Supplemental Information

Direct androgen receptor control of sexually dimorphic gene expression in the mammalian kidney

Lingyun Xiong[1,2,#], Jing Liu[1,#], Seung Yub Han[3], Kari Koppitch[1], Jin-Jin Guo[1], Megan Rommelfanger[2], Zhen Miao[3], Fan Gao[4], Ingileif B. Hallgrimsdottir[5], Lior Pachter[5,6], Junhyong Kim[7], Adam L. MacLean[2], Andrew P. McMahon[1,8,*]

[1]Department of Stem Cell Biology and Regenerative Medicine, Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research, Keck School of Medicine of the University of Southern California, Los Angeles, CA 90089, USA
[2]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA
[3]Graduate Program in Genomics and Computational Biology, Biomedical Graduate Studies, University of Pennsylvania, Philadelphia, PA 19104, USA
[4]Caltech Bioinformatics Resource Center at Beckman Institute, California Institute of Technology, Pasadena, CA 91125, USA
[5]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[6]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA
[7]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[8]Lead Contact
*Correspondence: amcmahon@med.usc.edu (A.P.M.)
#These authors contributed equally to the study

Figure S1

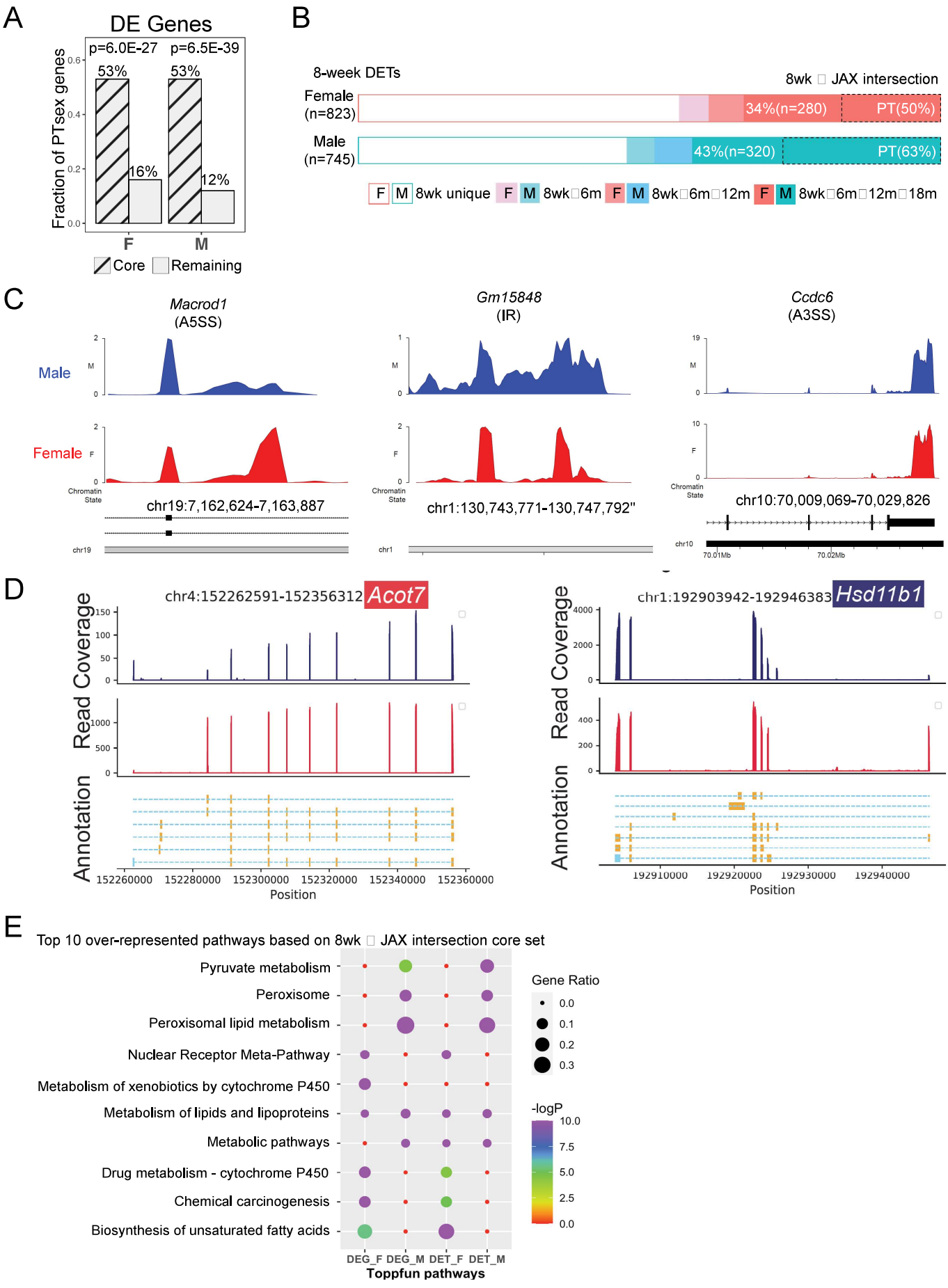

A
DE Genes
p=6.0E-27
p=6.5E-39
53%
53%
16%
12%
Fraction of PTsex genes
F
M
Core
Remaining
B
8-week DETs
8wk ∩ JAX intersection
Female (n=823)
34%(n=280)
PT(50%)
Male (n=745)
43%(n=320)
PT(63%)
F M 8wk unique
F M 8wk∩6m
F M 8wk∩6m∩12m
F M 8wk∩6m∩12m∩18m
C
Macrod1 (A5SS)
Gm15848 (IR)
Ccdc6 (A3SS)
Male
Female
Chromatin State
chr19:7,162,624-7,163,887
chr1:130,743,771-130,747,792"
chr10:70,009,069-70,029,826
D
chr4:152262591-152356312 Acot7
chr1:192903942-192946383 Hsd11b1
Read Coverage
Annotation
Position
E
Top 10 over-represented pathways based on 8wk ∩ JAX intersection core set
