.99
	R-5′gtcttctgggtggcagtga-3′				
B2M	F-5′ggtgctgcttagaggtctcg-3′	109	60°C	1.99	0.99
	R-5′acgctgagttcactcccaac-3′				
CRYL1	F-5′ tgttcgacattgagcctcgg-3′	124	60°C	2.05	0.99
	R-5′atgagtgacagctgctcctc-3′				
SORCS2	F-5′agatcagcttcctcctgcg-3′	97	60°C	2.06	0.99
	R-5′ cagtagccgcctccacttc-3′				

[1]Annealing temperature.
[2]Efficiency of PCR amplification
*Correlation coefficient.

score (*NES*) was also calculated considering the number of genes in the pathway. Genes enriched at the top of the ordered fold change list had a positive *NES*, while a negative *NES* indicates the opposite, i.e., enriched at the bottom. Pathways with an FDR 5% were considered significant.

Gene expression validation of RNA-Seq by RT-qPCR

The statistical analysis was done in the 59 rams sexually characterized in the first pen test. Furthermore, a second analysis considering only the 47 rams not used for RNA-seq was run. The corresponding mRNA levels were measured and analyzed by their quantification cycle (Cq), following the statistical method proposed by Steibel et al. (2009). The mixed model fitted was as follows:

$$y_{rigkm} = SG_{gi} + P_k + BC_m + A_m + LW_m + E_m + e_{rigkm}$$

where y_{rigkm} is the Cq value (transformed data taking into account gene amplification efficiencies) of the *g*th gene (DEGs and housekeeping genes) from the *r*th well (reactions were run in triplicate) in the *k*th plate corresponding to the *m*th animal and to the *i*th sexual activity group (NA without any mount or service or A with at least one mount or service in the first ram pen test); SG_{gi} is the fixed interaction between the *i*th sexual activity group and the *g*th gene; P_k is the fixed effect of the *k*th plate; BC_m, A_m, and LW_m are the effects of body condition score, age and the live weight of the *m*th animal, respectively, included as covariates; E_m is the random effect of the *m*th animal from which samples were collected $\left(E_m : \left(0, \sigma^2{}_E\right)\right)$; and e_{rigkm} is the random residual. Gene-specific residual variance (heterogeneous residual) was fitted to the gene by sexual behavior $\left(e_{rigkm} : N\left(0, \sigma^2{}_{egi}\right)\right)$.

To test differences in the expression rate of the genes of interest ($diff_{DEGs}$) between treatments in terms of fold change (FC), the approach suggested by Steibel et al. (2009) was used using the three housekeeping genes. The significance of the of the $diff_{DEGs}$ estimates was determined with the *t*-test ($P < 0.05$) after Bonferroni correction.

Results

Summary of the raw sequence reads

RNA-Seq data were obtained for 12 blood samples (6 A and 6 NA rams). The statistics of the sequencing and read alignments against the ovine reference genome Oar_rambouillet v1.0 (GCA_002742125.1) are listed in Supplementary Table S1. Illumina sequencing produced a total of 1,024,395,857 raw reads from blood samples. After removing low-quality bases and adaptor sequences, the remained clean reads were 905,265,453 which provided abundant data for further analysis. On average, 97.7% of all the clean reads were mapped to the ovine reference genome of which 90.1 % were uniquely mapped. These results demonstrated good data quality suitable for subsequent research analysis (scenario 1). However, we observed that for sample 1, more than a half of input raw reads were trimmed. In total, 76% of all read bases included in library 1 were trimmed off, while it was less than 10% for the other libraries (Supplementary Table 1). So, we performed a second analysis without this sample to check a putative effect on downstream analyses (scenario 2).

