PNAS
www.pnas.org

Supplementary Information for

Developmental and sexual dimorphic atlas of the prenatal mouse external genitalia at the single-cell level.

Ciro Maurizio Amato[1] and Humphrey Hung-Chang Yao[1*]

* Humphrey H-C Yao

Email: humphrey.yao@nih.gov

This PDF file includes:

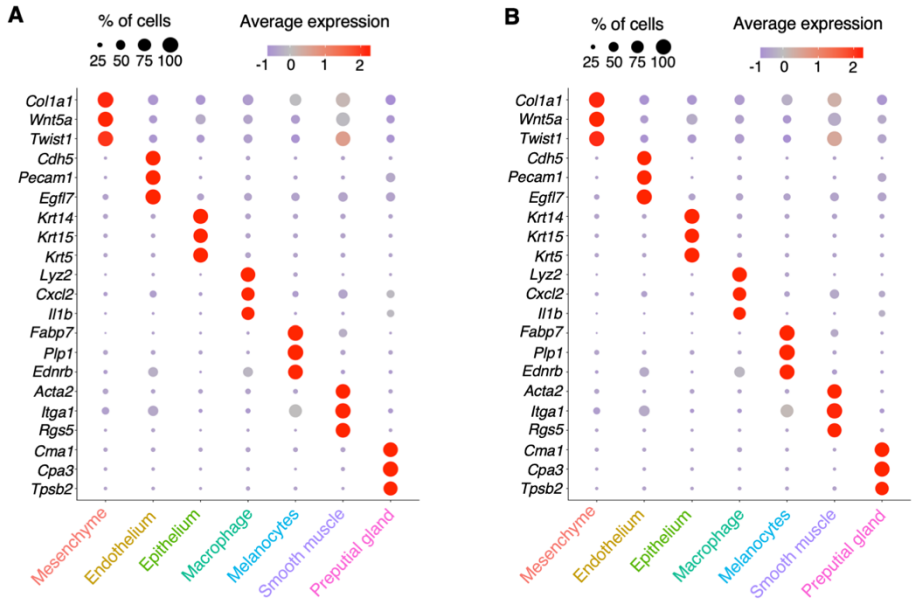
Supplemental methods
Figures S1 to S8
Table S1

Supplemental Methods

Cells with >1,000 and <7,500 genes per cell (nFeature_RNA) were kept for future analysis. Cells with <25% mitochondrial genes percentage were also maintained for downstream analysis. For Figs. 1–4, all time points for either male or female data were merged together. Standard log normalization, scaling, and clustering were conducted for both the male and female data. FindAllMarkers(test.use = Wilcoxon Rank Sum P value < 0.05, logfc.threshold = 0.25, min.cell > 0.1) was used to find genetic candidates for each cell cluster. Glanular and external prepuce subpopulations were isolated using subset and were rescaled, renormalized, and reclustered FindAllMarkers(test.use = Wilcoxon Rank Sum p-value < 0.05, logfc.threshold = 0.25, min.cell > 0.1) was used to identify markers for each subpopulation within the glans and external prepuce. To identify time-dependent changes in gene expression, the differential expression of E14.5 and E18.5 was calculated for each identified subpopulation. Differential gene lists for each subpopulation were then compared to each gene population gene list. Genes that did not overlap with other subpopulations and between sexes were kept and identified as the tissue and sex-specific development gene list. Resultant tissue and sex-specific gene lists were loaded into IPA to identify pathways associated with each subpopulation and tissue.

For sexual dimorphic comparisons, male and female data were combined and filtered and normalized as described above. External prepuce and glanular mesenchyme were separate and renormalized. Sexually dimorphic genes were found by conducting differential expression, FindAllMarkers(test.use = Wilcoxon Rank Sum, logfc.threshold = 0.25, min.cell > 0.1), between male and female for each developmental time point and each subpopulation. We then went through the sexually dimorphic gene lists and identified genes that were androgen or estrogen responsive, based on literature searches.