Pyruvate metabolism
Peroxisome
Peroxisomal lipid metabolism
Nuclear Receptor Meta-Pathway
Metabolism of xenobiotics by cytochrome P450
Metabolism of lipids and lipoproteins
Metabolic pathways
Drug metabolism - cytochrome P450
Chemical carcinogenesis
Biosynthesis of unsaturated fatty acids
DEG_F DEG_M DET_F DET_M
Toppfun pathways
Gene Ratio
-logP

Figure S1. Transcript-level renal sex differences in adult mice. Related to Fig. 1.
(A) Bar plot shows the fraction of sex-biased genes previously identified in the PT segments from our previous single-cell RNA-seq experiment[19] (PTsex genes) that are covered in the core set of sex-biased genes identified by whole-kidney bulk RNA-seq experiment in the current study. Enrichment analysis was based on Chi-square test.
(B) Stacked bar plot shows the proportion of sex-biased transcripts identified in 8-week C57BL/6 mice that continue to show dimorphic expression in the aging kidney in the diversity outbred (DO) mice (the JAX data[45]) measured at 6, 12, and 18 months. The degree of overlap between 8-week C57BL/6 and JAX data is indicated by the progressive shading of the bars as age advances. The core set of sex-biased transcripts is defined as those that are shared by all four time points. Dashed boxes indicate the number and percentage of core sex-biased transcripts that are shared with PTsex genes.
(C) Read coverages of bulk RNA-seq at representative genomic regions demonstrate alternative splicing events.
(D) Read coverages of bulk RNA-seq at representative genomic regions demonstrate alternative isoform usage.
(E) Dot plot shows the enrichment of ToppFun pathways for both the core set of sex-biased genes and transcripts. Gene count of individual pathways is shown by the relative size of the dot, and the color of the dot indicates the statistical significance.

