Cell Reports Medicine, Volume *2*

Supplemental information

Single-cell RNA-seq unravels alterations of the human spermatogonial stem cell compartment in patients with impaired spermatogenesis

Sara Di Persio, Tobias Tekath, Lara Marie Siebert-Kuss, Jann-Frederik Cremers, Joachim Wistuba, Xiaolin Li, Gerd Meyer zu Hörste, Hannes C.A. Drexler, Margot Julia Wyrwoll, Frank Tüttelmann, Martin Dugas, Sabine Kliesch, Stefan Schlatt, Sandra Laurentino, and Nina Neuhaus

SUPPLEMENTAL FIGURES

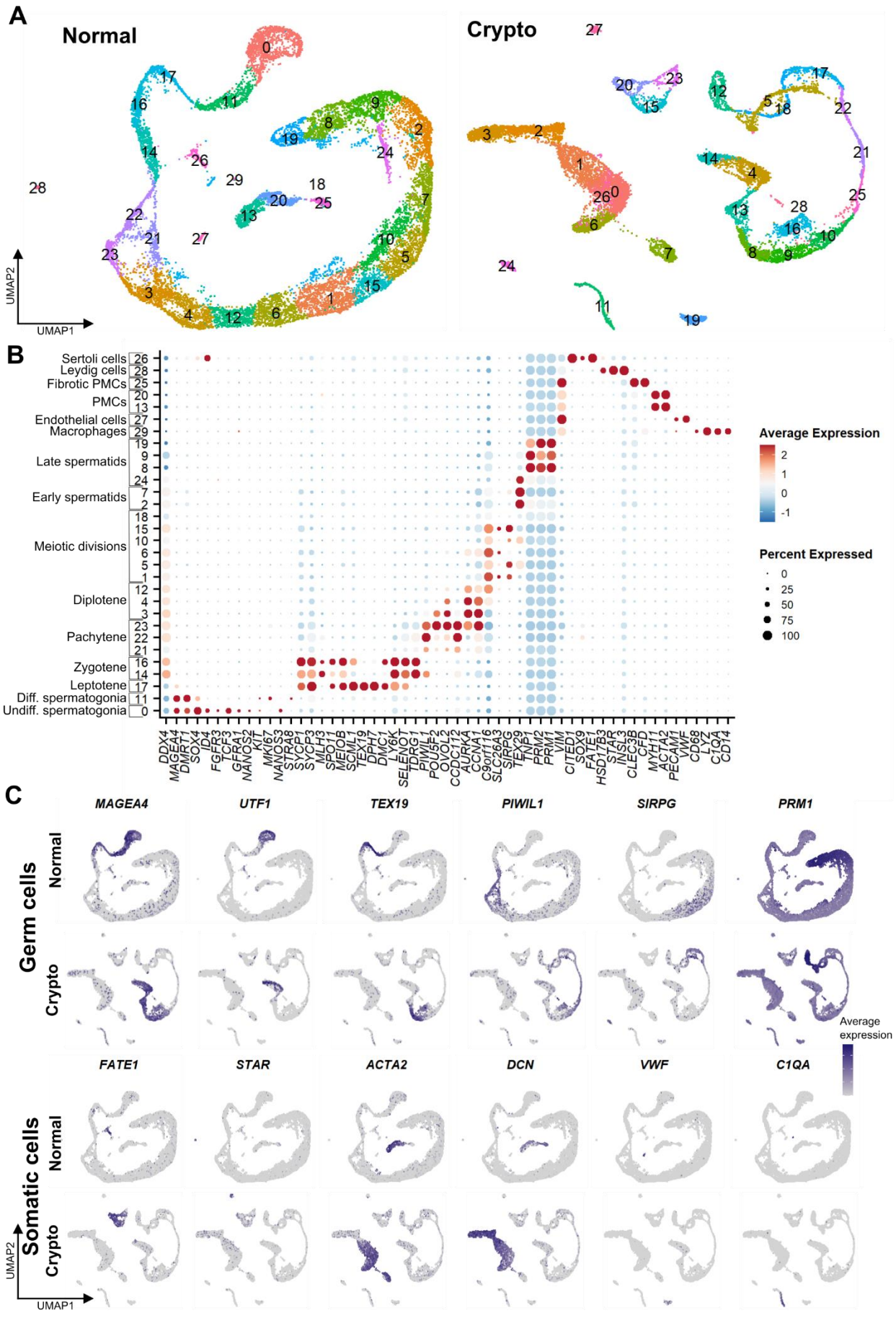
A
Normal
Crypto
UMAP2
UMAP1
B
Sertoli cells
Leydig cells
Fibrotic PMCs
PMCs
Endothelial cells
Macrophages
Late spermatids
Early spermatids
Meiotic divisions
Diplotene
Pachytene
Zygotene
Leptotene
Diff. spermatogonia
Undiff. spermatogonia
DDX4 MAGEA4 DMRT1 SOX4 ID4 FGFR3 TCF3 GFRA1 NANOS2 KIT MKI67 NANOS3 STRA8 SYCP1 SYCP3 MLH3 SPO11 MEIOB SCML1 TEX19 DPH7 DMC1 LY6K SELENOT TDRG1 PIWIL1 POU5F2 CCDC112 AURKA CCNA1 C9orf116 SLC26A3 SIRPG TEX29 TNP1 PRM2 PRM1 VIM CITED1 SOX9 FATE1 HSD17B3 STAR INSL3 CLEC3B CFD MYH11 ACTA2 PECAM1 VWF CD68 LYZ C1QA CD14
Average Expression
Percent Expressed
C
MAGEA4 UTF1 TEX19 PIWIL1 SIRPG PRM1
Germ cells
Normal
Crypto
Average expression
FATE1 STAR ACTA2 DCN VWF C1QA
Somatic cells
Normal
Crypto
UMAP2
UMAP1

Figure S1: Clustering analysis of normal and crypto datasets. (Related to Figure 1)

(A) Uniform manifold approximation and projection (UMAP) plot of the clustering analysis of the integrated (left) normal (15 546 cells, 30 clusters) and (right) crypto (13 144 cells, 29 clusters) datasets.