Supplemental Figure 1



Supplemental figure 1. Identification of subpopulations within the external genitalia. Dotplots of select significant marker genes in the male (A) and female (B) genitalia. Cell populations found are labeled with unique colors in both UMAP and dotplot. Dot size represent percentage of cells found with gene and color is the intensity of expression (red = high and blue = low).

Supplemental Figure 2

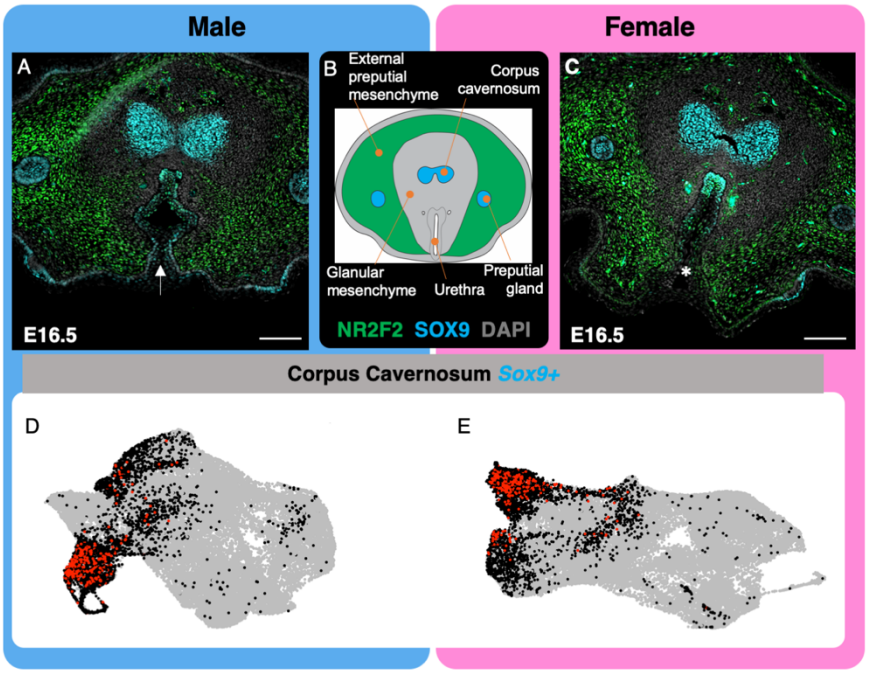


Fig. S2. Identification of subpopulations within the external genitalia. Immunohistochemistry of NR2F2 (green) and SOX9 (cyan) in male (A) and female (C) external genitalia. Representative cartoon of each structure in the external genitalia (B). UMAP representation of *Sox9* expression in male (D) and female (E) genitalia. Grey = little to no expression, black = moderate expression, and red = high gene expression. Scale bars represent 100 μm.