Figure S2

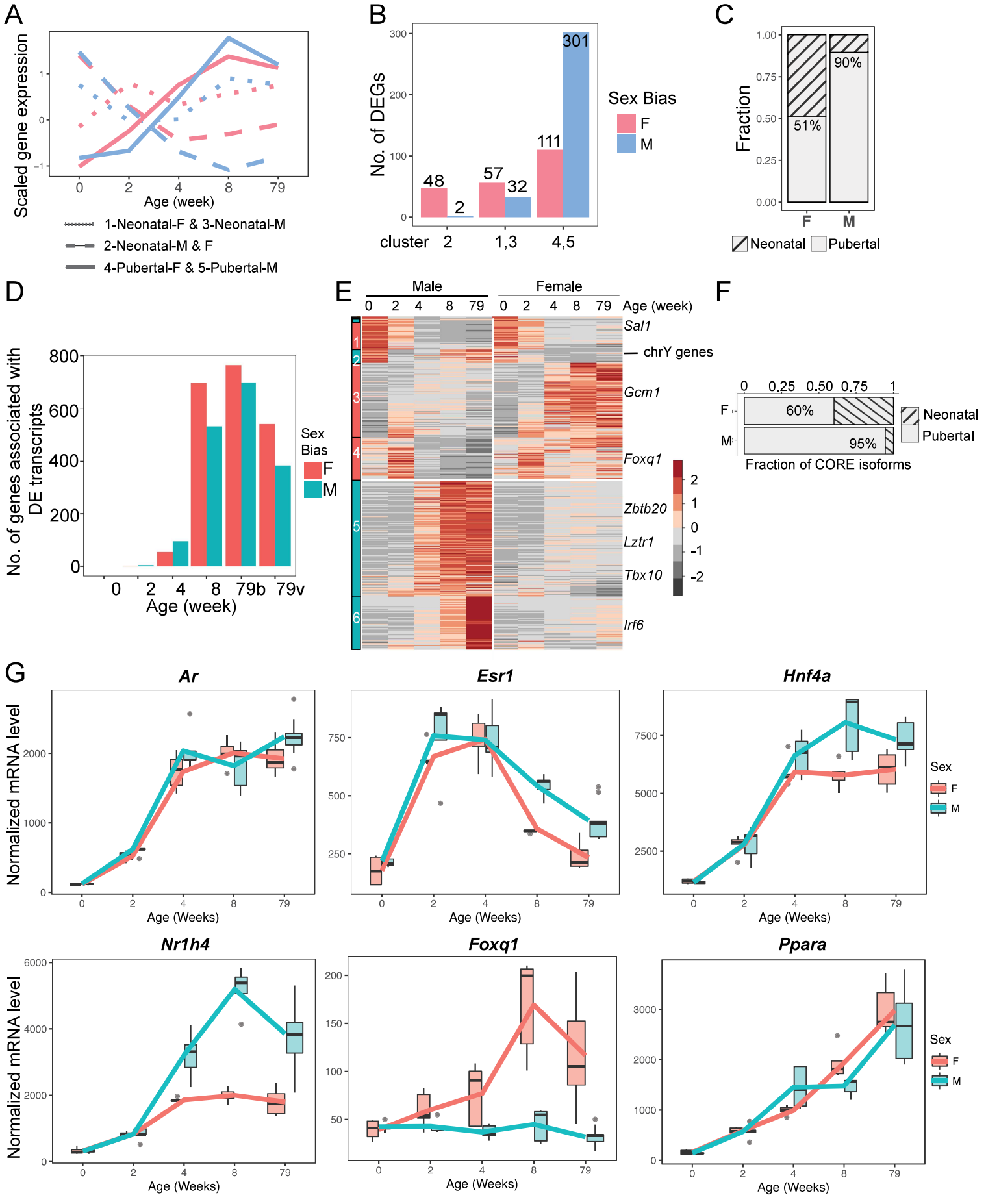

A
Scaled gene expression
Age (week)
1-Neonatal-F & 3-Neonatal-M
2-Neonatal-M & F
4-Pubertal-F & 5-Pubertal-M
B
No. of DEGs
Sex Bias
F
M
48
2
57
32
111
301
cluster 2 1,3 4,5
C
Fraction
90%
51%
F M
Neonatal
Pubertal
D
No. of genes associated with DE transcripts
Age (week)
Sex Bias
F
M
E
Male
Female
Age (week)
Sal1
chrY genes
Gcm1
Foxq1
Zbtb20
Lztr1
Tbx10
Irf6
F
60%
95%
Neonatal
Pubertal
Fraction of CORE isoforms
G
Ar
Esr1
Hnf4a
Nr1h4
Foxq1
Ppara
Normalized mRNA level
Age (Weeks)
Sex
F
M

Figure S2. Dynamical features of male and female renal transcriptomes during postnatal development. Related to Fig. 2.

(A) Scaled average expression of core sex-biased genes over the five timepoints we measured, grouped by clusters shown in Fig. 2D. Together with the heatmap, we can see that neonatal genes are already expressed at high levels at early timepoints (0 and 2 weeks); pubertal genes exhibit increased expression over time (4 weeks onwards).

(B) Bar plot shows the number of core sex-biased genes in each cluster.

(C) Stacked bar plot indicates the fraction of neonatal and pubertal genes in the core set.

(D) Bar plot demonstrates the number of sex-biased transcripts identified at individual timepoints. Bars are colored by sex bias: M-biased genes are in green and F-biased genes in salmon.

(E) Heatmap shows the scaled average expression levels of the core sex-biased transcripts in male and female samples at individual timepoints. Genes were clustered via hierarchical clustering.

(F) Stacked bar plot indicates the fraction of neonatal and pubertal transcripts in the core set.

(G) Normalized read count for *Ar, Esr1* and *Hnf4a* over the five timepoints we measured. Lines indicate average expression levels. Boxes show the variation among samples. Data from male samples are in green and data from female samples are in salmon.