(B) Dot plot showing the relative expression of 55 marker genes in the 30 normal clusters. Cell identity was assigned according to marker expression. The color and size of the dots represent the average expression and the percentage of cells expressing each marker in a cluster, as indicated in the key.

(C) Feature plots showing the expression of germ cell and somatic cell marker genes in the normal and crypto datasets.

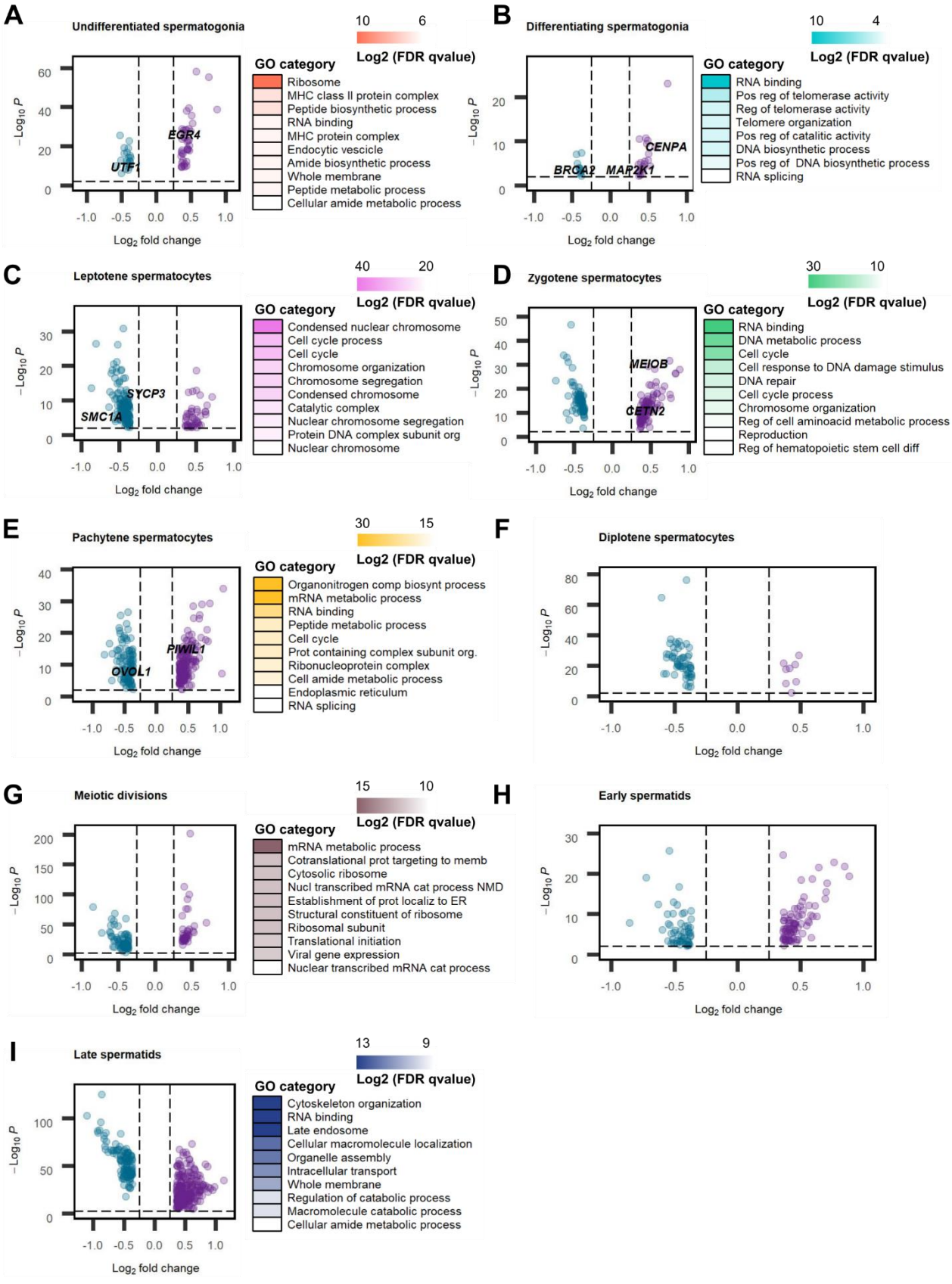
A
Undifferentiated spermatogonia
10
6
Log2 (FDR qvalue)
GO category
Ribosome
MHC class II protein complex
Peptide biosynthetic process
RNA binding
MHC protein complex
Endocytic vesicle
Amide biosynthetic process
Whole membrane
Peptide metabolic process
Cellular amide metabolic process
EGR4
UTF1
−Log10 P
Log2 fold change
B
Differentiating spermatogonia
10
4
Log2 (FDR qvalue)
GO category
RNA binding
Pos reg of telomerase activity
Reg of telomerase activity
Telomere organization
Pos reg of catalitic activity
DNA biosynthetic process
Pos reg of DNA biosynthetic process
RNA splicing
CENPA
BRCA2
MAP2K1
C
Leptotene spermatocytes
40
20
Log2 (FDR qvalue)
GO category
Condensed nuclear chromosome
Cell cycle process
Cell cycle
Chromosome organization
Chromosome segregation
Condensed chromosome
Catalytic complex
Nuclear chromosome segregation
Protein DNA complex subunit org
Nuclear chromosome
SYCP3
SMC1A
D
Zygotene spermatocytes
30
10
Log2 (FDR qvalue)
GO category
RNA binding
DNA metabolic process
Cell cycle
Cell response to DNA damage stimulus
DNA repair
Cell cycle process
Chromosome organization
Reg of cell aminoacid metabolic process
Reproduction
Reg of hematopoietic stem cell diff
MEIOB
CETN2
E
Pachytene spermatocytes
30
15
Log2 (FDR qvalue)
GO category
Organonitrogen comp biosynt process
mRNA metabolic process
RNA binding
Peptide metabolic process
Cell cycle
Prot containing complex subunit org.
Ribonucleoprotein complex
Cell amide metabolic process
Endoplasmic reticulum
RNA splicing
PIWIL1
OVOL1
F
Diplotene spermatocytes
G
Meiotic divisions
15
10
Log2 (FDR qvalue)
GO category
mRNA metabolic process
Cotranslational prot targeting to memb
Cytosolic ribosome
Nucl transcribed mRNA cat process NMD
Establishment of prot localiz to ER
Structural constituent of ribosome
Ribosomal subunit
Translational initiation
Viral gene expression
Nuclear transcribed mRNA cat process
H
Early spermatids
I
Late spermatids
13
9
Log2 (FDR qvalue)
GO category
Cytoskeleton organization
RNA binding
Late endosome
Cellular macromolecule localization
Organelle assembly
Intracellular transport
Whole membrane
Regulation of catabolic process
Macromolecule catabolic process
Cellular amide metabolic process

Figure S2: DGE and GO analysis between the different germ cell clusters of the normal and crypto datasets. (Related to Figure 2)

(A) Volcano plot (Left) and gene ontology (GO) analysis (Right) of the 61 genes uniquely differentially expressed in the undifferentiated spermatogonial cluster.