Supplemental Figure 3

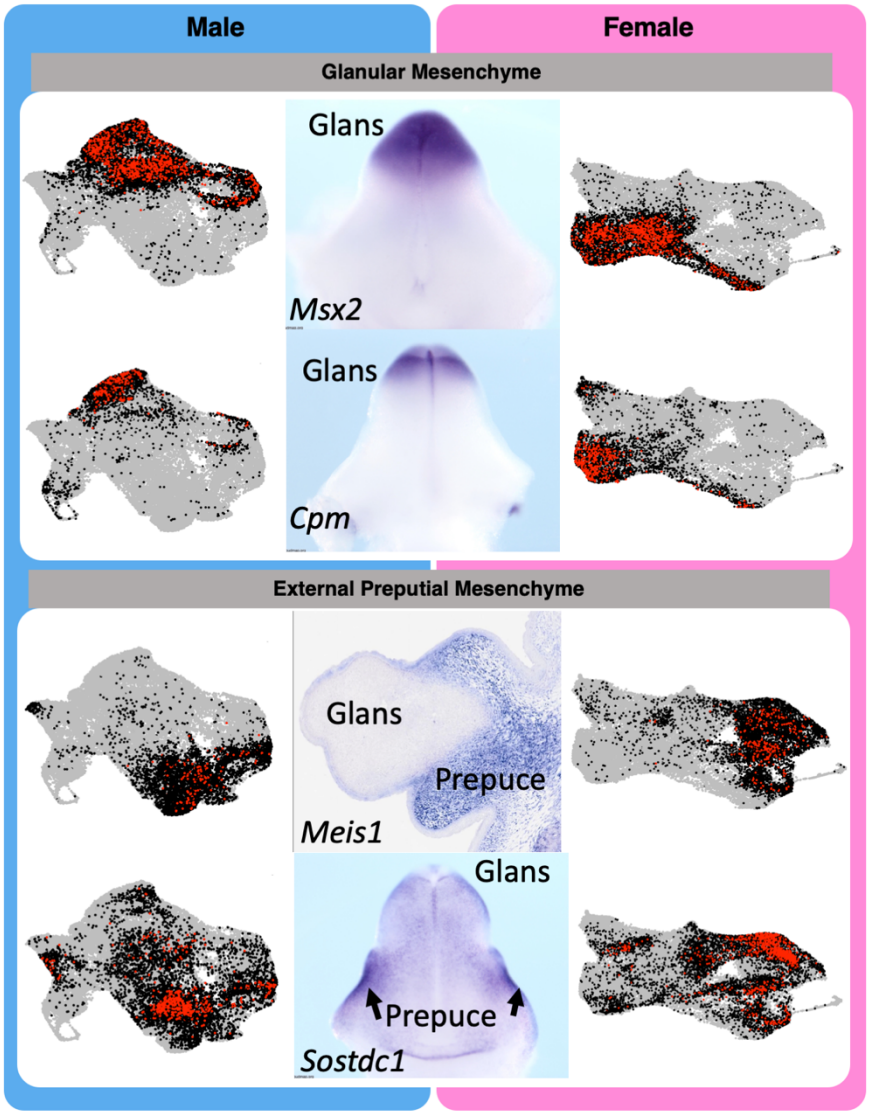


Fig. S3. Validation of marker gene expression in the glanular and preputial mesenchyme using GUDMAP data. Single cell UMAP gene expression for each marker gene in male and female genitalia. GUDMAP images are genitalia staged as E13.5 (*Msx2, Cpm, and Sostdc1*) and E15.5 (*Meis1*).

Supplemental Figure 4

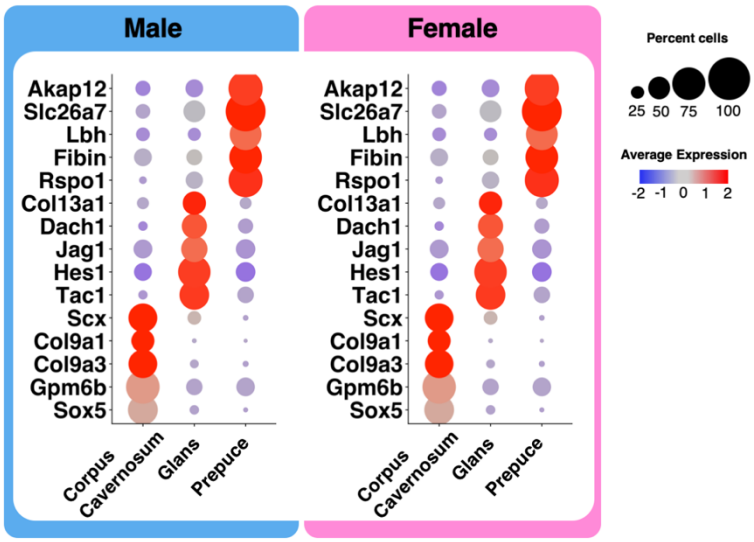


Fig. S4. Representative marker genes for corpus cavernosum, glans, and external prepuce. Five representative marker genes for the corpus cavernousm, glans and external prepuce in both male and female genitalia. Size of dot represents percentage of cell expressing the gene of interest. The color represents average expression, red= high and blue = low.