Figure S3

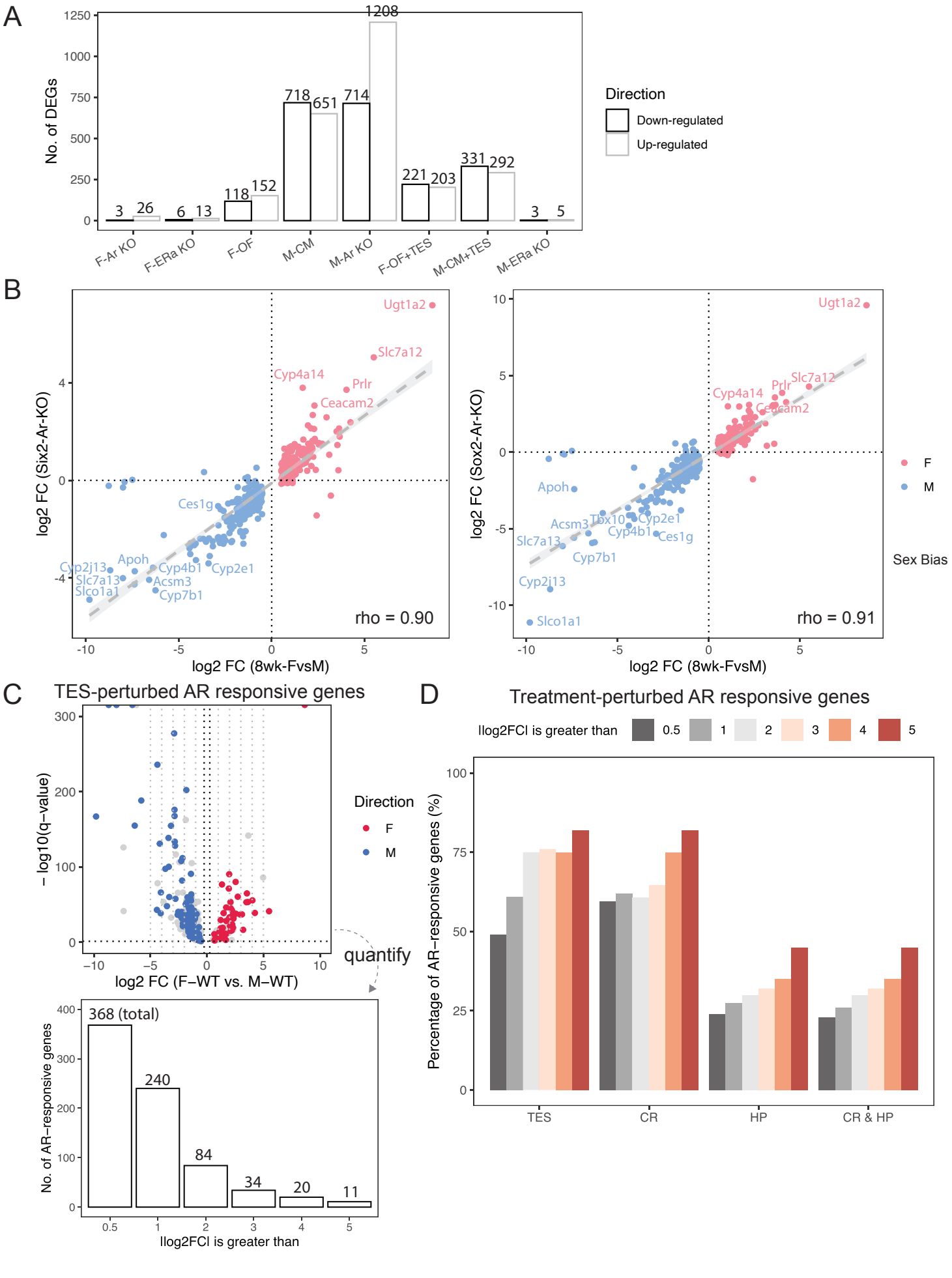

A
No. of DEGs
1250
1000
750
500
250
0
3
26
6
13
118
152
718
651
714
1208
221
203
331
292
3
5
F-Ar KO
F-ERa KO
F-OF
M-CM
M-Ar KO
F-OF+TES
M-CM+TES
M-ERa KO
Direction
Down-regulated
Up-regulated
B
log2 FC (Six2-Ar-KO)
log2 FC (8wk-FvsM)
Ugt1a2
Slc7a12
Cyp4a14
Prlr
Ceacam2
Ces1g
Cyp2j13
Apoh
Cyp4b1
Cyp2e1
Slc7a13
Acsm3
Slco1a1
Cyp7b1
rho = 0.90
log2 FC (Sox2-Ar-KO)
Tbx10
rho = 0.91
F
M
Sex Bias
C
TES-perturbed AR responsive genes
– log10(q–value)
log2 FC (F–WT vs. M–WT)
Direction
F
M
quantify
368 (total)
240
84
34
20
11
No. of AR–responsive genes
|log2FC| is greater than
D
Treatment-perturbed AR responsive genes
|log2FC| is greater than
0.5
1
2
3
4
5
Percentage of AR–responsive genes (%)
TES
CR
HP
CR & HP

Figure S3. AR-mediated signaling plays an important role in renal sex differences. Related to Fig. 3.

(A) Bar plot shows the total number of differentially expressed genes identified in each treatment when compared to controls. Direction of perturbation in gene expression is indicated by the shade of the bar: up-regulated genes are in gray and down-regulated genes are in black.

(B) Scatter plots compare difference in the expression of core sex-biased genes in either of the two types of AR knockout experiments (nephron-specific removal [Six2-Ar-KO] and whole-body removal [Sox2-Ar-KO]) versus male-female differences at 8 weeks. Each dot represents a sex-biased gene, where sex bias is indicated by colors. Changes in gene expression (KO vs. WT or F vs. M) are summarized by log2 fold change (FC), which we used to evaluate the overall concordance of the two experiments by Spearman correlation.

(C) Volcano plot highlights AR-responsive genes that were perturbed by TES treatment. Bar plot quantifies the number of genes using different effect size cutoffs.

(D) Grouped bar plot shows percentage of AR-responsive genes that were perturbed by different treatments, using the gene sets identified in (C) as reference.