(B) Volcano plot (Left) and GO analysis (Right) of the 32 genes uniquely differentially expressed in the differentiating spermatogonial cluster.

(C) Volcano plot (Left) and GO analysis (Right) of the 164 genes uniquely differentially expressed in the leptotene spermatocyte cluster.

(D) Volcano plot (Left) and GO analysis (Right) of the 180 genes uniquely differentially expressed in the zygotene spermatocyte cluster.

(E) Volcano plot (Left) and GO analysis (Right) of the 271 genes uniquely differentially expressed in the pachytene spermatocyte cluster.

(F) Volcano plot of the 66 genes uniquely differentially expressed in the diplotene spermatocyte cluster.

(G) Volcano plot (Left) and GO analysis (Right) of the 121 genes uniquely differentially expressed in the meiotic division cluster.

(H) Volcano plot of the 142 genes uniquely differentially expressed in the early spermatid cluster.

(I) Volcano plot (Left) and GO analysis (Right) of the 405 genes uniquely differentially expressed in the late spermatid cluster. See Table S4 for the statistical details.

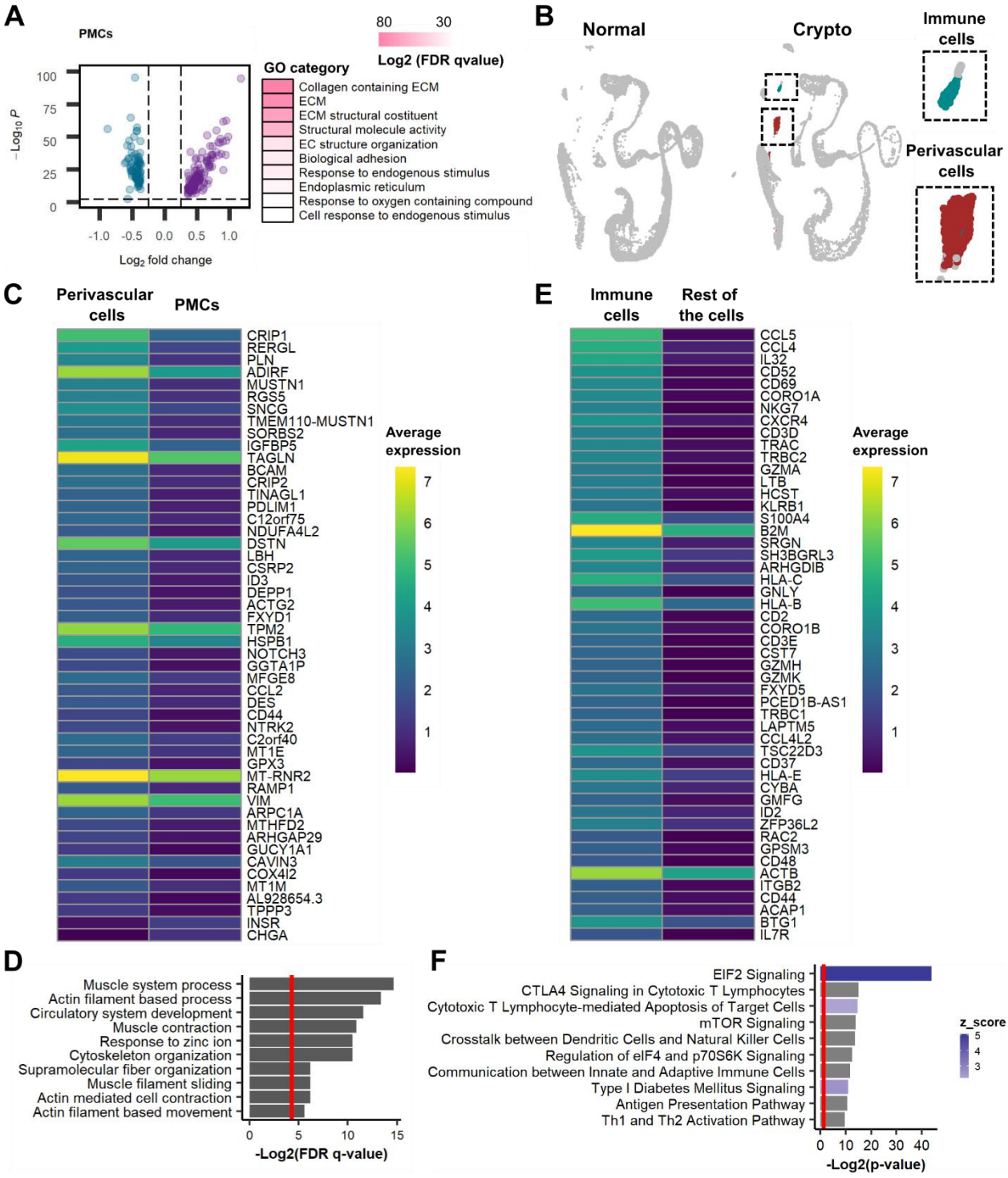
A
PMCs
80
30
Log2 (FDR qvalue)
GO category
Collagen containing ECM
ECM
ECM structural costituent
Structural molecule activity
EC structure organization
Biological adhesion
Response to endogenous stimulus
Endoplasmic reticulum
Response to oxygen containing compound
Cell response to endogenous stimulus
−Log10 P
Log2 fold change
B
Normal
Crypto
Immune cells
Perivascular cells
C
Perivascular cells
PMCs
CRIP1
RERGL
PLN
ADIRF
MUSTN1
RGS5
SNCG
TMEM110-MUSTN1
SORBS2
IGFBP5
TAGLN
BCAM
CRIP2
TINAGL1
PDLIM1
C12orf75
NDUFA4L2
DSTN
LBH
CSRP2
ID3
DEPP1
ACTG2
FXYD1
TPM2
HSPB1
NOTCH3
GGTA1P
MFGE8
CCL2
DES
CD44
NTRK2
C2orf40
MT1E
GPX3
MT-RNR2
RAMP1
VIM
ARPC1A
MTHFD2
ARHGAP29
GUCY1A1
CAVIN3
COX4I2
MT1M
AL928654.3
TPPP3
INSR
CHGA
Average expression
E
Immune cells
Rest of the cells
CCL5
CCL4
IL32
CD52
CD69
CORO1A
NKG7
CXCR4
CD3D
TRAC
TRBC2
GZMA
LTB
HCST
KLRB1
S100A4
B2M
SRGN
SH3BGRL3
ARHGDIB
HLA-C
GNLY
HLA-B
CD2
CORO1B
CD3E
CST7
GZMH
GZMK
FXYD5
PCED1B-AS1
TRBC1
LAPTM5
CCL4L2
TSC22D3
CD37
HLA-E
CYBA
GMFG
ID2
ZFP36L2
RAC2
GPSM3
CD48
ACTB
ITGB2
CD44
ACAP1
BTG1
IL7R
Average expression
D
Muscle system process
Actin filament based process
Circulatory system development
Muscle contraction
Response to zinc ion
Cytoskeleton organization
Supramolecular fiber organization
Muscle filament sliding
Actin mediated cell contraction
Actin filament based movement
-Log2(FDR q-value)
F
EIF2 Signaling
CTLA4 Signaling in Cytotoxic T Lymphocytes
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells
mTOR Signaling
Crosstalk between Dendritic Cells and Natural Killer Cells
Regulation of eIF4 and p70S6K Signaling
Communication between Innate and Adaptive Immune Cells
Type I Diabetes Mellitus Signaling
Antigen Presentation Pathway
Th1 and Th2 Activation Pathway
-Log2(p-value)
z_score

Figure S3: DGE and GO analysis between the different somatic cell clusters of the normal and crypto datasets. (Related to Figure 2)

(A) Volcano plot (Left) and GO analysis (Right) of the 232 genes uniquely differentially expressed in the PMC cluster. See Supplementary information, Table S4 for the statistical details.

(B) UMAP plot of the normal (Left) and crypto (Right) datasets showing the unique presence of the peritubular (Red) and immune (Green) cell clusters in the crypto dataset.

(C) Heatmap showing the differential gene expression analysis between the perivascular cell clusters and the PMCs in the crypto dataset. The analysis identified 48 differentially expressed genes (DEG), including *MUSTN1*, with a higher expression in perivascular cells (Table S5).

(D) The bar plot shows the GO analysis of the 48 DEG found in perivascular cells. The results were enriched in GO terms related to muscle and cytoskeletal contraction. The red line on the bar plot represent p=0.05.

(E) Heatmap showing the first 50 genes of the differential gene expression analysis between immune cells and the rest of the cells in the crypto dataset. A total of 195 genes showed higher expression in cluster 24 (A complete list of the DEG is available Table S5).

(F) The bar plot shows the pathway analysis of the 195 DEG. The results were enriched in immune cell associated pathways. The red line on the bar plot represents the p=0.05. The z_score represents the activation level of each pathway.

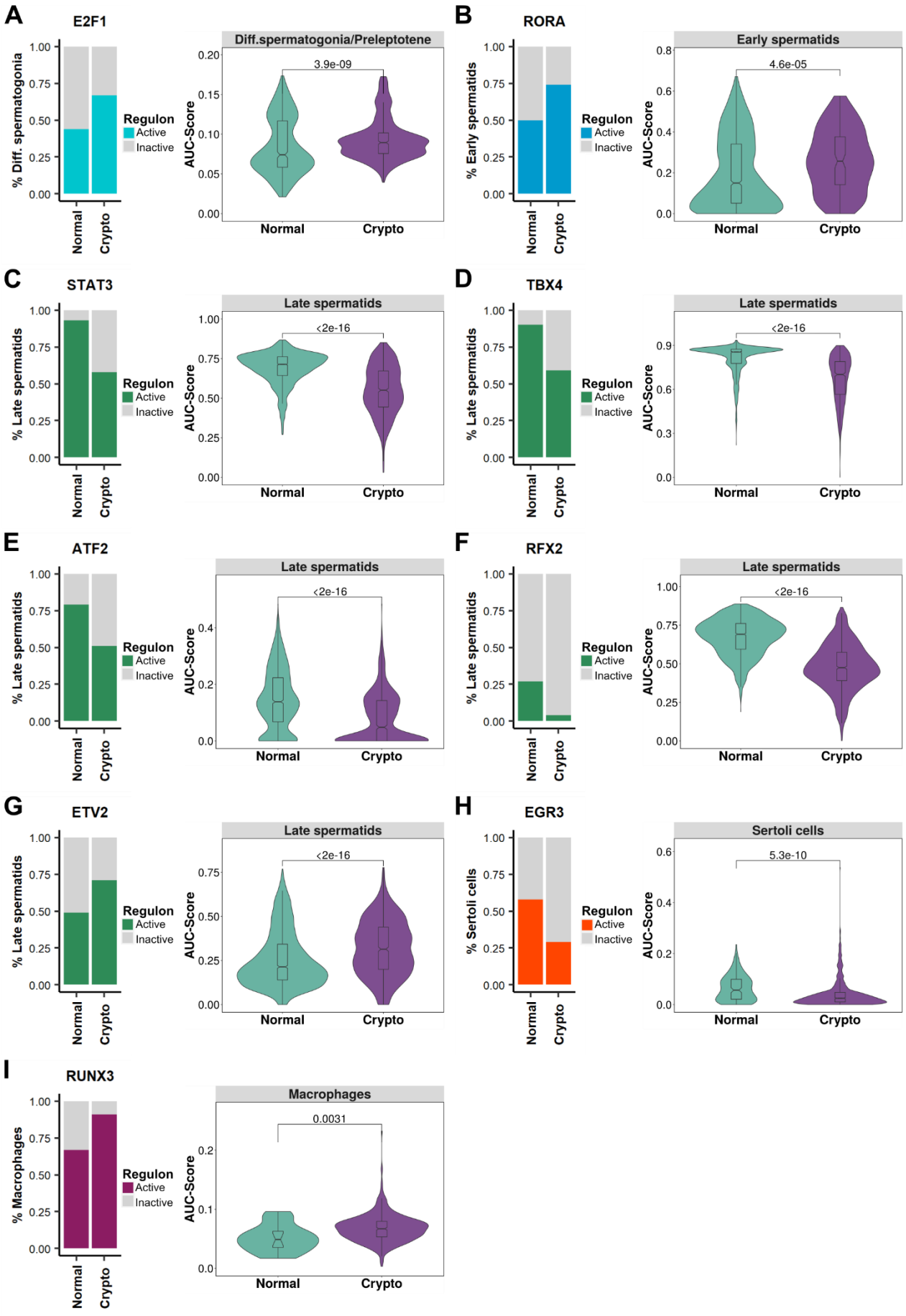
A
E2F1
% Diff. spermatogonia
Regulon
Active
Inactive
Normal
Crypto
Diff.spermatogonia/Preleptotene
AUC-Score
3.9e-09
B
RORA
% Early spermatids
Early spermatids
4.6e-05
C
STAT3
% Late spermatids
Late spermatids
<2e-16
D
TBX4
E
ATF2
F
RFX2
G
ETV2
H
EGR3
% Sertoli cells
Sertoli cells
5.3e-10
I
RUNX3
% Macrophages
Macrophages
0.0031

Figure S4: Differential regulon activation in the normal and crypto datasets. (Related to Figure 4)

(A-I) Left panel: Stacked bar plot comparing the proportion of cells with active regulons. Right panel: Violin plots comparing the AUC score of the regulons in the normal and crypto dataset. The regulons are organized as follow: **(A)** E2F1 in the differentiating spermatogonia/preleptotene, **(B)** RORA in the early spermatids, **(C)** STAT3, **(D)** TBX4, **(E)** ATF2, **(F)** RFX2 and **(G)** ETV2 in the late spermatids, **(H)** EGR3 in the Sertoli cells and **(I)** RUNX3 in the macrophages. A significant change was found for all the nine regulons in the crypto dataset.

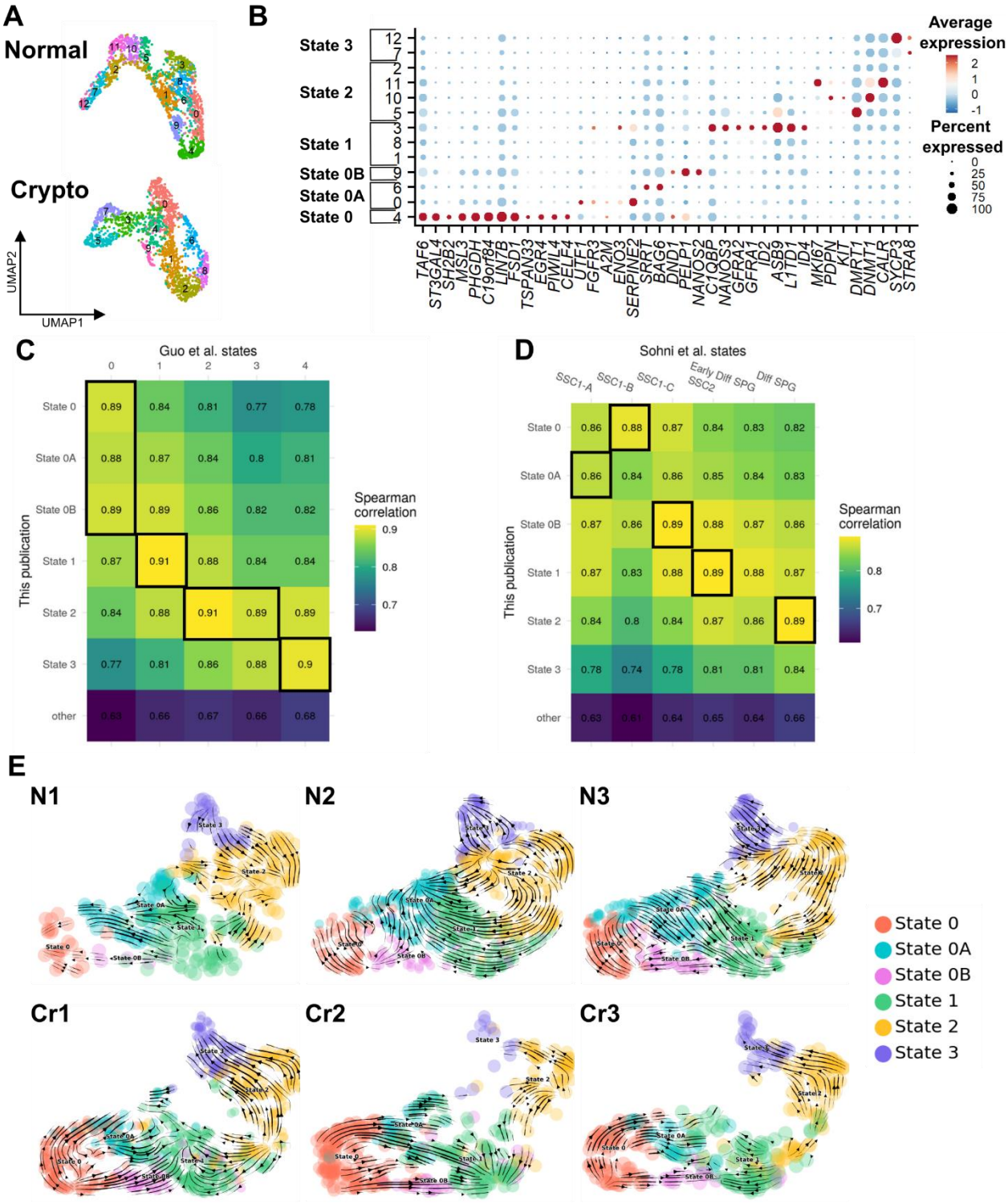
A
Normal
Crypto
UMAP2
UMAP1
B
State 3
State 2
State 1
State 0B
State 0A
State 0
TAF6
ST3GAL4
SH2B2
MSL3
PHGDH
C19orf84
LIN7B
FSD1
TSPAN33
EGR4
PIWIL4
CELF4
UTF1
FGFR3
A2M
ENO3
SERPINE2
SRRT
BAG6
DND1
PELP1
NANOS2
C1QBP
NANOS3
GFRA2
GFRA1
ID2
ASB9
L1TD1
ID4
MKI67
PDPN
KIT
DMRT1
DNMT1
CALR
SYCP3
STRA8
Average expression
Percent expressed
C
Guo et al. states
This publication
Spearman correlation
D
Sohni et al. states
SSC1-A
SSC1-B
SSC1-C
SSC2
Early Diff SPG
Diff SPG
This publication
Spearman correlation
E
N1
N2
N3
Cr1
Cr2
Cr3
State 0
State 0A
State 0B
State 1
State 2
State 3

Figure S5: Cluster analysis and assignment of normal and crypto spermatogonial datasets. (Related to Figure 5)

(A) UMAP plot of the integrated normal (upper panel, 13 clusters) and crypto (lower panel, 10 clusters) spermatogonial datasets. The cells are color coded according to the cluster.

(B) Dot plot showing the relative expression of selected marker genes in the 13 normal clusters. The color and size of the dots represent the average expression and the percentage of cells expressing each marker in a cluster, as indicated in the key. According to the marker expression, six clusters were defined: State 0, 0A, 0B, 1, 2, and 3.

(C-D) Correlation analysis comparing the transcriptome of the different spermatogonial states defined in this publication with those defined in Guo et al., 2018 **(C)** and Sohni et al., 2019 **(D)**. The 'other' state is comprised of all non-spermatogonial cell types available in this publication. The black boxes show the states where we expected the highest correlation in each state.

(E) Contribution of each normal and crypto sample to the RNA velocity derived from the scVelo dynamical model visualized as streamlines in UMAP plots.

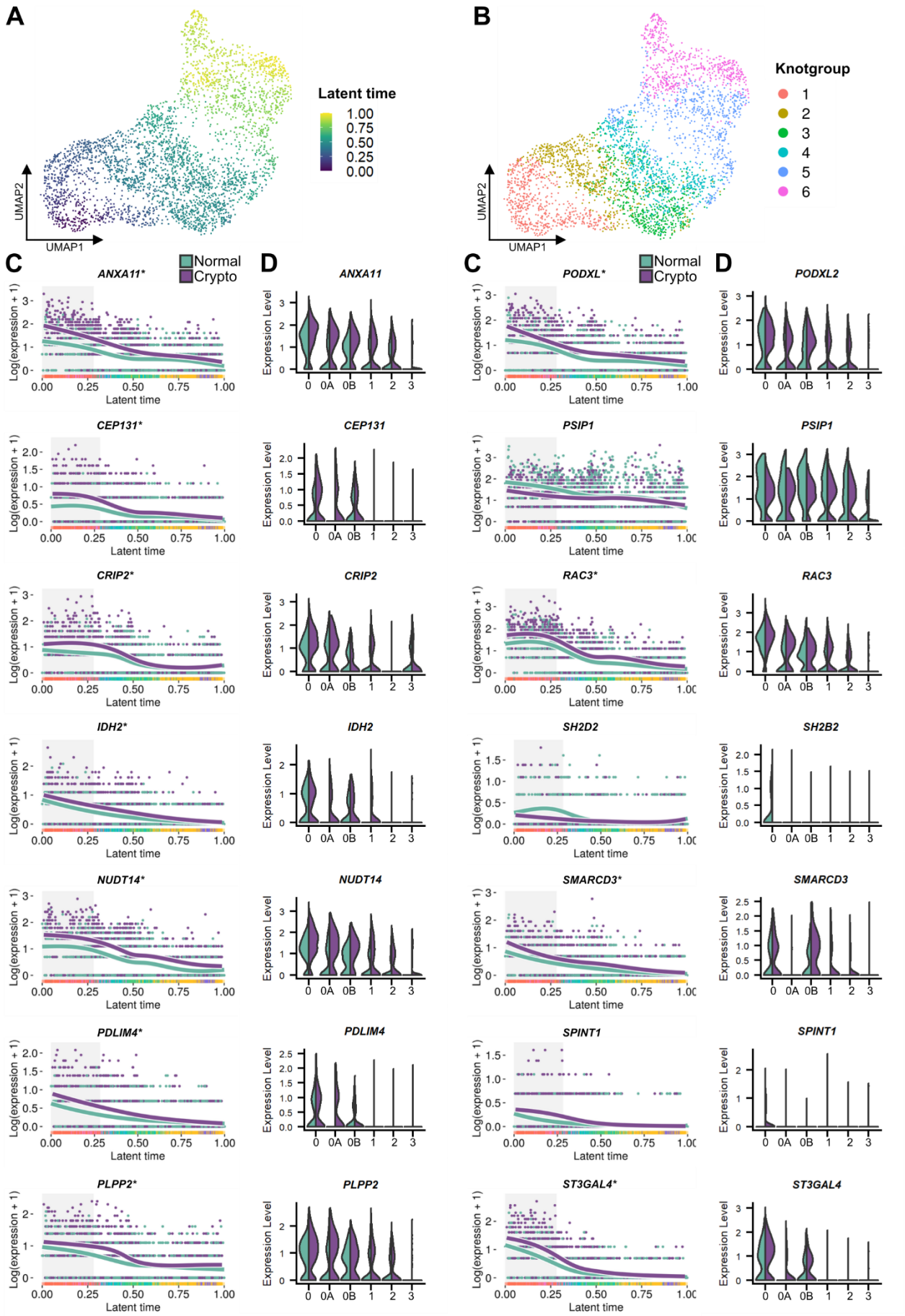
A
Latent time
1.00
0.75
0.50
0.25
0.00
UMAP2
UMAP1
B
Knotgroup
1
2
3
4
5
6
UMAP2
UMAP1
C
Normal
Crypto
D
ANXA11*
ANXA11
PODXL*
PODXL2
CEP131*
CEP131
PSIP1
PSIP1
CRIP2*
CRIP2
RAC3*
RAC3
IDH2*
IDH2
SH2D2
SH2B2
NUDT14*
NUDT14
SMARCD3*
SMARCD3
PDLIM4*
PDLIM4
SPINT1
SPINT1
PLPP2*
PLPP2
ST3GAL4*
ST3GAL4
Log(expression + 1)
Expression Level
Latent time
0 0A 0B 1 2 3

Figure S6: EGR4-regulated genes in the knotgroup 1 spermatogonia. (Related to Figure 7)

(A) UMAP plot showing the integrated normal-crypto spermatogonial dataset aligned along the latent time. The cells are color-coded according to their progression along the latent time. State 0 was set as starting point of the differentiation process.

(B) UMAP plot showing the subdivision of the integrated normal-crypto spermatogonial dataset into six knotgroups. The cells are color coded according to their respective knotgroup.

(C) Line plots show the expression along the latent time of the EGR4-regulated differentially expressed genes in normal (teal) and crypto (purple) spermatogonia included in the knotgroup 1: *ANXA11*, *CEP131*, *CRIP2*, *IDH2*, *NUDT14*, *PDLIM4*, *PLPP2*, *PODXL2*, *PSIP1*, *RAC3*, *SH2B2*, *SMARCD3*, *SPINT1*, and *ST3GAL4*. The gray area highlights the cells belonging to the knotgroup 1. The asterisk indicates EGR4-regulated genes that were also differentially expressed in knotgroup 2.

(D) Double violin plots comparing the expression levels of EGR4-regulated *ANXA11*, *CEP131*, *CRIP2*, *IDH2*, *NUDT14*, *PDLIM4*, *PLPP2*, *PODXL2*, *PSIP1*, *RAC3*, *SH2B2*, *SMARCD3*, *SPINT1*, and *ST3GAL4* in the spermatogonial states of the normal (teal) and crypto (purple) datasets.

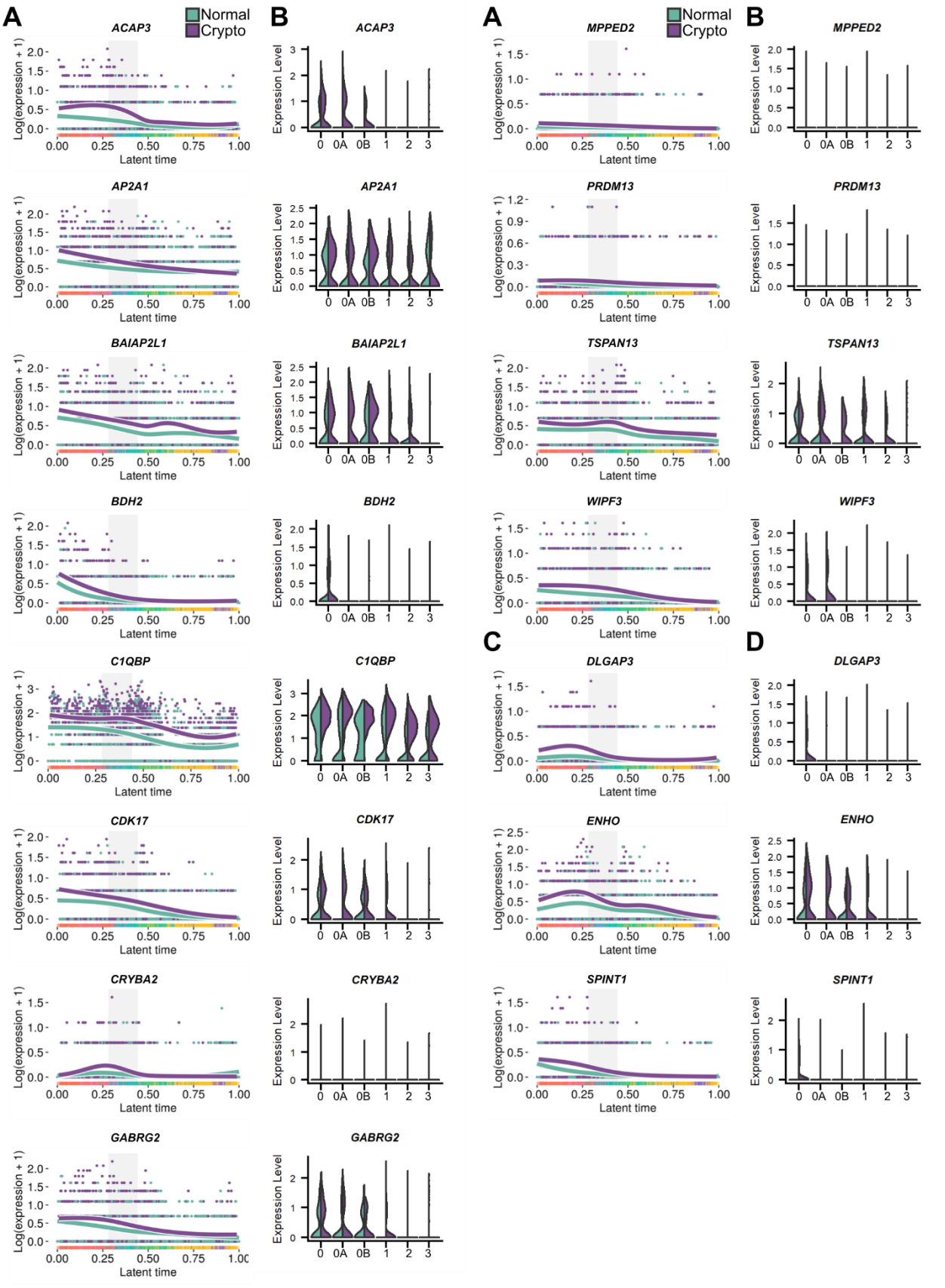
A
B
Normal
Crypto
ACAP3
AP2A1
BAIAP2L1
BDH2
C1QBP
CDK17
CRYBA2
GABRG2
Log(expression + 1)
Latent time
Expression Level
0 0A 0B 1 2 3
A
B
Normal
Crypto
MPPED2
PRDM13
TSPAN13
WIPF3
C
D
DLGAP3
ENHO
SPINT1
Log(expression + 1)
Latent time
Expression Level
0 0A 0B 1 2 3

Figure S7: EGR4- and HOXC9-regulated genes in knotgroup 2 spermatogonia. (Related to Figure 7)

(A) Line plots showing the expression along the latent time of EGR4-regulated, differentially expressed genes between the normal (teal) and crypto (purple) spermatogonia in the knotgroup 2: *ACAP3*, *AP2A1*, *BAIAP2L1*, *BDH2*, *C1QBP*, *CDK17*, *CRYBA2*, *GABRG2*, *MPPED2*, *PRDM13*, *TSPAN13*, and *WIPF3*. The gray area highlights the cells belonging to the knotgroup 2.

(B) Double violin plots comparing the expression levels of EGR4-regulated genes *ACAP3*, *AP2A1*, *BAIAP2L1*, *BDH2*, *C1QBP*, *CDK17*, *CRYBA2*, *GABRG2*, *MPPED2*, *PRDM13*, *TSPAN13*, and *WIPF3* between the normal (teal) and crypto (purple) spermatogonial states.

(C) Line plots showing the expression along the latent time of the EGR4- and HOXC9-regulated, differentially expressed genes between the normal (teal) and crypto (purple) spermatogonia in knotgroup 2: *DLGAP3*, *ENHO*, and *SPINT1*. The gray area highlights the knotgroup 2.

(D) double violin plots comparing the expression levels of EGR4- and HOXC9-regulated genes *DLGAP3*, *ENHO*, and *SPINT1* between the normal (teal) and crypto (purple) spermatogonial states.

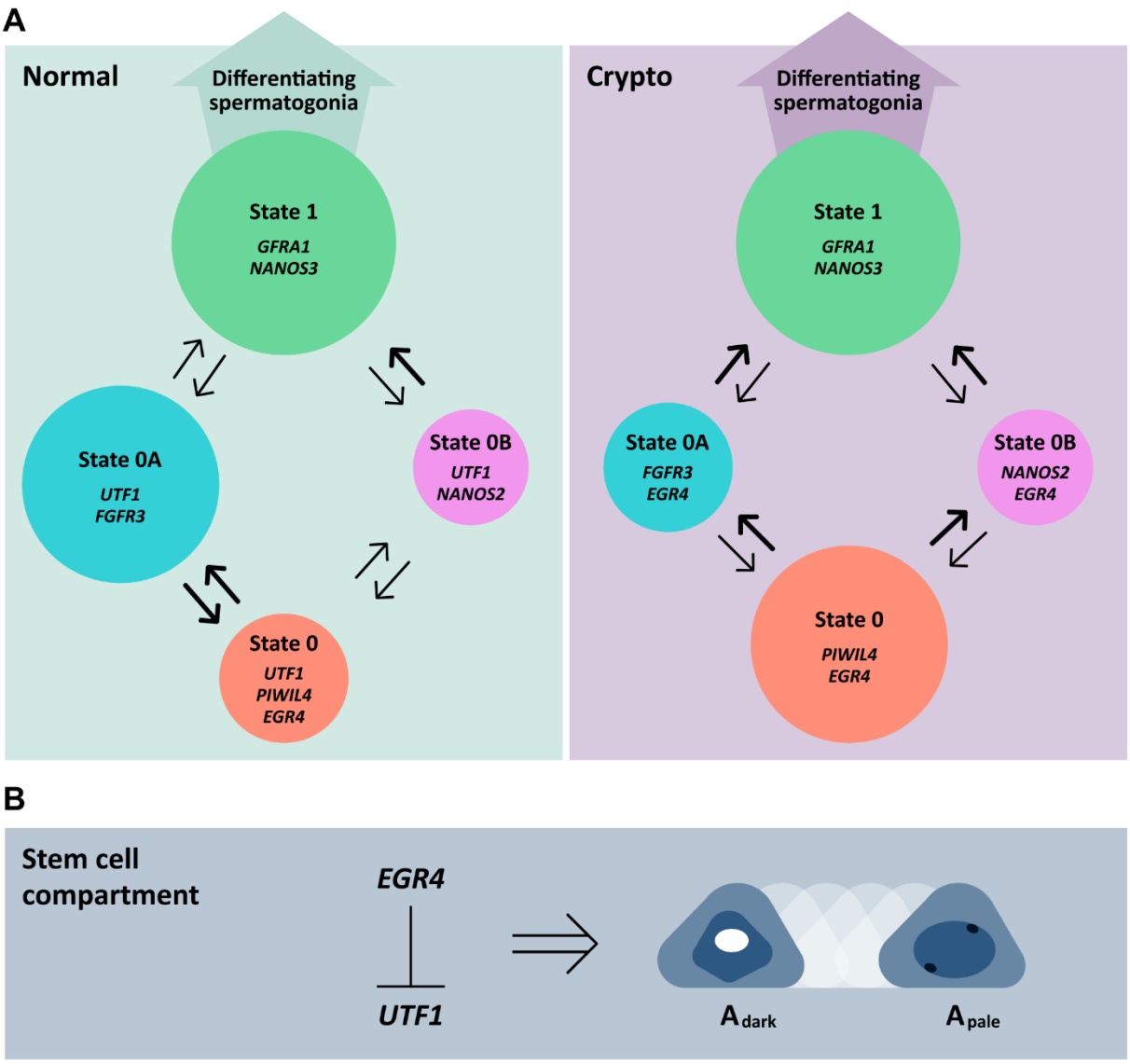

Figure S8: Model of the human spermatogonial compartment in normal and crypto samples. (Related to Discussion)

(A) The circles represent the State 0 (red), State 0A (blue), State 0B (magenta) and State 1 (green) spermatogonia. The thickness of the black arrows indicates the proportion of cells changing their expression profile. We showed that the A_{dark} spermatogonia are within all the different spermatogonial subpopulations with a low proportion within the $PIWIL4^{+}$ cells (State 0).

(B) Suggested mechanism: the EGR4 mediated UTF1 downregulation induces the chromatin remodeling resulting in the transition from A_{dark} to A_{pale} morphology.