Supplemental Figure 5

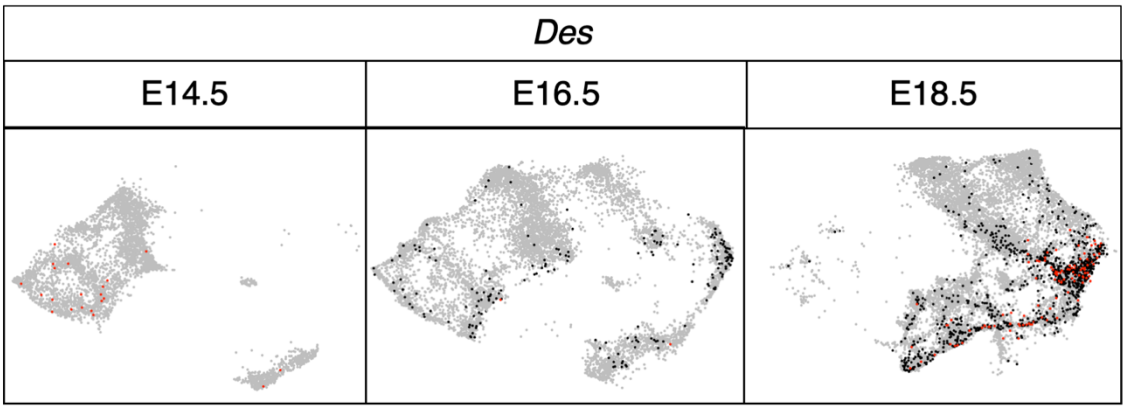


Fig. S5. *Des* UMAP. Visualization of *Des* in glanular mesenchyme subcluster separated by embryonic day. Grey = little to no expression, black = moderate expression, and red = high gene expression.

Supplemental Figure 6

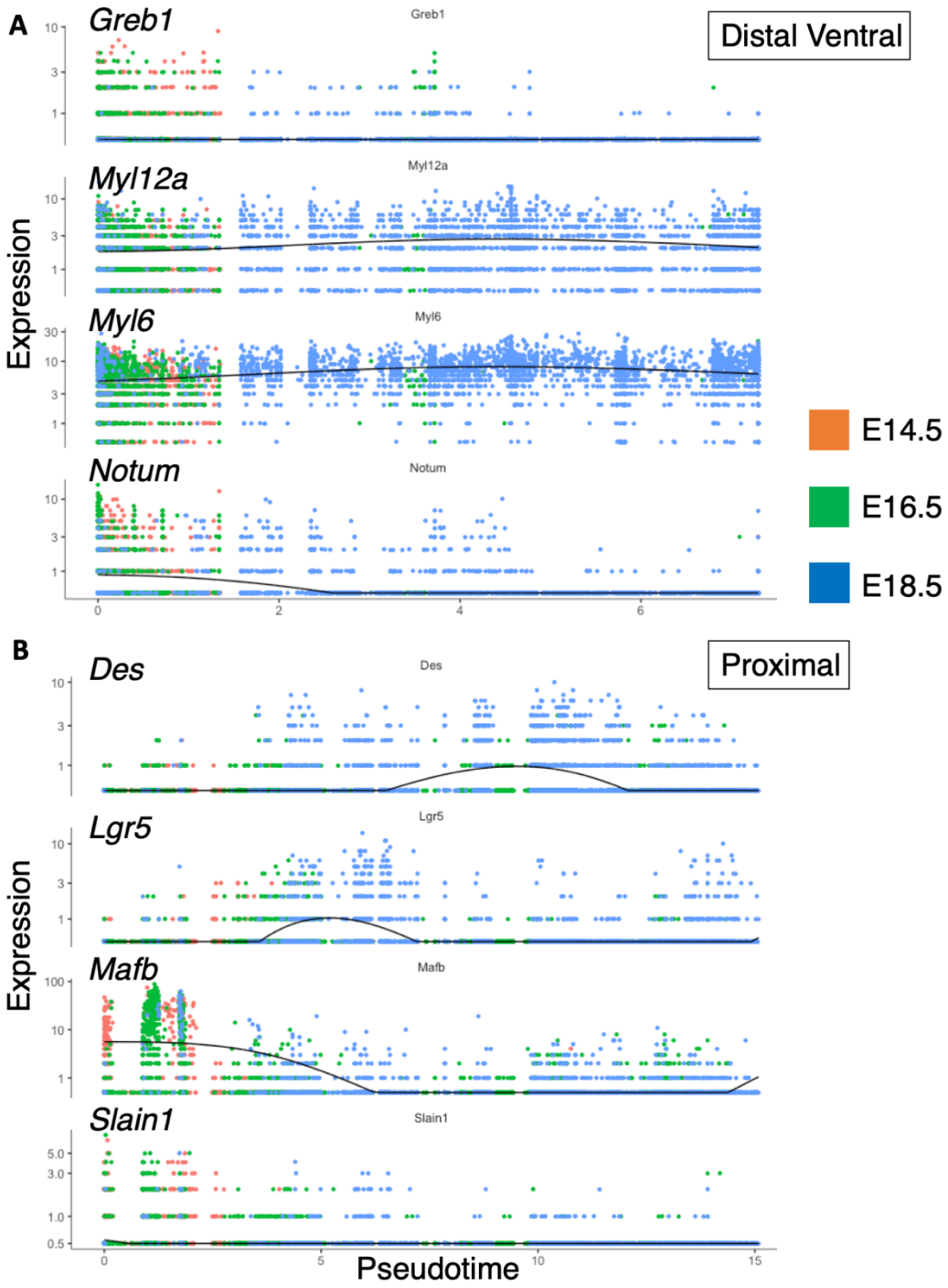


Fig. S6. Pseudotime of 2 glanular mesenchyme subpopluations. A) Pseudotime of the distal ventral glanular mesenchyme subpopulation. B) Pseudotime of the proximal glanular mesenchyme subpopulation. Red points = E14.5 samples, green points = E16.5 samples, and blue points = E18.5 samples. Black line represents a smoothened line fit to the pseudotime data.

Supplemental Figure 7

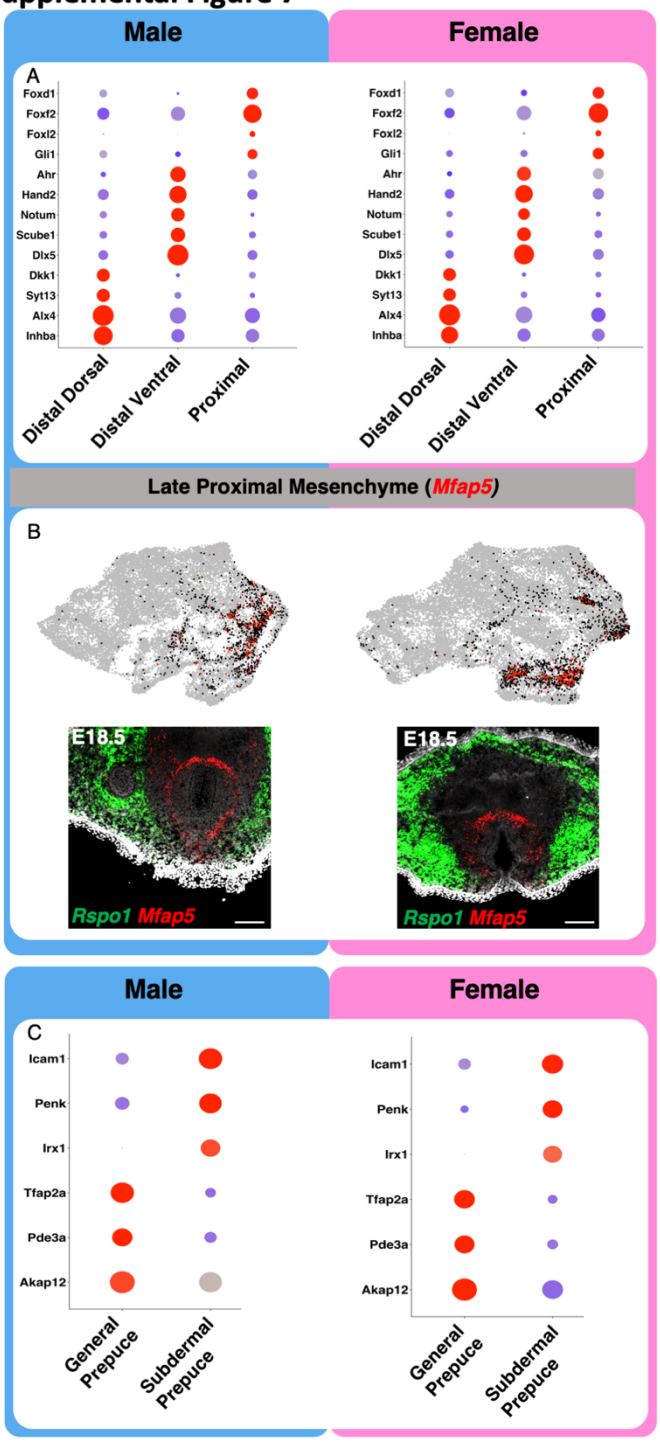


Fig. S7. Representative marker genes for the subpopulations within the glans and external prepuce. Marker genes for distal dorsal, distal ventral, and proximal glans in male and female genitalia (A). UMAP gene expression and validation of late proximal mesenchyme marker, *Mfap5* (red). *Rspo1 (*green) was used as a counterstain to mark the preputial mesenchyme. Marker genes for the general prepuce and subdermal prepuce in the male and female genitalia (C). Scale bars represent 100 μm.

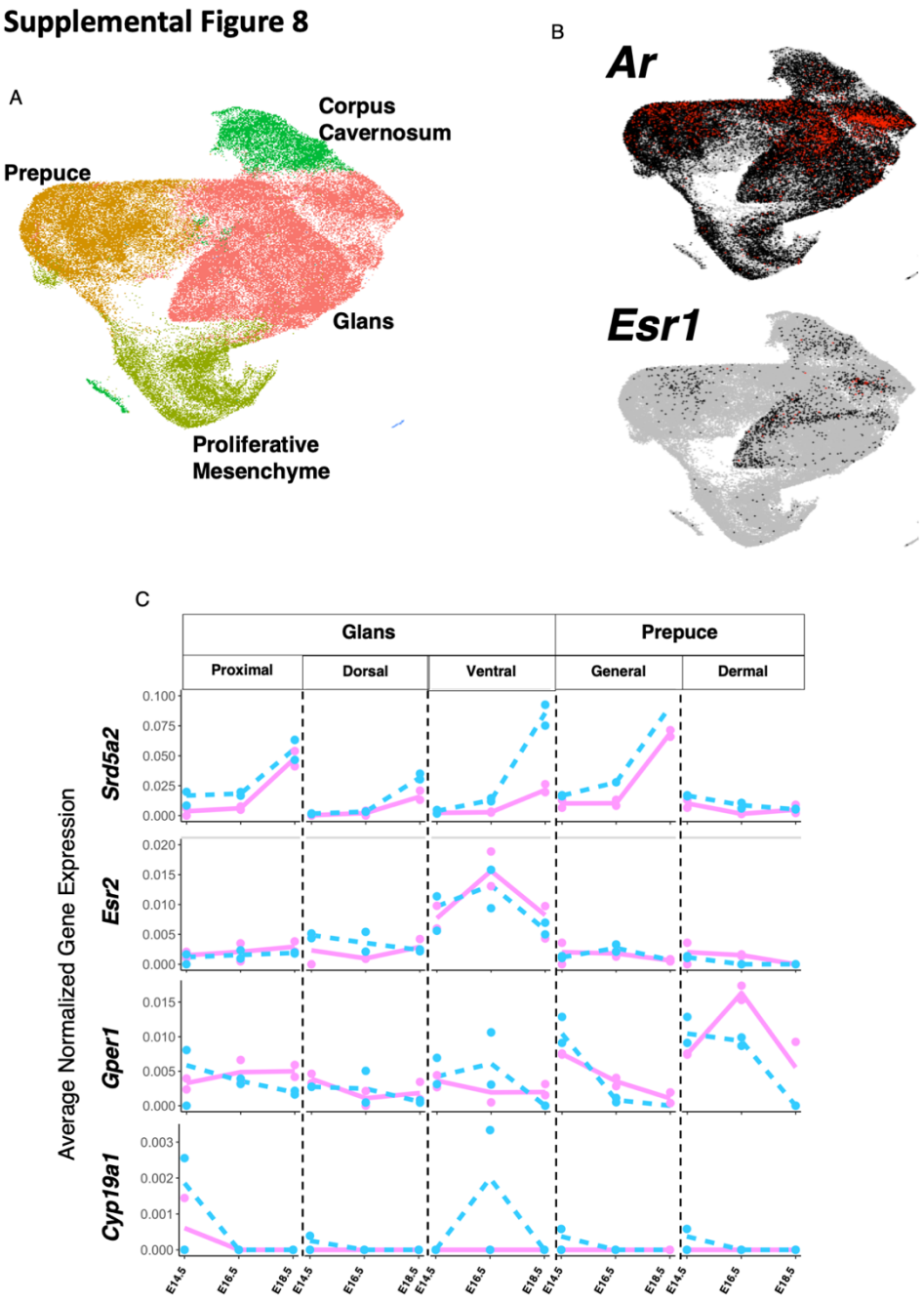


Fig. S8. Sexual dimorphism investigation of the external genitalia. Cell population identification of both the male and female external genitalia in one UMAP (A). Dot color represents cell population identification. Gene expression of *Ar* and *Esr1* in the full external genitalia UMAP (B). Gene expression kinetics of *Srd5a2, Esr2, Gper1,* and *Cyp19a1* in each subpopulation. Blue = male and pink = female.

Supplemental Figure 9

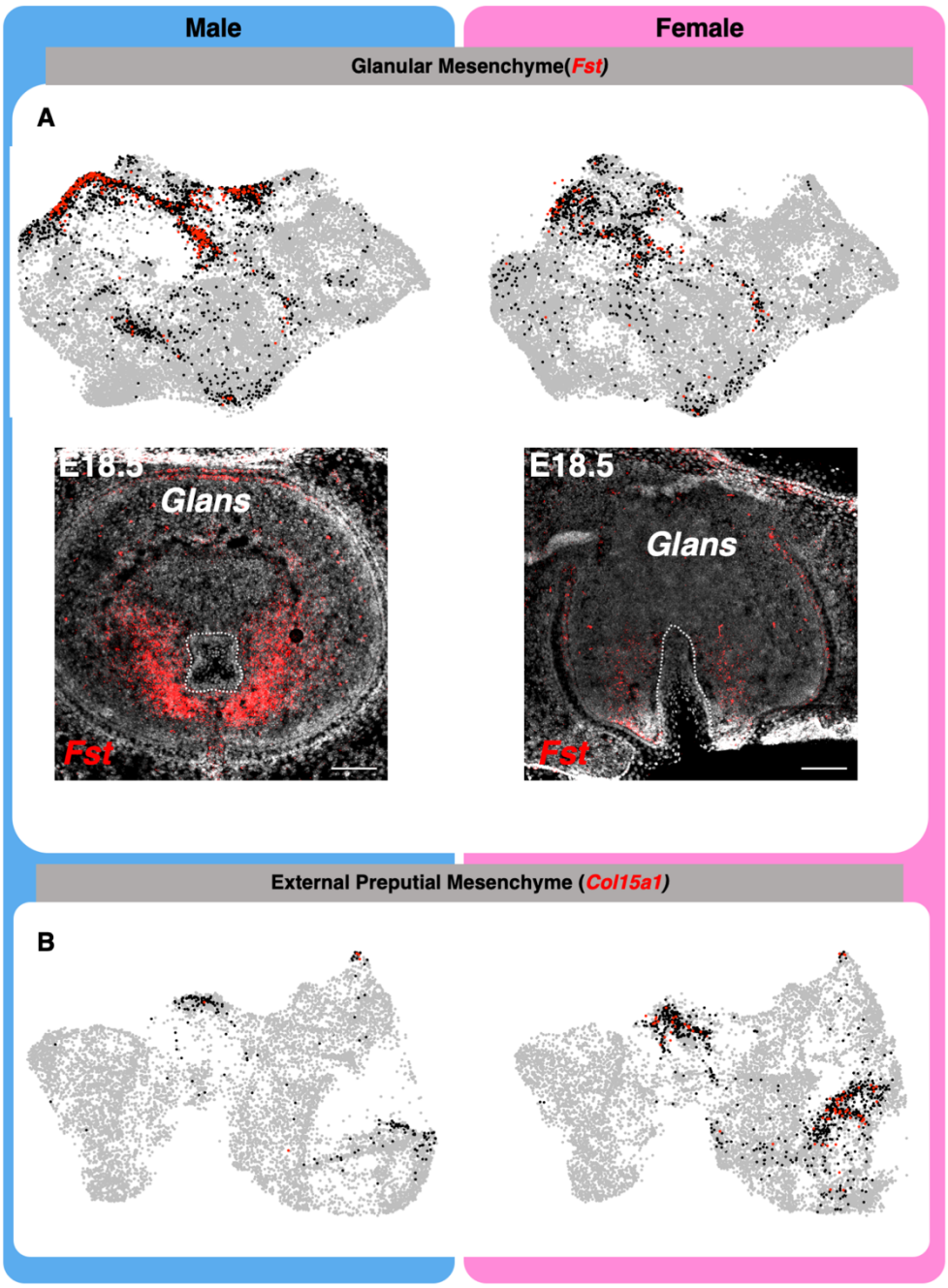


Fig. S9. Identification of other marker genes that identify sex specific cell clusters. *Fst* expression in male and female external genitalia UMAP and validation of *Fst* expression using RNAscope on E18.5 external genitalia (A). Urethra is outlined with white dotted line. *Col15a1* gene expression in male and female UMAPs (B). Scale bars represent 100 μm.

Table S1. Cell and read counts for each sample

Sample	Raw.Reads	Barcode	Mapped	Saturation	Cells	Post-QC Cells	Reads/Cell	Median umi/cell	Median genes/cell
E14.5-Male1	371,073,792	97%	92%	48%	9,635	8,479	38,513	8,601	2,914
E14.5-Male2	385,305,871	96%	91%	59%	5,989	4,814	64,336	10,252	3,210
E14.5-Female1	392,551,274	97%	93%	46%	10,289	9,302	38,153	8,809	2,948
E14.5-Female2	424,685,854	97%	92%	49%	10,283	9,179	41,300	9,211	3,031
E16.5-Male1	351,249,915	97%	82%	49%	8,376	7,318	41,935	8,404	2,871
E16.5-Male2	325,789,676	97%	93%	47%	8,208	7,583	39,692	8,799	2,971
E16.5-Female1	346,782,679	97%	93%	46%	9,217	8,330	37,624	7,887	2,734
E16.5-Female2	323,186,403	97%	92%	47%	8,171	7,434	39,553	8,313	8,171
E18.5-Male1	328,476,663	97%	90%	38%	11,246	7,953	29,208	7,175	2,428
E18.5-Male2	369,414,155	97%	89%	40%	11,815	7,960	31,267	7,781	2,486
E18.5-Female1	333,104,659	96%	90%	42%	11,470	8,912	29,041	5,957	2,200
E18.5-Female2	339,173,048	96%	91%	46%	9,743	8,048	34,812	6,792	2,414

Dataset S1 (separate file). Gene markers for male single cell sequencing data

Dataset S2 (separate file). Gene markers for female single cell sequencing data

Dataset S3 (separate file). Gene markers that differentiate glans subpopulations in the male

Dataset S4 (separate file). Gene markers that differentiate glans subpopulations in the female

Dataset S5 (separate file). Gene markers that differentiate prepuce subpopulations in the male

Dataset S6 (separate file). Gene markers that differentiate prepuce subpopulations in the female

Dataset S7 (separate file). Sexual dimorphic genes in the glans subpopulations

Dataset S8 (separate file). Sexual dimorphic genes in the prepuce subpopulations