Figure S4

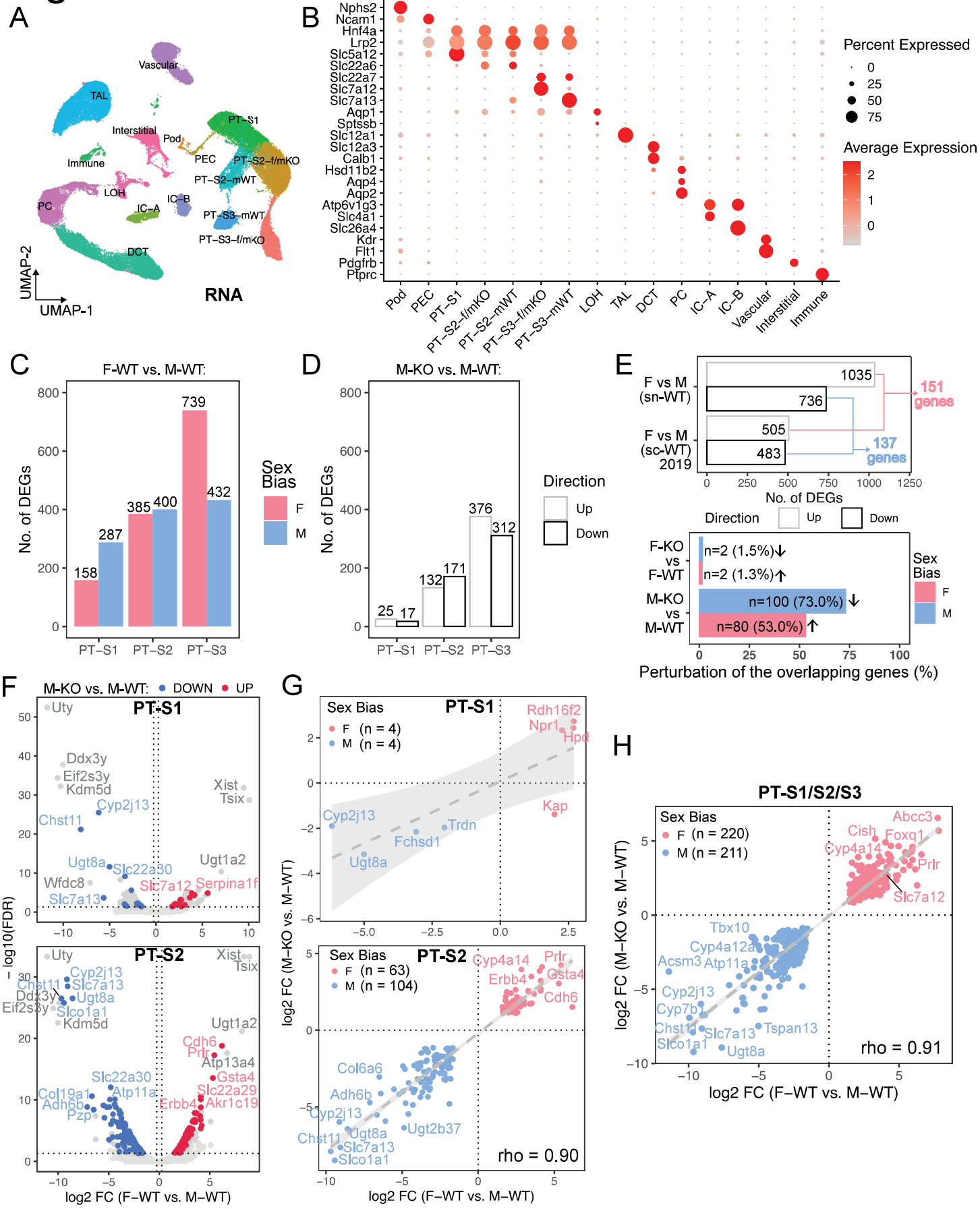

A
Vascular
TAL
Interstitial
Pod
PT–S1
Immune
PEC
PT–S2–f/mKO
PT–S2–mWT
LOH
IC–B
PC
IC–A
PT–S3–mWT
PT–S3–f/mKO
DCT
UMAP-2
UMAP-1
RNA
B
Nphs2
Ncam1
Hnf4a
Lrp2
Slc5a12
Slc22a6
Slc22a7
Slc7a12
Slc7a13
Aqp1
Sptssb
Slc12a1
Slc12a3
Calb1
Hsd11b2
Aqp4
Aqp2
Atp6v1g3
Slc4a1
Slc26a4
Kdr
Flt1
Pdgfrb
Ptprc
Pod
PEC
PT–S1
PT–S2–f/mKO
PT–S2–mWT
PT–S3–f/mKO
PT–S3–mWT
LOH
TAL
DCT
PC
IC–A
IC–B
Vascular
Interstitial
Immune
Percent Expressed
0
25
50
75
Average Expression
2
1
0
C
F-WT vs. M-WT:
No. of DEGs
800
600
400
200
0
158
287
385
400
739
432
PT–S1
PT–S2
PT–S3
Sex Bias
F
M
D
M-KO vs. M-WT:
No. of DEGs
800
600
400
200
0
25
17
132
171
376
312
PT–S1
PT–S2
PT–S3
Direction
Up
Down
E
F vs M (sn-WT)
F vs M (sc-WT) 2019
1035
736
505
483
151 genes
137 genes
0
250
500
750
1000
1250
No. of DEGs
Direction
Up
Down
F-KO vs F-WT
n=2 (1.5%)↓
n=2 (1.3%)↑
M-KO vs M-WT
n=100 (73.0%) ↓
n=80 (53.0%) ↑
Sex Bias
F
M
0
25
50
75
100
Perturbation of the overlapping genes (%)
F
M-KO vs. M-WT: DOWN UP
PT-S1
Uty
Ddx3y
Eif2s3y
Kdm5d
Cyp2j13
Chst11
Xist
Tsix
Ugt8a
Slc22a30
Ugt1a2
Wfdc8
Slc7a12
Serpina1f
Slc7a13
− log10(FDR)
PT-S2
Uty
Cyp2j13
Chst11
Slc7a13
Ddx3y
Ugt8a
Eif2s3y
Slco1a1
Kdm5d
Xist
Tsix
Ugt1a2
Cdh6
Prlr
Atp13a4
Slc22a30
Gsta4
Col19a1
Atp11a
Slc22a29
Adh6b
Erbb4
Akr1c19
Pzp
log2 FC (F–WT vs. M–WT)
G
Sex Bias
PT-S1
F (n = 4)
M (n = 4)
Rdh16f2
Npr1
Hpd
Kap
Cyp2j13
Trdn
Fchsd1
Ugt8a
Sex Bias
PT-S2
F (n = 63)
M (n = 104)
Cyp4a14
Prlr
Erbb4
Gsta4
Cdh6
Col6a6
Adh6b
Cyp2j13
Chst11
Ugt8a
Ugt2b37
Slc7a13
Slco1a1
rho = 0.90
log2 FC (M–KO vs. M–WT)
log2 FC (F–WT vs. M–WT)
H
PT-S1/S2/S3
Sex Bias
F (n = 220)
M (n = 211)
Abcc3
Cish
Foxq1
Cyp4a14
Prlr
Slc7a12
Tbx10
Cyp4a12a
Acsm3
Atp11a
Cyp2j13
Cyp7b
Chst11
Slc7a13
Tspan13
Slco1a1
Ugt8a
rho = 0.91
log2 FC (M–KO vs. M–WT)
log2 FC (F–WT vs. M–WT)

Figure S4. Nephron-specific AR removal reverses dimorphic gene expression in males. Related to Fig. 4.

(A) UMAP plot shows clustering outcome the RNA modality of the multimodal data. Cluster nomenclature and color scheme are consistent with Fig. 4B.

(B) Dot plot of established marker expression for all cell populations.

(C) Bar plot shows the number of sex-biased genes identified in each PT segment.

(D) Bar plot shows the number of differentially expressed genes comparing M-KO to M-WT in each PT segment.

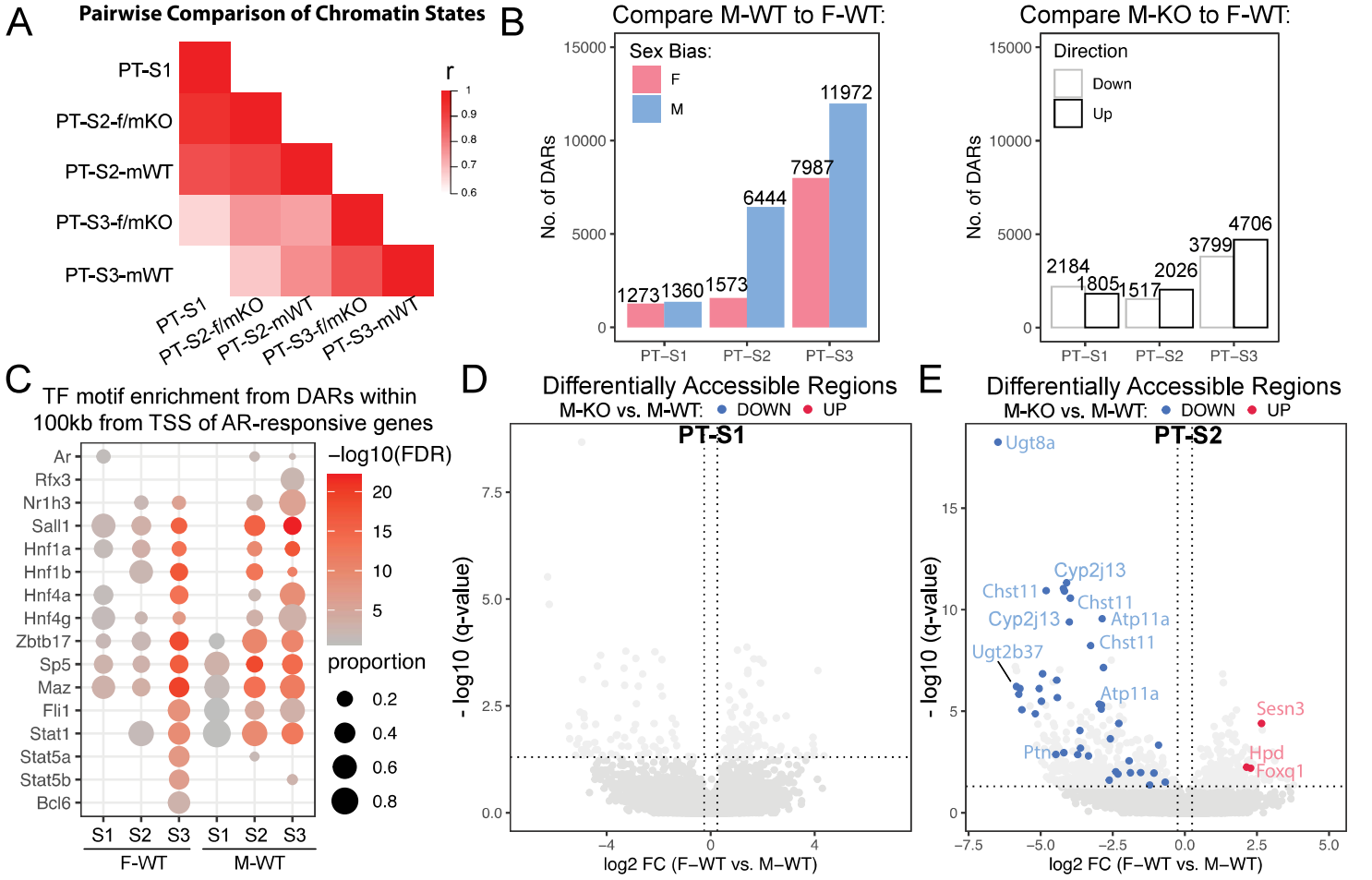
(E) Top bar plot shows the number of multimodal sex-biased genes and those identified from our previous single-cell RNA-seq experiment, with the number of overlapping sex-biased genes indicated on the right. Bottom bar plot shows the percentage of the overlapping sex-biased genes that were perturbed by nephron-specific AR removal. The number of genes that are perturbed in each treatment and the corresponding percentages are listed. Arrows indicate the direction of perturbation in gene expression as compared to controls.

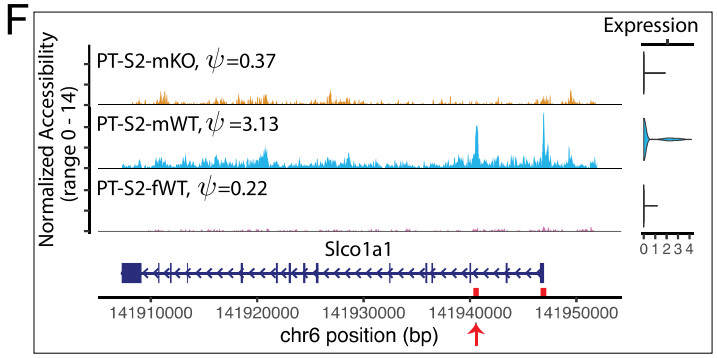
(F) Volcano plots show multimodal sex-biased genes identified in PT-S1 and PT-S2 separately, where genes that are perturbed by AR removal in male are highlighted: comparing M-KO to M-WT, up-regulated genes are in red and down-regulated genes are in blue.

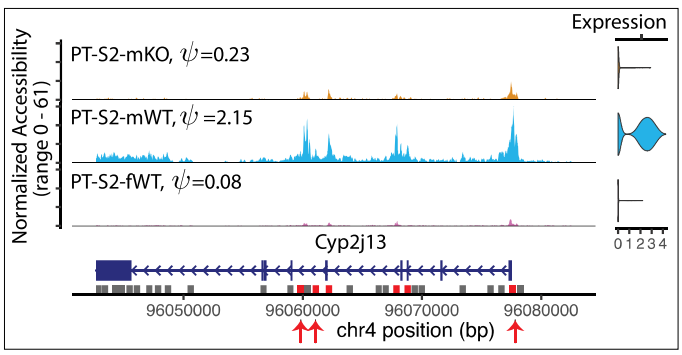
(G) Scatter plots contrast the effect of nephron-specific AR removal in male to the observed sex biases, among multimodal AR-responsive genes identified in PT-S1 and PT-S2, as shown in (E). Each dot represents a gene, where sex bias differs by color. Changes in gene expression (F-WT vs. M-WT or M-KO vs. M-WT) are summarized by log2 FC. Spearman's rank correlation coefficient was calculated to assess the overall concordance of the two comparisons.

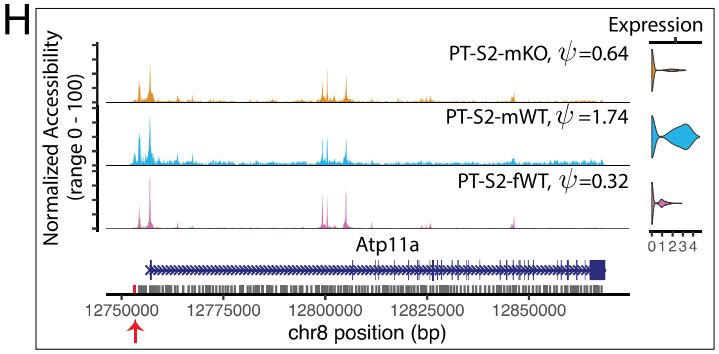
(H) Scatter plot compares the effect of nephron-specific AR removal in male to the observed sex biases, among multimodal AR-responsive genes identified in PT-S1, PT-S2, and PT-S3 combined. Each dot represents a gene, where sex bias differs by color. Changes in gene expression (F-WT vs. M-WT or M-KO vs. M-WT) are summarized by log2 FC. For genes that were perturbed in more than one segments, the maximum effect sizes are shown.

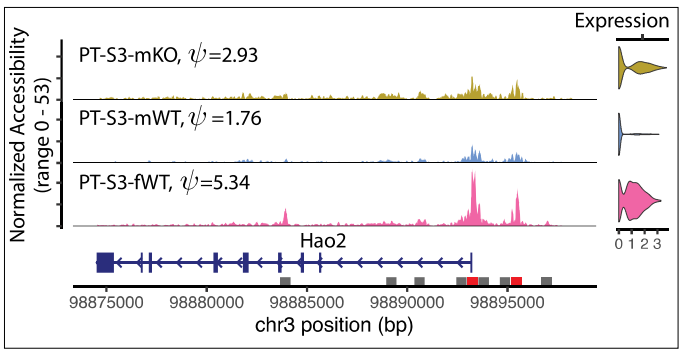
Figure S5

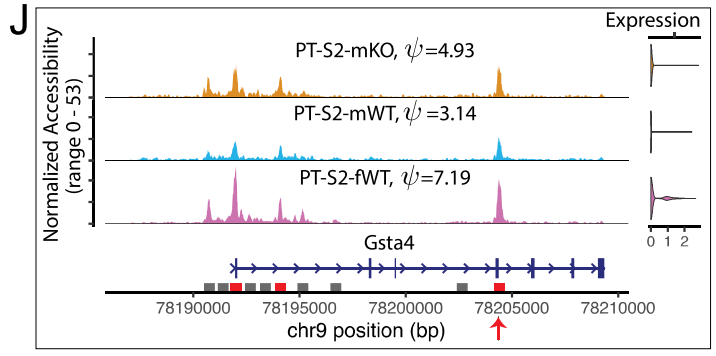

A
Pairwise Comparison of Chromatin States
PT-S1
PT-S2-f/mKO
PT-S2-mWT
PT-S3-f/mKO
PT-S3-mWT
r
B
Compare M-WT to F-WT:
Sex Bias:
F
M
No. of DARs
1273
1360
1573
6444
7987
11972
PT-S1
PT-S2
PT-S3
Compare M-KO to F-WT:
Direction
Down
Up
2184
1805
1517
2026
3799
4706
C
TF motif enrichment from DARs within 100kb from TSS of AR-responsive genes
Ar
Rfx3
Nr1h3
Sall1
Hnf1a
Hnf1b
Hnf4a
Hnf4g
Zbtb17
Sp5
Maz
Fli1
Stat1
Stat5a
Stat5b
Bcl6
–log10(FDR)
proportion
S1
S2
S3
F-WT
M-WT
D
Differentially Accessible Regions
M-KO vs. M-WT: DOWN UP
PT-S1
- log10 (q-value)
log2 FC (F–WT vs. M–WT)
E
Differentially Accessible Regions
M-KO vs. M-WT: DOWN UP
PT-S2
Ugt8a
Cyp2j13
Chst11
Atp11a
Ugt2b37
Sesn3
Ptn
Hpd
Foxq1
- log10 (q-value)
log2 FC (F–WT vs. M–WT)


F
Expression
Normalized Accessibility (range 0 - 14)
PT-S2-mKO, ψ=0.37
PT-S2-mWT, ψ =3.13
PT-S2-fWT, ψ=0.22
Slco1a1
141910000
141920000
141930000
141940000
141950000
chr6 position (bp)


G
Expression
Normalized Accessibility (range 0 - 61)
PT-S2-mKO, ψ=0.23
PT-S2-mWT, ψ =2.15
PT-S2-fWT, ψ=0.08
Cyp2j13
96050000
96060000
96070000
96080000
chr4 position (bp)


H
Expression
Normalized Accessibility (range 0 - 100)
PT-S2-mKO, ψ=0.64
PT-S2-mWT, ψ =1.74
PT-S2-fWT, ψ=0.32
Atp11a
12750000
12775000
12800000
12825000
12850000
chr8 position (bp)


I
Expression
Normalized Accessibility (range 0 - 53)
PT-S3-mKO, ψ=2.93
PT-S3-mWT, ψ =1.76
PT-S3-fWT, ψ=5.34
Hao2
98875000
98880000
98885000
98890000
98895000
chr3 position (bp)


J
Expression
Normalized Accessibility (range 0 - 53)
PT-S2-mKO, ψ=4.93
PT-S2-mWT, ψ =3.14
PT-S2-fWT, ψ=7.19
Gsta4
78190000
78195000
78200000
78205000
78210000
chr9 position (bp)

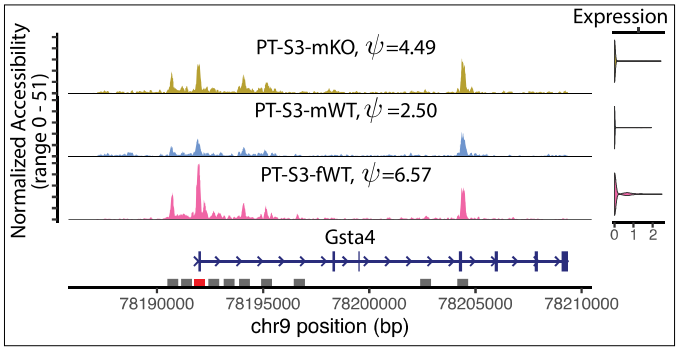

K
Expression
Normalized Accessibility (range 0 - 51)
PT-S3-mKO, ψ=4.49
PT-S3-mWT, ψ =2.50
PT-S3-fWT, ψ=6.57
Gsta4
78190000
78195000
78200000
78205000
78210000
chr9 position (bp)

Figure S5. Multimodal data reveals putative AR response elements near sex-biased genes. Related to Fig. 5.

(A) Heatmap shows pairwise comparison of chromatin state between individual PT clusters. Scatter plot on the right exemplifies the correlation of normalized average peak counts for all PT marker peaks between PT-S3-mWT and PT-S3-f/mKO cluster.

(B) Bar plot shows number of significantly open regions for each condition of each pair-wise comparison (absolute Log2FC > 0.25, adjusted p-value < 0.05).

(C) Dot plot summarizes TF motif enrichment in the DARs within 100KB to TSS of the AR-responsive sex-biased genes in M-WT compared to F-WT PT segments.

(D) Volcano plot shows DARs within 100KB of sex-biased genes in PT-S1. 1,360 peaks were differentially open in male (left) and 1,273 peaks were differentially open in female (right). We colored F-biased peaks that are preferentially open in M-KO in red, and M-biased peaks that are preferentially closed in M-KO in blue, if any. Each dot represents a 500-bp region, where the nearest gene is annotated.

(E) Volcano plot shows DARs within 100KB of sex-biased genes in PT-S2. 6,444 peaks were differentially open in male (left) and 1,573 peaks were differentially open in female (right). We colored F-biased peaks that are preferentially open in M-KO in red, and M-biased peaks that are preferentially closed in M-KO in blue. Each dot represents a 500-bp region, where the nearest gene is annotated.

(F-K) