

Supplemental Information

Detailed Materials and Methods

Evaluation of spermatozoa motility using a computer-assisted sperm analysis system

Spermatozoa motility was assessed using a computer-assisted sperm assay (CASA) method according to World Health Organization guidelines [57,58]. After euthanasia, spermatozoa were collected from the cauda epididymis of mice and suspended in DMEM/F12 medium with 10% FBS and incubated at 37.5 °C for 30 min; samples were then placed in a pre-warmed counting chamber [57,58]. A microscopic sperm class analyzer (CASA; MICROPTIC S.L. Barcelona, Spain) was used in this investigation. It was equipped with a 20-fold objective, a camera adaptor (Eclipse E200, Nikon, Japan), and a camera (acA780-75gc, Basler, Germany). The classification of sperm motility was as follows: Grade A linear velocity $>22 \mu\text{m s}^{-1}$; Grade B $<22 \mu\text{m s}^{-1}$ and curvilinear velocity $>5 \mu\text{m s}^{-1}$; Grade C curvilinear velocity $<5 \mu\text{m s}^{-1}$; and Grade D = immotile spermatozoa. The spermatozoa motility data represented only Grade A + Grade B since only these two grades are considered to be functional.

Morphological observations of spermatozoa

The extracted murine caudal epididymides were placed in RPMI medium and finely chopped; subsequently, Eosin Y (1%) was added for staining as described previously [57,58]. Spermatozoon abnormalities were then viewed using an optical microscope and were classified into head or tail morphological abnormalities: two heads, two tails, blunt hooks, and short tails. The examinations were repeated three times, and 500 spermatozoa per animal were scored.

Assessment of acrosome integrity

After harvest, mouse spermatozoa were incubated at 37.5 °C for 30 min, after

which a drop of sperm suspension was uniformly smeared on a clean glass slide. Smeared slides were air dried and incubated in methanol for 2 min for fixation. After fixation, the slides were washed with PBS three times. Assessment of an intact acrosome was accomplished by staining the sperm with 0.025% Coomassie brilliant blue G-250 in 40% methanol for 20 min at room temperature (RT). The slides were then washed three times with PBS and mounted with 50% glycerol in PBS. Acrosomal integrity was determined by an intense staining on the anterior region of the sperm head under bright-field microscopy (AH3-RFCA, Olympus, Tokyo, Japan) and scored accordingly [57,58].

Single cell library preparation, sequencing, and data analysis [single cell RNA-sequencing (scRNA-seq)]

Single cell library preparation and sequencing. Single cell libraries were constructed with the 10x Genomics Chromium Single Cell 3' Library and Gel Bead Kit v2 (10× Genomics Inc., Pleasanton, CA, USA, 120237) following the manufacturer's instructions. Single cell sample collection followed the methods reported by Wang et al. [1]. Briefly, mouse testes were collected, seminiferous tubules were cut into small pieces, and then washed with PBS three times to remove the sperm. Subsequently, the tissue was digested using TrypLE express (Invitrogen) for 15 min at 37 °C (in culture, in an incubator). The single cells were collected by filtration using a 40 µm filter. Cells were then washed twice with PBS solution supplemented with 0.04% bovine serum albumin (BSA; Sigma, St. Louis, MO, USA, A1933). Trypan blue staining with a hemocytometer (Bio-Rad, Hercules, CA, USA, TC20) was used to detect cell viability. Six individual murine testicular cells were collected and combined together; subsequently, a concentration of 1000 cells/µl was used for loading onto the single cell chip (one/group). A Chromium 10x Single Cell System (10×Genomics, Pleasanton, CA,

USA) was used to form the Gel-Bead in Emulsions (GEMs). Cells were then barcoded and a cDNA library was constructed. The sequencing protocol used an Illumina HiSeq X Ten sequencer (Illumina, San Diego, CA, USA) with pair end 150 bp (PE150) reads.

Single sample analysis and aggregation. CellRanger v2.2.0 software (<https://www.10xgenomics.com/>) was used to process the datasets with the “--force-cells = 5000” argument. The 10x Genomics pre-built mouse genome for mm10-3.0.0 (<https://support.10xgenomics.com/single-cell-geneexpression/software/downloads/latest>) was referenced. After the CellRanger analysis, the gene-barcode matrices were processed using the Seurat single cell RNA-seq analysis R package in Rstudio (v3.0) [59]. Cells with minimal genes <200 and genes expressed in <3 cells were removed to obtain high-quality datasets for downstream analysis. After normalization, the three datasets (from the three treatment groups) were merged together using the Seurat “RunMultiCCA” function. The characterized cell clusters were reviewed by the Seurat “RunTSNE” function based on the t-distributed Stochastic Neighbor Embedding (tSNE) algorithm with default settings. The cell clusters were calculated using the “FindClusters” function. The Seurat “FindAllMarkers” function was applied to find cell cluster markers.

Subclustering and gene ontology enrichment analysis. After characterization of all cell clusters in murine small intestine samples, cells were further clustered based on their cell identity. To obtain the same type of cells for downstream analysis, the “SubsetData” function was applied. After clustering, cluster-specific marker genes were identified using the “FindAllMarkers” function. The marker genes were used for enrichment analysis in Metascape (<http://metascape.org>).

Single-cell pseudo-time trajectory analysis. Monocle 2 (v2.8.0) was used to determine the single-cell pseudo-time trajectory (<http://cole-trapnell->

lab.github.io/monocle-release/tutorials/) [60,61]. The Monocle object was formed by implementing the Monocle “newCellDataSet” function from the Seurat object with a “lowerDetectionLimit” = 0.5. The variable genes for ordering were obtained using Seurat. Dimensionality was constructed using the DDRTree method with regression of the number of UMIs. Root states were collected following their Seurat cell identity information and branch-specific gene expression was calculated using the Monocle implemented BEAM function. The branched heatmap was further constructed using the “plot_genes_branched_heatmap” function.

Single cell regulatory network analysis. To identify gene regulatory networks active during the development of the small intestine, we performed regulatory network inference and clustering using SCENIC (<https://github.com/aertslab/SCENIC>), a modified method for inferring with gene regulatory networks from single-cell RNA-seq data [31]. The single-cell RNA-seq expression matrix, where each column represents a cell ID and each row represents a gene, was used for analysis. Then, the “geneFiltering” function was used to remove genes with UMI counts across all samples <100 and was expressed in less than 1% of cells. GENIE3 was then used to infer the co-expression matrix containing potential regulators. To identify potential direct-binding targets, RcisTarget was used based on DNA-motif analysis and we used databases (mm 10) that scored the motifs in the gene promoters (up to 500 bp upstream the TSS), and in the 10 kb around the TSS (+/-10 kb). The AUCell algorithm was used to calculate regulon activity in each cell and to convert the network activity into ON/OFF (binary activity matrix) with default settings.

RNA-seq analysis for ex vivo testes samples

Transcriptomics were analyzed as described in our earlier article [62]. Briefly, total RNA was isolated using TRIzol Reagent (Invitrogen) and purified using a Pure-

Link1 RNA Mini Kit (Cat: 12183018A; Life Technologies) following the manufacturer's protocol. The total RNA samples were first treated with DNase I to degrade any possible DNA contamination. Then the mRNA was enriched using oligo(dT) magnetic beads. When mixed with the fragmentation buffer, the mRNA was broken into short fragments (about 200 bp), after which the first strand of cDNA was synthesized using a random hexamer-primer. Buffer, dNTPs, RNase H, and DNA polymerase I were added to synthesize the second strand. The double strand cDNA was purified with magnetic beads. Subsequently, 3'-end single nucleotide A (adenine) addition was performed. Finally, sequencing adaptors were ligated to the fragments. The fragments were enriched by PCR amplification. During the QC step, Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System were used to qualify and quantify the sample library. The library products were ready for sequencing via Illumina HiSeq™ 2500. The reads were mapped to reference genes using SOAPaligner (v. 2.20) with a maximum of two nucleotide mismatches allowed at the parameters of “-m 0 -x 1000 -s 40 -l 35 -v 3 -r 2”. The read number of each gene was transformed into RPKM (Reads Per Kilo bases per Million reads), and then differentially expressed genes were identified using the DEGseq package and the MARS (MA-plot-based method with the Random Sampling model) method. The threshold was set as $FDR \leq 0.001$ and an absolute value of \log_2 ratio ≥ 1 to judge the significance of differences in gene expression.

Sequencing of microbiota from small intestine digesta samples and data analysis [63]

DNA Extraction. Total genomic DNA of small intestine digesta was isolated using an E.Z.N.A.R Stool DNA Kit (Omega Bio-tek Inc., USA) following the manufacturer's instructions. DNA quantity and quality were analyzed using a NanoDrop 2000 (Thermo Scientific) and 1% agarose gel. Ten samples/groups were determined.

Library preparation and sequencing. The V3–V4 region of the 16S rRNA gene was amplified using the primers MPRK341F (50- ACTCCTACGGGAGGCAGCAG - 30) and MPRK806R: (50- GGACTACHVGGGTWTCTAAT -30) with Barcode. The PCR reactions (total 30 μ L) included 15 μ L of PhusionR High-Fidelity PCR Master Mix (New England Biolabs), 0.2 mM of primers, and 10 ng of DNA. The thermal cycle was carried out with an initial denaturation at 98 °C, followed by 30 cycles of 98 °C for 10 s, 50 °C for 30 s, 72 °C for 30 s, and a final extension at 72 °C for 5 min. PCR products were purified using a GeneJET Gel Extraction Kit (Thermo Scientific). The sequencing libraries were constructed with a NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer’s instructions and index codes were added. Then, the library was sequenced on a Illumina HiSeq 2500 platform and 300 bp paired-end reads were generated at the Novo gene. The paired-end reads were merged using FLASH (v. 1.2.71). The quality of the tags was controlled in QIIME (v. 1.7.02), meanwhile all chimeras were removed. The “Core Set” of the Greengenes database³ was used for classification, and sequences with >97% similarity were assigned to the same operational taxonomic units (OTUs).

Analysis of sequencing data Operational taxonomic unit abundance information was normalized using a standard of sequence number corresponding to the sample with the least sequences. The alpha diversity index was calculated with QIIME (v. 1.7.0). The Unifrac distance was obtained using QIIME (v. 1.7.0), and PCoA (principal coordinate analysis) was performed using R software (v. 2.15.3). The linear discriminate analysis effect size (LEfSe) was performed to determine differences in abundance; the threshold LDA score was 4.0. GraphPad Prism7 software was used to produce the graphs.

Plasma and testis metabolite measurements by LC-MS/MS

Plasma samples were collected and immediately stored at -80 °C. Before LC-MS/MS analysis, the samples were thawed on ice and processed to remove proteins. Testis samples were collected and the same amount of tissue from each mouse testis was used to isolate the metabolites using CH₃OH: H₂O (V:V) = 4:1. Then samples were analyzed by ACQUITY UPLC and AB Sciex Triple TOF 5600 (LC/MS) as reported previously [39]. Fifteen samples/group were analyzed for plasma or testis tissues.

HPLC used an ACQUITY UPLC BEH C18 column (100 mm × 2.1 mm, 1.7 μm), solvent A [aqueous solution with 0.1% (v/v) formic acid], and solvent B [acetonitrile with 0.1% (v/v) formic acid] with a gradient program:

Time	A%	B%
0	95	5
2	80	20
4	75	25
9	40	60
17	0	100
19	0	100
19.1	95	5
20.1	95	5

The flow rate was 0.4 mL/min and the injection volume was 5 μL.

The mass spectrometry program with ESI was:

Parameters	Positive ion	Positive ion
Nebulizer Gas (GS1, PSI)	40	40
Auxiliary Gas (GS2, PSI)	40	40
Curtain Gas (CUR, PSI)	35	35
Ion Source Temperature (°C)	550	550
Ion Spray Voltage (V)	5500	4500
Declustering Potential (DP, V)	100	-100
Mass Scan Range (TOF MS scan)	70–1000	70–1000
Collision Energy (TOF MS scan, eV)	10	-10
Mass Scan Range (Product Ion scan)	50–1000	50–1000
Collision Energy (Product Ion scan, eV)	30	30
Interface Heater Temperature (°C)	550	600

Progenesis QI v. 2.3 (Nonlinear Dynamics, Newcastle, UK) was used to normalize the peaks. Then the Human Metabolome Database (HMDB), Lipidmaps (v. 2.3), and METLIN software were used to qualify the data. Moreover, the data were processed with SIMCA software (v. 14.0, Umetrics, Umeå, Sweden) following pathway enrichment analysis using the KEGG database (<http://www.genome.jp/KEGG/pathway.html>).

Histopathological analysis

Testicular tissues were fixed in 10% neutral buffered formalin, paraffin embedded, cut into 5 µm sections and subsequently stained with hematoxylin and eosin for histopathological analysis.

Immunofluorescence staining (IHF)

The procedure for immunofluorescence staining is reported in our recent publications [58,64]. Table S8 lists the primary antibodies. Briefly, testis sections (5 µm) were prepared and subjected to antigen retrieval. Sections were then blocked with normal goat serum in PBS, followed by incubation (1:150 in PBS-1% BSA) with primary antibodies at 4 °C overnight. After a brief wash, sections were incubated with goat anti-rabbit or donkey anti-goat secondary Abs (1:100 in PBS; Beyotime Institute of Biotechnology, Shanghai, P.R. China) at RT for 30 min and then counterstained with 4',6-diamidino-2-phenylindole (DAPI). The stained sections were visualized using a Nikon Eclipse TE2000-U fluorescence microscope (Nikon, Inc., Melville, NY), and the captured fluorescence images were analyzed using MetaMorph software. Six samples/group were determined.

Western blotting

Western blotting analysis followed the procedure reported in our previous publications [58,64]. Briefly, testis tissue samples were lysed in RIPA buffer containing

a protease inhibitor cocktail from Sangong Biotech, Ltd. (Shanghai, China). Protein concentration was determined by BCA kit (Beyotime Institute of Biotechnology). Information for primary antibodies is given in Table S8. Secondary donkey anti-goat Ab (Cat no.: A0181) was purchased from Beyotime Institute of Biotechnology, and goat anti-rabbit (Cat no.: A24531) Abs were purchased from Novex[®] by Life Technologies (USA). Fifty micrograms of total protein per sample were loaded onto 10% SDS polyacrylamide electrophoresis gels. The gels were transferred to a polyvinylidene fluoride (PVDF) membrane at 300 mA for 2.5 h, at 4 °C. Then, the membranes were blocked with 5% BSA for 1 h at RT, followed by three washes with 0.1% Tween-20 in TBS (TBST). The membranes were incubated with primary Abs diluted with 1:500 in TBST with 1% BSA overnight at 4 °C. After three washes with TBST, the blots were incubated with the HRP-labelled secondary goat anti-rabbit or donkey anti-goat Abs respectively for 1 h at RT. After three washes, the blots were imaged. The experiment was performed with six individual samples/group.

Statistical analysis

Data were analyzed using SPSS statistical software (IBM Co., NY, USA) with one-way analysis of variance (ANOVA) followed by LSD multiple comparison tests. All groups were compared with each other for every parameter. The data were shown as the mean \pm SEM. Statistical significance was based on $p < 0.05$.

Supplementary Materials

Figure S1. Additional data for single cell RNA seq (scRNA seq) analysis of all cells in four group samples.

Figure S2. Additional data for sub-clustering of SPGs, SPCs, and STs in single cell RNA seq (scRNA seq) analysis.

Figure S3. Heatmap of gene expression for SPG, SPC, ST, and LC/SC clusters in each group sample.

Figure S4. Enrichment analysis data for each group sampled individually.

Figure S5. GO enrichment analysis for down-regulated genes in each of the four comparisons (*ex vivo*).

Table S1. Differentially expressed genes in scRNA seq analysis for all the samples.

Table S2. Marker genes in scRNA seq analysis for all the samples.

Table S3. Motif Enrichment data for transcriptional factor analysis.

Table S4. Regulon target information in transcriptional factor analysis.

Table S5. Testis metabolites.

Table S6. Plasma metabolites.

Table S7. Correlation of testis metabolites and plasma metabolites.

Table S8. Relative amounts of microbiota in small intestine samples.

Table S9. Information for primary antibodies.

Supplementary Table 9. Primary antibody information

Gene symbol	Name	Cat. #	Predicted size	Source (Animal)	Company
DDX4 (VASA)	DEAD (Asp Glu Ala Asp) box polypeptide	ab13840	76kDa	Rabbit (polyclonal)	Abcam
DAZL	DAZ like autosomal	ab34139	33kDa	Rabbit (polyclonal)	Abcam
SCP3/SYCP3	Synaptonemal complex protein 3	NB300-232	28kDa	Rabbit (polyclonal)	Novus Biologicals
SOX9	SRY (sex-determining region Y)-box 9 protein	AB5535	65kDa	Rabbit (polyclonal)	Merck Millipore
TNP1(TR1)	Transition protein-1	ab73135		Rabbit (polyclonal)	Abcam
Jun D	Transcription factor jun-D	bs-1393R	38kDa	Rabbit (polyclonal)	Beijing Biosynthesis Biotechnology CO.
Wt1	Wilms tumor protein	bs-6983R	55kDa	Rabbit (polyclonal)	Beijing Biosynthesis Biotechnology CO.
PIWIL1	Piwi like protein 1	ab94917	99kDa	Rabbit (polyclonal)	Abcam
ZFP37	Zinc finger protein 37 homolog	bs-19125R	71kDa	Mouse(monoclonal)	Beijing Biosynthesis Biotechnology CO.
P-ERK	Phospho-Erk1 (Thr202 + Tyr204)	bs-1645R	43kDa	Rabbit (polyclonal)	Beijing Biosynthesis Biotechnology CO.
GPX1	Glutathione peroxidase 1	bs-3882R	22kDa	Rabbit (polyclonal)	Beijing Biosynthesis Biotechnology CO.
Caspase 8	Caspase 8	bs-0052R	12/55kDa	Rabbit (polyclonal)	Beijing Biosynthesis Biotechnology CO.
ACTB	Anti-ACTB rabbit polyclonal antibody	D110001-0200	42kDa	Rabbit (polyclonal)	Sangon Biotech.
actin	Actin	Ab3280	42kDa	Rabbit (polyclonal)	Abcam

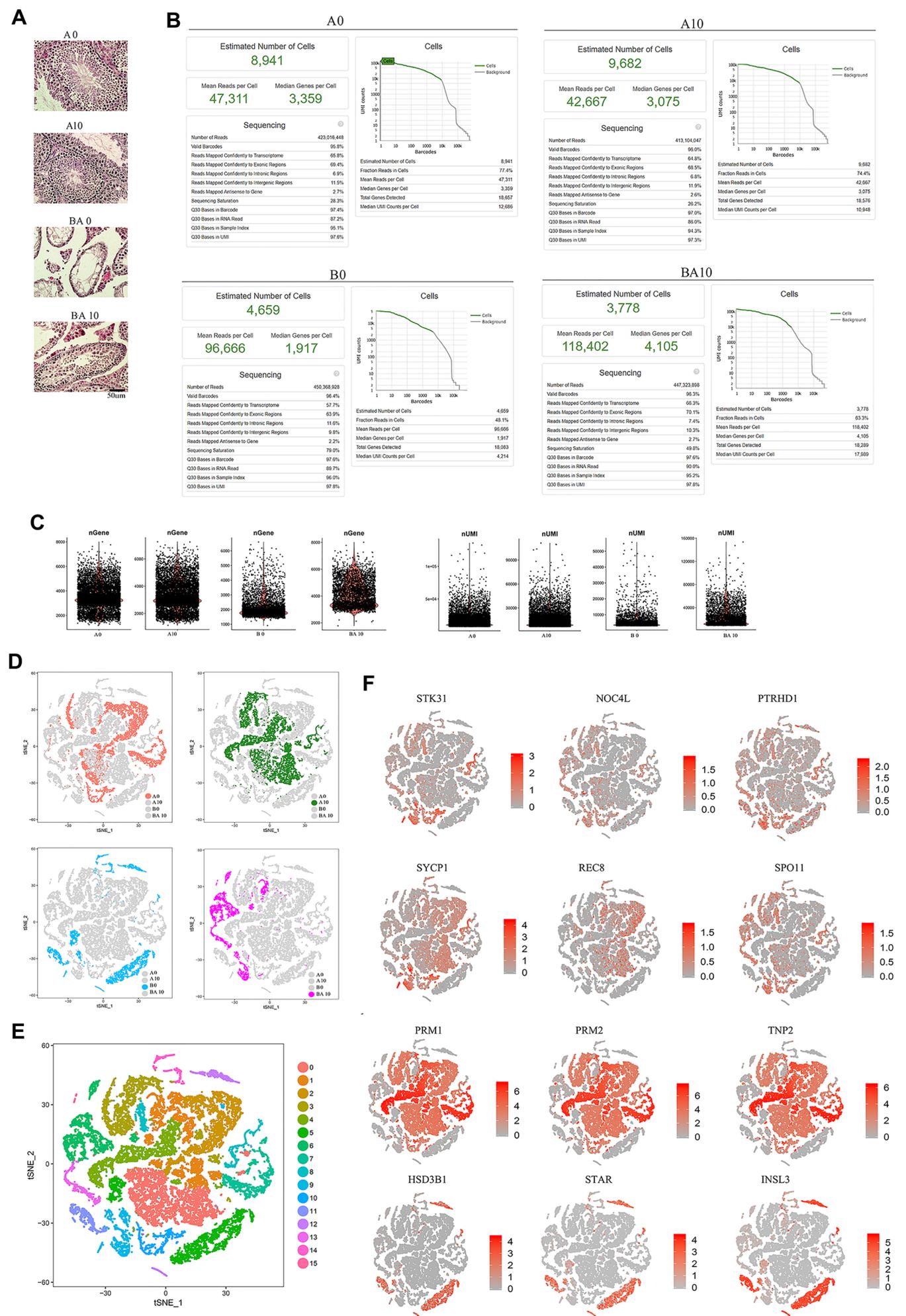


Figure S1. Additional data for single cell RNA seq (scRNA seq) analysis of all cells in four group samples.

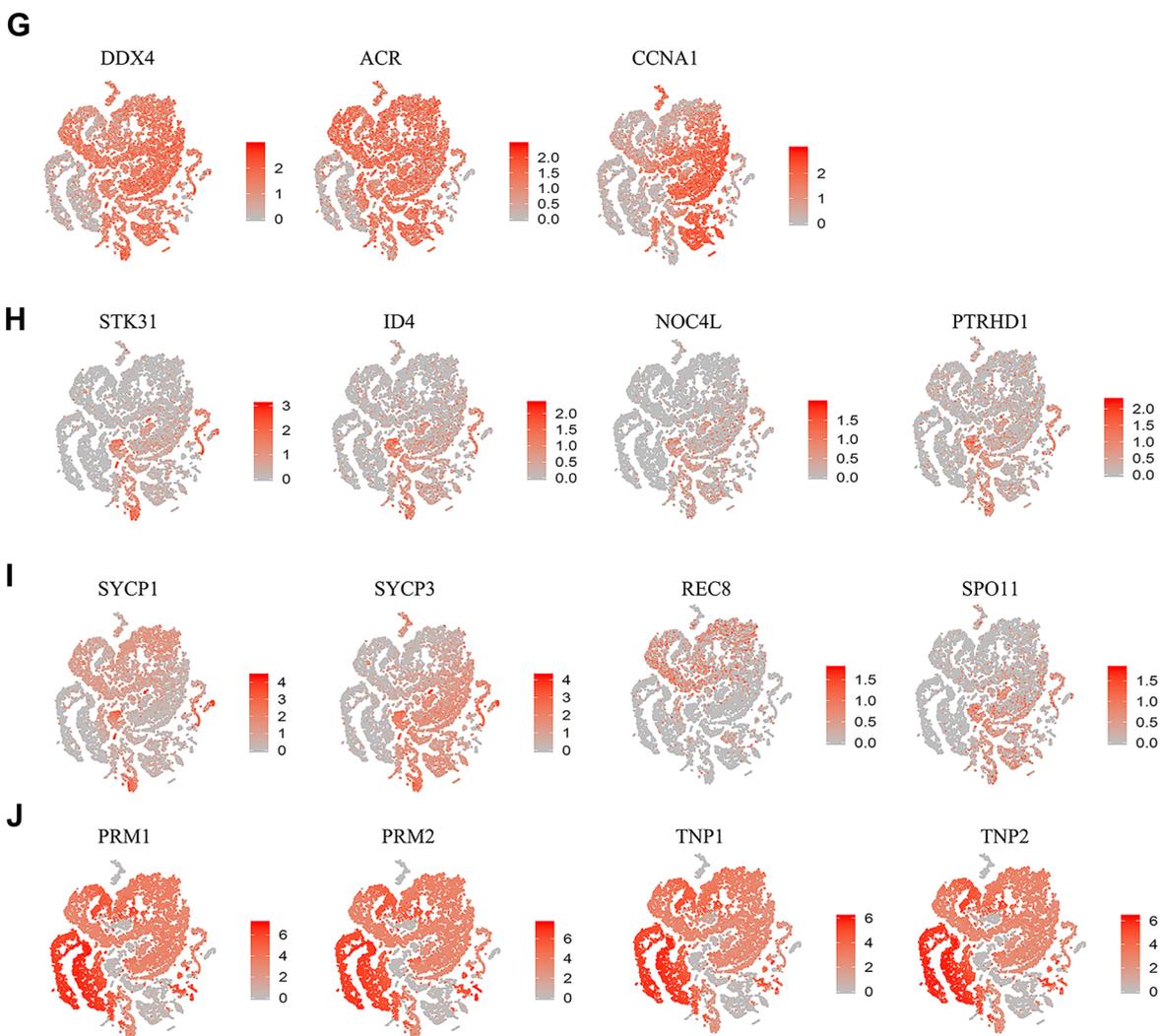
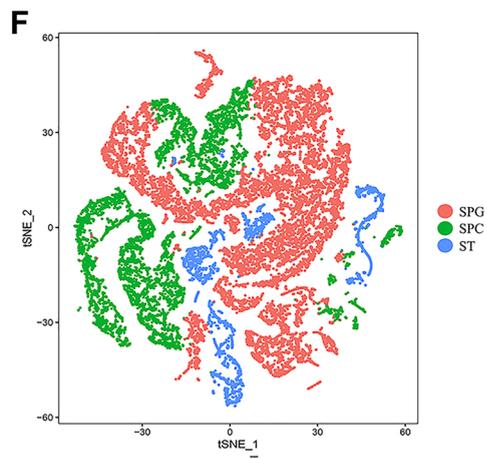
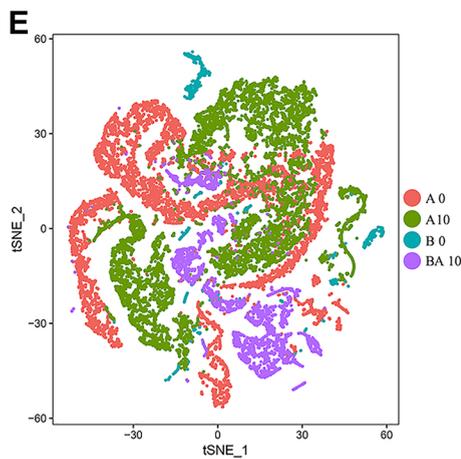
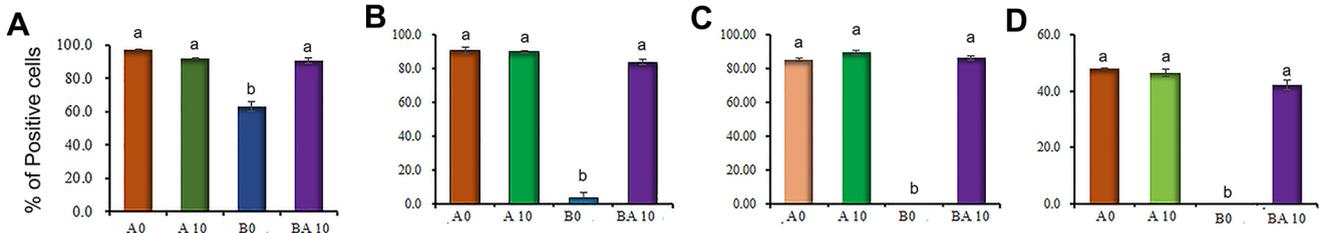


Figure S2. Additional data for sub-clustering of SPGs, SPCs, and STs in single cell RNA seq (scRNA seq) analysis.

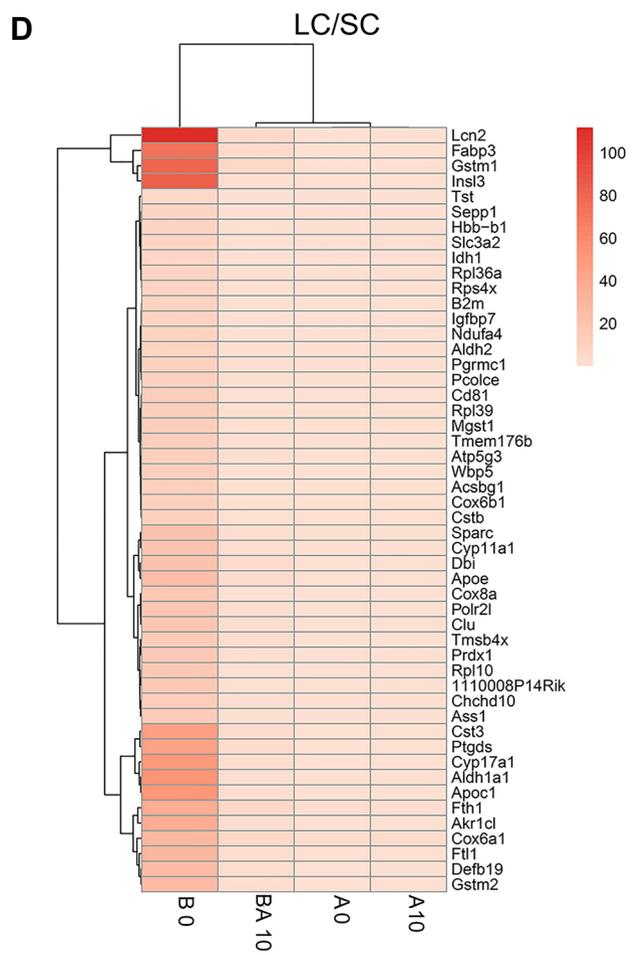
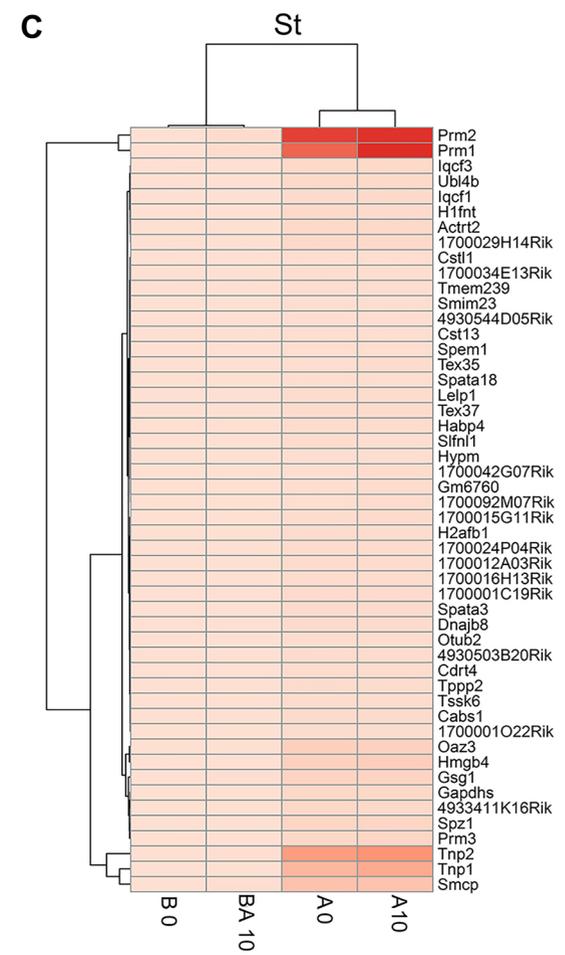
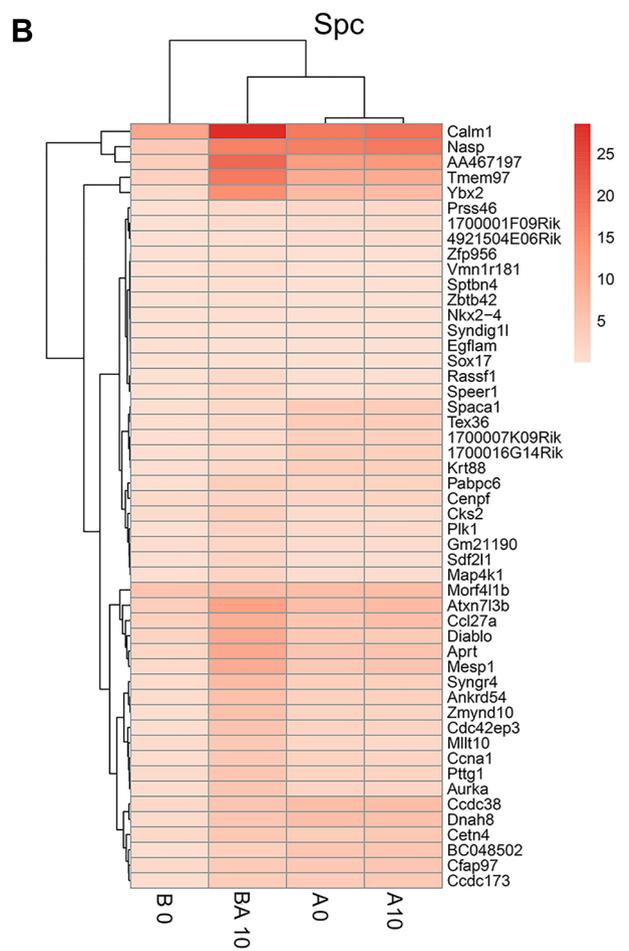
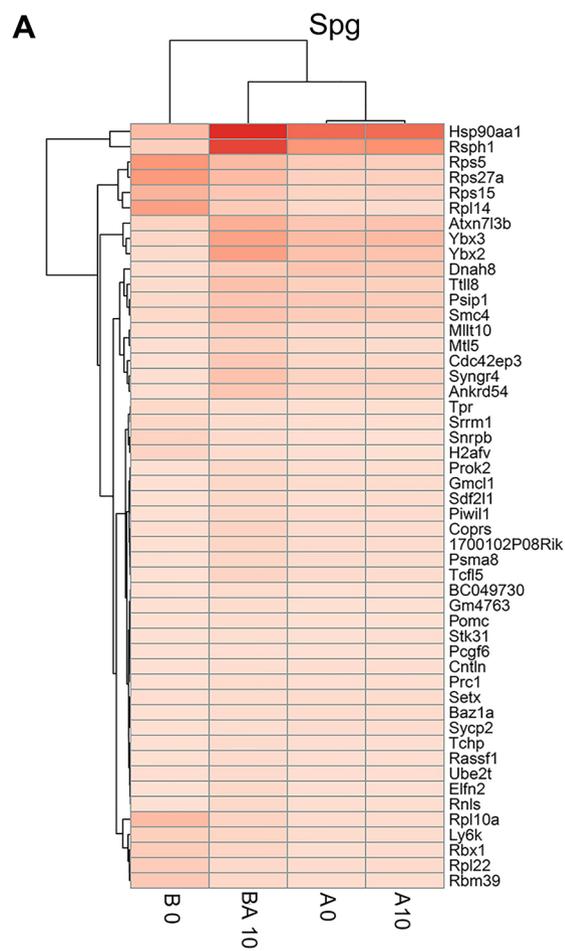


Figure S3. Heatmap of gene expression for SPG, SPC, ST, and LC/SC clusters in each group sample.

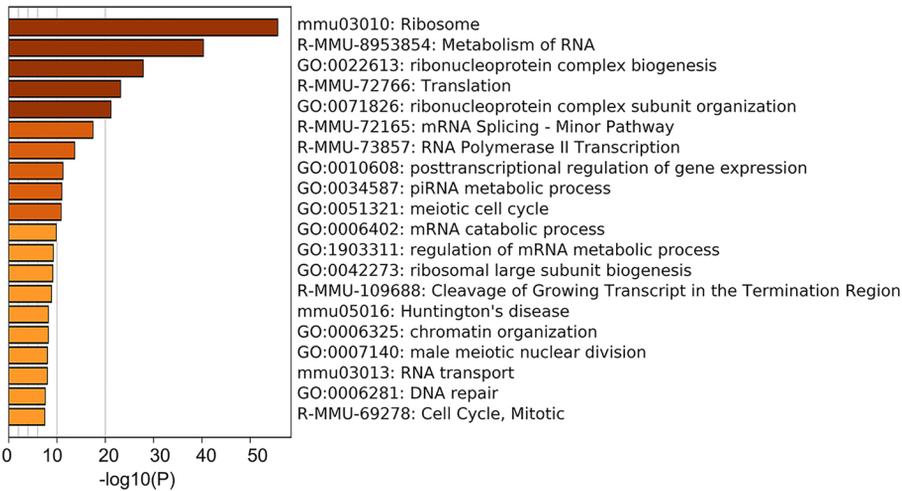
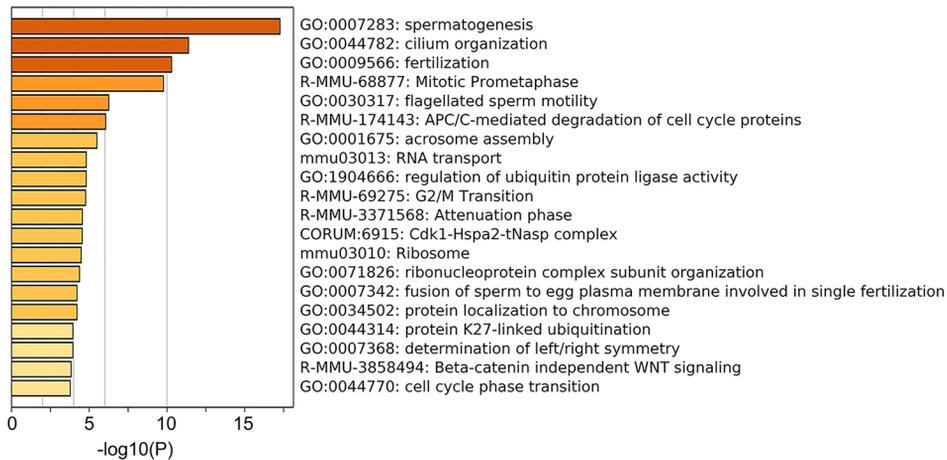
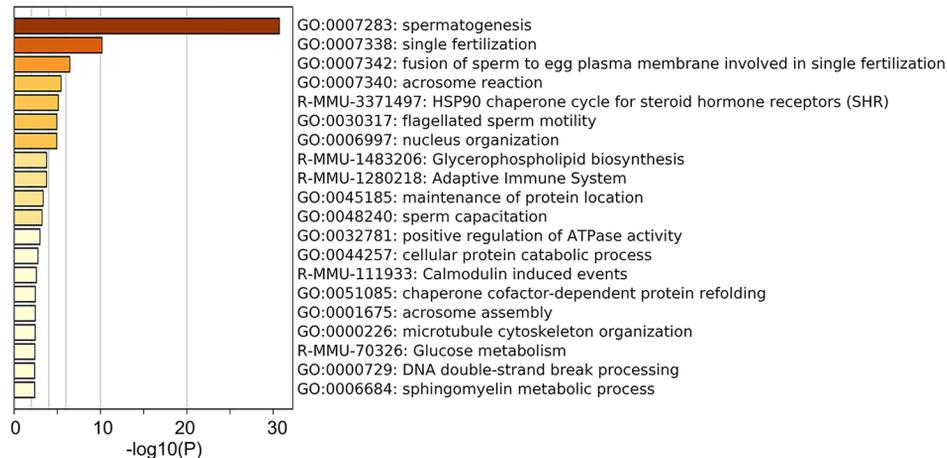
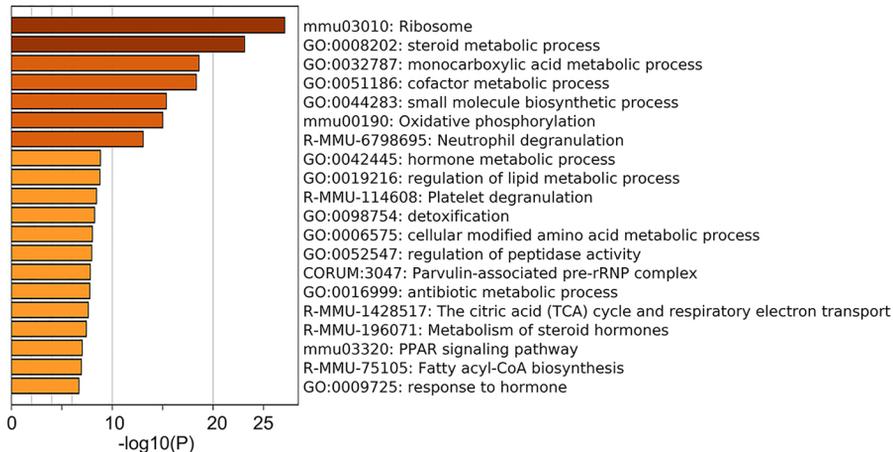
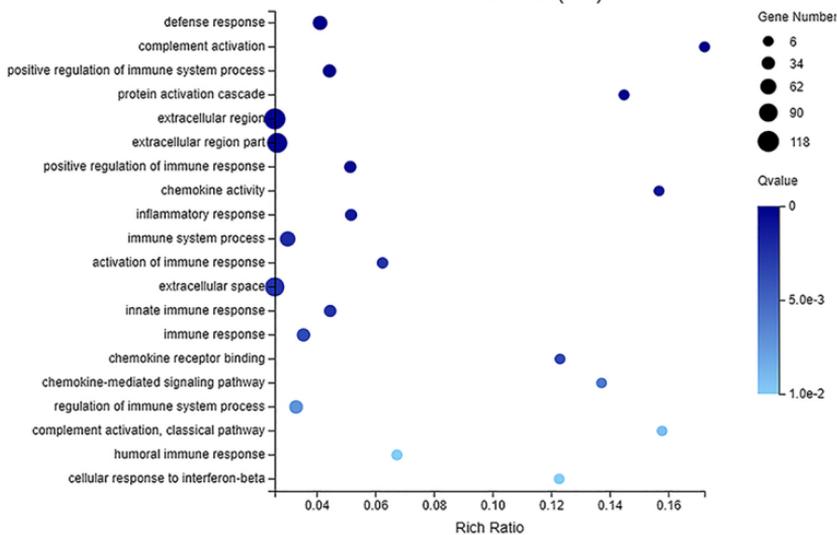
A**Spg****B****Spc****C****St****D****LC/SC**

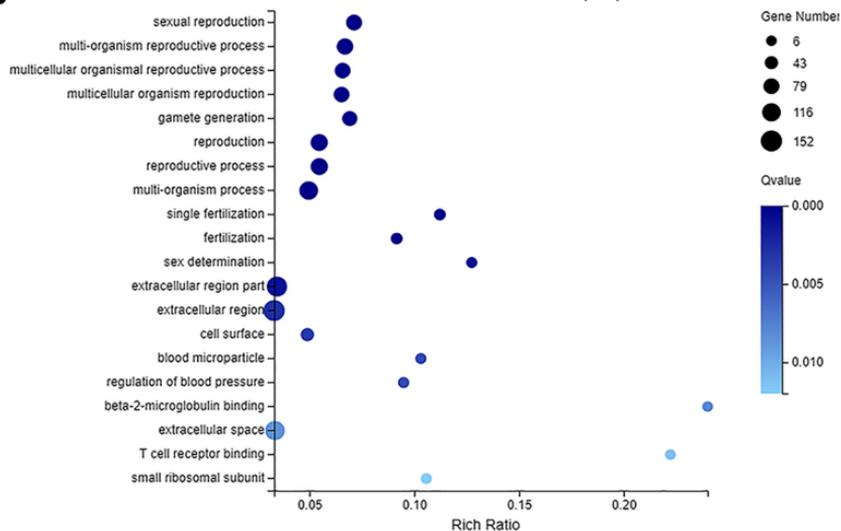
Figure S4. Enrichment analysis data for each group sampled individually.

A

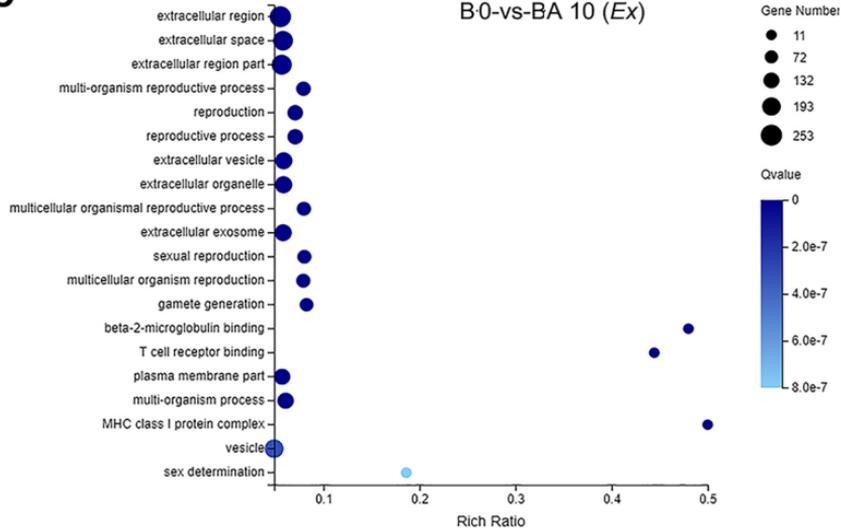
A0-vs-A10 (Ex)

**B**

A0-vs-A50 (Ex)

**C**

B0-vs-BA10 (Ex)

**D**

B0-vs-BA50 (Ex)

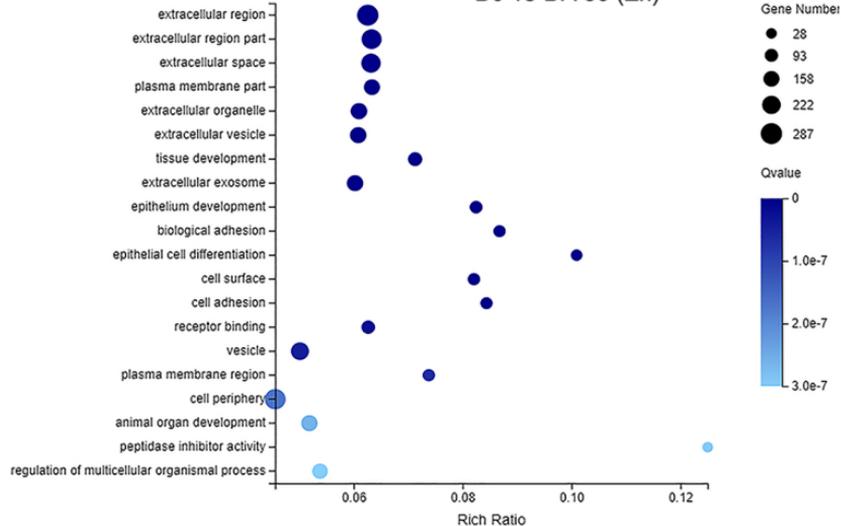


Figure S5. GO enrichment analysis for down-regulated genes in each of the four comparisons (*ex vivo*).