Differentially expressed genes (DEGs)

We next investigated the differences in gene expression data between A and NA rams. As a result, 26,478 genes were expressed in the blood samples. After filtering, a total of 14,078 were retained for the differential expression analysis. Only four differentially expressed genes (FDR = 0.1) were identified of which two were significant with a FDR < 0.05 (*CRYL1* and *acrosin inhibitor 1*) between A and NA rams (Table 2) in the scenario 1. The *acrosin inhibitor 1* and *SORCS2* genes were upregulated (log2FC > 1) in the A rams whereas *CRYL1* and *immunoglobulin lambda-1 light chain isoform X47* (*ENSOARG00020025518*) genes were downregulated (log2FC < -1) in this same group of rams. The results of the second analysis (5 A and 6 NA rams) revealed that the results of the study do not change greatly in relation to DEGs. A total of 14,021 genes remained after filtering. Similarly, only four DEGs were detected (FDR = 0.1). Three genes were upregulated, and one were downregulated (Table 2). The results of this analysis also showed *CRYL1* and *Acrosin inhibitor 1* genes in common between the two analyses. On the other hand, the *GPC3* and *OSBP2* genes were found upregulated under this scenario. The FC were similar between the two analyses. In this sense, the correlation coefficient between the four significant DEGs in the first analysis with respect to the results of the same genes in the second analysis was 0.99 and vice versa. The FCs for *CRYL1*, *LOC121819654*, *SORCS2*, and *ENSOARG00020025518* were –5.6, 36.30, 2.72, and

Table 2. Differentially expressed genes (DEGs) in the blood through the comparison between active and nonactive rams in scenario 1 (*n* = 12; 6 active vs 6 nonactive rams) and 2 (*n* = 11; 5 active vs 6 nonactive rams)

	Genes	Description	FC	Log2FC[1]	*P*-adj[2]
Scenario 1	*CRYL1*	*Crystallin lambda 1*	–5.60	–2.48	2.19E–08
	LOC121819654	*Acrosin inhibitor 1*	36.30	5.18	0.006
	SORCS2	*Sortilin related VPS10 domain containing receptor 2*	2.73	1.45	0.095
	ENSOARG00020025518	*Immunoglobulin lambda-1 light chain isoform X47*	–2.48	–1.30	0.099
Scenario 2	*CRYL1*	*Crystallin lambda 1*	–5.46	–2.45	0
	GPC3	*Glypican 3*	3.28	1.71	0.0041
	LOC121819654	*Acrosin inhibitor 1*	42.17	5.40	0.0002
	OSBP2	*Oxysterol binding protein 2*	41.70	5.38	0.0932

[1]Log2Fold Change.
[2]*P* adjusted.

–2.48, respectively, while in the second analysis were –5.46, 42.17, 2.71, and –2.39.

Enrichment analysis

The GSEA revealed 428 signaling pathways, mainly dominated by biological processes (291 BP, 74 MF, and 63 CC). The top 30 most enriched terms between the A and NA rams are shown in Figure 1. The enrichment score and the gene set size details for these GOs are described in Supplementary Table S2. The most enriched pathways were the lysosome (GO:0005764), external side of plasma membrane (GO:0009897), fibrillar center (GO:0001650), and focal adhesion (GO:0005925).

Verification of DEGs by RT-qPCR

To validate the RNA-Seq results, the expression fold changes of three genes (*acrosin inhibitor 1*, *CRYL1*, and *SORCS2*) were tested by using RT-qPCR in RNA samples of 59 rams of the first pen test. Unfortunately, the *acrosin inhibitor 1* could not be validated due to its very low expression in blood samples, with an average of 12.7 counts (number of reads aligned to each genomic feature). The gene stability analysis with NormFinder program revealed that the housekeeping genes were more stable than the DEGs (Supplementary Figure S1). Therefore, the RT-qPCR results were normalized with the three more stable housekeeping genes: *B2M*, *GAPDH*, and *RPL32*. The comparison A vs. NA rams was performed considering active those that had at least 1 mount or service. The active ones presented 14.15 ± 9.61 mounts + service (n = 35), whereas the NA rams did not have any mount or services (0, n = 24). The two genes *CRYL1* and *SORCS2* were significant as in RNA-Seq but in lower magnitude (FC of –1.48 and 1.64 vs –5.60 and 2.73 for *CRYL1* and *SORCS2*, in RT-qPCR vs RNA-Seq) considering the 59 rams (Figure 2). The analysis considering only the 47 rams not used for RNA-seq also confirmed that the expression of both genes was significant affected (FC of –1.29 and 1.99 for *CRYL1* and *SORCS2*, respectively). In general, the expression patterns of these two genes were consistent with those obtained with the RNA-Seq technique.

Discussion

Understanding the molecular mechanisms underlying the sexual behavior of rams is crucial to improve flock sustainability and profitability. RNA-Seq technology is a valuable tool allowing large scale analysis of the transcriptional changes related to biological conditions of interest (Crow et al., 2019). Ram's blood transcriptome is expected to provide complete

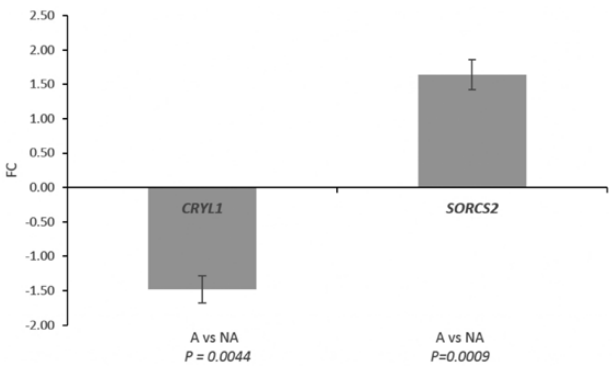

Figure 2. Differences in the *CRYL1* and *SORCS2* expression rate (measure as a fold change, FC) between active (A; n = 35) and nonactive (NA; n = 24) rams using RT-qPCR. Segments indicate the error standard of the estimate.

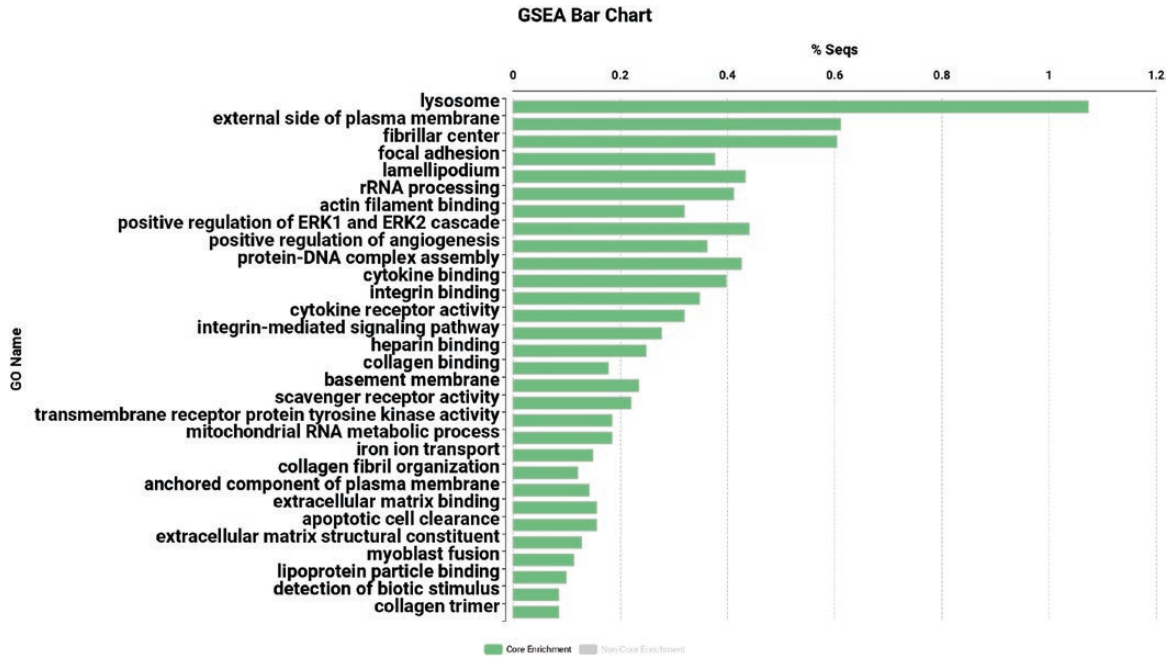

Figure 1. Top 30 enriched Gene Ontology (GO) terms of the blood between active and nonactive rams. The vertical axis represents the top 30 enriched pathway categories (GO) according to their normalized enrichment score (NES) at the significance threshold (false discovery rate; FDR < 0.05), and the horizontal axis represents the percentage calculated as the number of sequences (core or noncore) annotated with each GO. Core enrichment: subset of genes within the gene set that contributes most to the enrichment result. Noncore enrichment: subset of genes within the gene list that have not been enriched. Only core enrichment is shown in the figure.

information on physiological differences about the state of sexual behavior with potential outcomes for applications in animal breeding such as biomarkers.

In this study, we investigated transcriptional differences in peripheral blood samples from rams of Rasa Aragonesa breed with different sexual behavior using RNA-Seq technology. In a previous study, we analyzed the factors influencing the sexual activity phenotype for these 59 rams (Lakhssassi et al., 2023). Testosterone is considered as the predominant androgen for expressing and maintaining libido in rams in numerous studies (Perkins and Roselli, 2007). However, the differences observed in sexual behavior between rams were not due to testosterone in the current study, as no significant effect between testosterone concentration and rams sexual behavior in the pen test was found (Lakhssassi et al., 2023). In that study, statistical analysis showed that rams sexual behavior was influenced by age ($P < 0.01$) and LW ($P < 0.05$), and these results were also confirmed by a tree-based regression model analysis for the sexual activity phenotype (Lakhssassi et al., 2023). For this reason, for the RNA-Seq analysis we choose a population subgroup from the total population with similar values for these significant variables but different sexual activity phenotype.

Unfortunately, as evidenced by read trimming, more than 50% of sample 1 was trimmed (Supplementary Table S1). Best practices for ENCODE2 RNA-Seq experiments indicate that experiments whose purpose is to look for gene expression profiling, require 20 to 25M mappable reads/sample to the genome or known transcriptome (https://genome.ucsc.edu/ENCODE/protocols/dataStandards/ENCODE_RNAseq_Standards_V1.0.pdf). In our case, surviving reads for sample 1 reached 22,708,884. Therefore, we decided to perform the transcriptomic analysis including this sample ($n = 12$) (scenario 1). Despite that, we also wanted to determine whether similar results (DEGs) would be obtained if we had excluded sample 1 from the analyses (scenario 2). Likewise, analyses with the 11 samples (scenario 2) identified also 4 DEGs (FDR < 0.10) where three were upregulated, and one was downregulated (Supplementary Table S2). *CRYL1* and *Acrosin inhibitor 1* genes appears in common between the two analyses in addition to two new upregulated genes appeared in scenario 2, *GPC3* and *OSBP2*. *GPC3* and *oxysterol binding protein 2* (*OSBP2)* genes do not seem to be related to our phenotype according to literature. In this case, the *SORCS2* gene did not reach the statistical significance (*P-adj* = 0.11). Therefore, only common DEGs to the two transcriptomic analyses were chosen for RNA-Seq validation (*CRYL1* and *acrosin inhibitor 1*). Besides, *SORCS2* gene, that was only significant in scenario 1, was also chosen for validation due to its putative involvement in sexual behavior according to literature. It is outstanding that samples used for validation of the RNA-Seq ($n = 59$) were taken for the total population the day of the first individual sexual behavior pen test (sexual behavior phenotype from the first pen test), while the RNA-Seq samples correspond to the animals that were submitted to two pen tests (sexual behavior phenotype from the two-pen test), and were taken the day after the second pen tests. In this sense, population validation samples ($n = 59$) and RNA-seq samples ($n = 12$) were collected 15 days apart. With the second individual sexual behavior pen test performed fifteen days later than the first one in subpopulation of selected rams with similar BCS and age ($n = 21$; BCS > 3 and age ≤ 2.521; Lakhssassi et al., 2023), we confirmed that the animals selected for RNA-seq studies were kept A or NA rams. The genes *acrosin inhibitor 1* (*ENSOARG00020023278*) and *SORCS2* were upregulated (log2FC > 1) in active rams. According to the literature, acrosin is considered to play a critical role in the reproduction process. However, this gene could not be validated in the total ram population ($n = 59$) by RT-qPCR. This gene was characterized by relatively low expression levels, as indicated by lower RNA-Seq read counts, with an average of 12.7 (number of reads aligned to each genomic feature). *SORCS2* which encodes the sortilin-related VPS10 domain containing receptor 2 was reported to be involved in neural plasticity, development and regeneration of neuronal circuits in mammals (Glerup et al., 2014, 2016). Furthermore, the *SORCS2* gene is mainly expressed in central nervous system but also in peripheral tissues (https://www.proteinatlas.org/ENSG00000184985-SORCS2/tissue; last accessed on 14 November 2022), including bone marrow tissue. Glerup et al. (2014) showed that *SORCS2* regulated dopaminergic axon guidance and peripheral sensory neuron apoptosis, respectively, beside interacting with pro-BDNF/p75NTR, and controlling the activity of Trk receptors. Knowing that the brain-derived neurotrophic factor (BDNF) and its receptor, tyrosine receptor kinase B (trkB), related to synaptic plasticity, are implicated in sexual behavior in male rats (Sanna et al., 2019), the gene expression changes in *SORCS2* gene in peripheral blood cells could be a reflection of those changes occurring in the brain and a marker for sexual activity phenotype. On the other hand, the *CRYL1* and *immunoglobulin lambda-1 light chain isoform X47* (*ENSOARG00020025518*) genes were downregulated. The *crystallin*, *lambda 1* (*CRYL1*) belongs to the family of crystallins, many of which function as small heat-shock proteins involved in stress-protection (Wistow, 2012). A genome-wide association study analyzing coping behaviors in humans identified the *CRYL1* gene as associated with the emotional support seeking behavior (Shimanoe et al., 2019). Biological roles for activin have been proposed in a number of reproductive organs including the testis, where it regulates spermatogenesis (Welt et al., 2002). Moreover, *CRYL1* expression levels decreased after administration of activin A in testis in a study of Sertoli cell maturation (Itman et al., 2009). In that study, Activin modulates the number of GnRH receptors and hence, response to GnRH (Gregory and Kaiser, 2004), as well as a number of other genes in LβT2 cells including *inhbb*, *inha*, *gdf9*, and the *17β-HSD* gene (Zhang et al., 2006). Thus, activin has a major role in modulating neuroendocrine reproductive control (Xia and Schneyer, 2009). In our study, this gene was downregulated in A rams, both in RNA-Seq ($n = 12$) and the validation study ($n = 59$). Thus, we can hypothesize that *CRYL1* gene expression levels could be a biomarker of low mating behavior.

DEGs detected in this study were not found in our previous research on hypothalamus, *pars tuberalis* and pineal gland transcriptome using the same animals (Lakhssassi et al., 2023), even if some of them were also expressed in brain (*CRYL1* and *SORCS2*).

Gene enrichment of DEGs was not allowed since we detected fewer DEGs ($n = 4$) in 12 blood samples. To identify overrepresented pathways related to rams' sexual behavior, GSEA was performed including all the expressed genes. Therefore, the lysosome pathway (GO:0005764) comprising

218 genes and considered as a cellular component, was the most enriched (Supplementary Table S2). According to Xie et al. (2006), lysosomes play a crucial role in steroidogenesis. The genes comprised in this pathway might be an indirect or direct effect on steroid hormone biosynthesis. In fact, infertility and sub-infertility due to some lysosome storage disorder in mice was reported, particularly those that present a shortage in lipid catabolism, which interfere in steroidogenesis since steroidal hormones are synthesized from cholesterol (Butler et al., 2002; Xu et al., 2011). In addition, Almeida et al. (2021) described that *SORCS2* deletion leads to altered neuronal lysosome activity in mouse thus leading to many neurodegenerative diseases. It is to be reminded that *SORCS2* was a DEG upregulated in A rams. Therefore, it is tempting to speculate that any disorders of lysosomal hydrolases could damage hormone production and affect consequently sexual behavior and fertility. Another interesting pathway was the positive regulation of ERK1 and ERK2 cascade (GO:0070374) considered as biological process category. Many studies indicated that ERK1 and ERK2 are important modulators of hypothalamic GnRH-mediated regulation of pituitary gonadotropin production and fertility (Liu et al., 2002; Saba-El-Leil et al., 2003; Nekrasova et al., 2005; Bliss et al., 2009). Furthermore, Bliss et al. (2009) reported that ERK signaling is required in females for ovulation and fertility, whereas male reproductive function is unaffected by this signaling deficiency. Additionally, external side of plasma membrane (GO:0009897), fibrillar center (GO:0001650), focal adhesion (GO:0005925), and lamellipodium (GO:0030027) pathways have also been identified, suggesting that some molecules of these pathways might also be involved in rams' sexual behavior.

The DEGs detected in this study may act as factors controlling sexual behavior through the regulation of dopaminergic pathways and BDNF and also modulating activin expression levels. The mechanism and the regulatory functions of these candidate genes found in this work related to sexual behavior is still unclear and should be investigated in future studies.

Conclusion

Blood transcriptome investigation offers an accessible and less invasive alternative to study behavioral change in rams. Few genes differentially expressed were identified throughout blood transcriptome comparison between active and non-active rams. These results will provide new clues for understanding the molecular regulation of sexual behavior in rams. Nonetheless, it is recommended to study more in depth how the expression variability of these genes affects the sexual behavior in rams. Extending this approach to other breeds and a much larger data set will help to get a detailed picture on physiological change in sexual behavior between rams.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Acknowledgments

The authors would like to thank the technical staff of CITA for the animal caring during the experiment.

Funding

This study was partially financed by the Ministry of Education and Science of Spain research project INIA-RTA2015-00090, the Research Group Funds of the Aragón Government (Ref. A14_20R), and FEDER funds. K. Lakhssassi is supported by a doctoral (BES-2017-080154) grant.

Data Availability

The datasets analyzed during the current study are available in NCBI's Gene Expression Omnibus repository and are accessible through GEO Series accession number GSE218667 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE218667).

Conflict of Interest Statement

The authors declare that they have no competing interests.

Literature Cited

Alexander, B. M., N. E. Cockett, D. J. Burton, T. L. Hadfield, and G. E. Moss. 2012. Reproductive performance of rams in three producer range flocks: evidence of poor sexual behavior in the field. *Small Rumin. Res*. 107:117–120. doi:10.1016/J.SMALLRUMRES.2012.04.003.

Almeida, S., A. M. Miranda, A. E. Tóth, M. S. Nielsen, and T. G. Oliveira. 2021. SorCS2 deletion leads to altered neuronal lysosome activity. *BioRxiv* doi:10.1101/2021.04.08.439000.

Anders, S., P. T. Pyl, and W. Huber. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 31:166–169. doi:10.1093/bioinformatics/btu638.

Andersen, C. L., J. L. Jensen, and T. F. Ørntoft. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*. 64:5245–5250. doi:10.1158/0008-5472.CAN-04-0496.

Andrews, S. 2015. FastQC. A quality control tool for high throughput sequence data. Babraham Inst. 1:1. https://www.scienceopen.com/document?vid=95c0b572-b20b-45da-b380-4227fc75a153

Barrett, T., S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, et al. 2013. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 41:D991–D995. doi:10.1093/nar/gks1193.

Bliss, S. P., A. Miller, A. M. Navratil, J. Xie, S. P. McDonough, P. J. Fisher, G. E. Landreth, and M. S. Roberson. 2009. ERK signaling in the pituitary is required for female but not male fertility. *Mol. Endocrinol*. 23:1092–1101. doi:10.1210/me.2009-0030.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–2120. doi:10.1093/bioinformatics/btu170.

Butler, A., X. He, R. E. Gordon, S. Gatt, and E. H. Schuchman. 2002. Reproductive pathology and sperm physiology in acid sphingomyelinase-deficient mice. *Am. J. Pathol*. 161:1061–1075. doi:10.1016/S0002-9440(10)64267-8.

Chemineau, P., J. Pelletier, Y. Guerin, G. Colas, J. P. Ravault, G. Toure, G. Almeida, J. Thimonier, and R. Ortavant. 1988. Photoperiodic and melatonin treatments for the control of seasonal reproduction in sheep and goats. *Reprod. Nutr. Dev*. 28:409–422. doi:10.1051/rnd:19880307.

Crow, M., N. Lim, S. Ballouz, P. Pavlidis, and J. Gillis. 2019. Predictability of human differential gene expression. *Proc. Natl. Acad. Sci. USA*. 116:6491–6500. doi:10.1073/pnas.1802973116.

Delgadillo, J. A., J. Vielma, J. A. Flores, F. G. Véliz, G. Duarte, and H. Hernández. 2008. La calidad del estímulo emitido por el macho

determina la respuesta de las cabras sometidas al efecto macho the stimulus quality provided by the buck determines the response of the female goats submitted to the male effect. *Trop. Subtrop. Agroecosystems*. 12:139–143. http://www.redalyc.org/articulo.oa?id=93911227004

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 29:15–21. doi:10.1093/bioinformatics/bts635.

Fabre-Nys, C., K. M. Kendrick, and R. J. Scaramuzzi. 2015. The "ram effect": new insights into neural modulation of the gonadotropic axis by male odors and socio-sexual interactions. *Front. Neurosci.* 9:111. doi:10.3389/fnins.2015.00111.

Folch, J., and M. Roca Bernaus. 1982. Caracteristicas sexuales del morueco de raza "Aragonesa" durante el primer año de vida. Influencia de la epoca de nacimiento. *Ann. INIA*. 16:99–112.

Glerup, S., U. Bolcho, S. Mlgaard, S. Bggild, C. B. Vaegter, A. H. Smith, J. L. Nieto-Gonzalez, P. L. Ovesen, L. F. Pedersen, A. N. Fjorback, et al. 2016. SorCS2 is required for BDNF-dependent plasticity in the hippocampus. *Mol. Psychiatry*. 21:1740–1751. doi:10.1038/mp.2016.108.

Glerup, S., D. Olsen, C. B. Vaegter, C. Gustafsen, S. S. Sjoegaard, G. Hermey, M. Kjolby, S. Molgaard, M. Ulrichsen, S. Boggild, et al. 2014. SorCS2 regulates dopaminergic wiring and is processed into an apoptotic two-chain receptor in peripheral glia. *Neuron*. 82:1074–1087. doi:10.1016/j.neuron.2014.04.022.

Gregory, S. J., and U. B. Kaiser. 2004. Regulation of gonadotropins by inhibin and activin. *Semin. Reprod. Med*. 22:253–267. doi:10.1055/s-2004-831901.

Hazard, D. 2010. Seasonality of reproduction in sheep: a) Experimental project for QTL detection influencing seasonality; b) preliminary results of an experimental genetic selection for seasonality. Internship report. AgroParisTech, Paris, France. https://hal.inrae.fr/hal-02819759

Itman, C., C. Small, M. Griswold, A. K. Nagaraja, M. M. Matzuk, C. W. Brown, D. A. Jans, and K. L. Loveland. 2009. Developmentally regulated SMAD2 and SMAD3 utilization directs activin signaling outcomes. *Dev. Dyn*. 238:1688–1700. doi:10.1002/dvdy.21995.

Lakhssassi, K., M. P. Sarto, B. Marín, B. Lahoz, J. Folch, J. L. Alabart, M. Serrano, and J. H. Calvo. 2023. Exploring differentially expressed genes in hypothalamic, pars tuberalis and pineal gland transcriptomes in different sexual behavior phenotypes in rams using RNA-Seq. *J. Anim. Sci*. 101:skac365. doi:10.1093/jas/skac365.

Lindsay, D. R. 1996. Environment and reproductive behaviour. *Anim. Reprod. Sci*. 42:1–12. doi:10.1016/0378-4320(96)01527-8.

Liu, F., D. A. Austin, P. L. Mellon, J. M. Olefsky, and N. J. G. Webster. 2002. GnRH activates ERK1/2 leading to the induction of c-fos and LHbeta protein expression in LbetaT2 cells. *Mol. Endocrinol*. 16:419–434. doi:10.1210/MEND.16.3.0791.

Martin, G. B., and H. Kadokawa. 2006. "Clean, green and ethical" animal production. Case study: reproductive efficiency in small ruminants. *J. Reprod. Dev*. 52:145–152. doi:10.1262/jrd.17086-2

Nekrasova, T., C. Shive, Y. Gao, K. Kawamura, R. Guardia, G. Landreth, and T. G. Forsthuber. 2005. ERK1-deficient mice show normal T cell effector function and are highly susceptible to experimental autoimmune encephalomyelitis. *J. Immunol*. 175:2374–2380. doi:10.4049/jimmunol.175.4.2374.

Perkins, A., and C. E. Roselli. 2007. The ram as a model for behavioral neuroendocrinology. *Horm. Behav*. 52:70–77. doi:10.1016/j.yhbeh.2007.03.016.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26:139–140. doi:10.1093/bioinformatics/btp616.

Russel, A. J. F., J. M. Doney, and R. G. Gunn. 1969. Subjective assessment of body fat in live sheep. *J. Agric. Sci*. 72:451–454. doi:10.1017/s0021859600024874.

Saba-El-Leil, M. K., F. D. J. Vella, B. Vernay, L. Voisin, L. Chen, N. Labrecque, S. L. Ang, and S. Meloche. 2003. An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. *EMBO Rep*. 4:964–968. doi:10.1038/sj.embor.embor939.

Sanna, F., L. Poddighe, M. P. Serra, M. Boi, J. Bratzu, S. Francesco, M. G. Corda, O. Giorgi, M. R. Melis, A. Argiolas, et al. 2019. c-Fos, ΔFosB, BDNF, trkB and arc expression in the limbic system of male roman high- and low-avoidance rats that show differences in sexual behavior: effect of sexual activity. *Neuroscience*. 396:1–23. doi:10.1016/j.neuroscience.2018.11.002.

Shimanoe, C., T. Hachiya, M. Hara, Y. Nishida, K. Tanaka, Y. Sutoh, A. Shimizu, A. Hishida, S. Kawai, R. Okada, et al. 2019. A genome-wide association study of coping behaviors suggests FBXO45 is associated with emotional expression. *Genes Brain Behav*. 18:1–10. doi:10.1111/gbb.12481.

Simitzis, P. E., S. G. Deligeorgis, and J. A. Bizelis. 2006. Effect of breed and age on sexual behaviour of rams. *Theriogenology*. 65:1480–1491. doi:10.1016/j.theriogenology.2005.09.001.

Steibel, J. P., R. Poletto, P. M. Coussens, and G. J. M. Rosa. 2009. A powerful and flexible linear mixed model framework for the analysis of relative quantification RT-PCR data. *Genomics*. 94:146–152. doi:10.1016/j.ygeno.2009.04.008.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*. 102:15545–15550. doi:10.1073/PNAS.0506580102.

Welt, C., Y. Sidis, H. Keutmann, and A. Schneyer. 2002. Activins, inhibins, and follistatins: from endocrinology to signaling. A paradigm for the new millennium. *Exp. Biol. Med*. 227:724–752. doi:10.1177/153537020222700905.

Wistow, G. 2012. The human crystallin gene families. *Hum. Genomics*. 6:1–10. doi:10.1186/1479-7364-6-26.

Xia, Y., and A. L. Schneyer. 2009. The biology of activin: recent advances in structure, regulation and function. *J. Endocrinol*. 202:1–12. doi:10.1677/JOE-08-0549.

Xie, C., J. A. Richardson, S. D. Turley, and J. M. Dietschy. 2006. Cholesterol substrate pools and steroid hormone levels are normal in the face of mutational inactivation of NPC1 protein. *J. Lipid Res*. 47:953–963. doi:10.1194/jlr.M500534-JLR200.

Xu, H., K. Kongmanas, S. Kadunganattil, C. E. Smith, T. Rupar, N. Goto-Inoue, L. Hermo, K. F. Faull, and N. Tanphaichitr. 2011. Arylsulfatase A deficiency causes seminolipid accumulation and a lysosomal storage disorder in Sertoli cells. *J. Lipid Res*. 52:2187–2197. doi:10.1194/jlr.M019661.

Zhang, H., J. S. Bailey, D. Coss, B. Lin, R. Tsutsumi, M. A. Lawson, P. L. Mellon, and N. J. G. Webster. 2006. Activin modulates the transcriptional response of LβT2 cells to gonadotropin-releasing hormone and alters cellular proliferation. *Mol. Endocrinol*. 20:2909–2930. doi:10.1210/me.2006-0109.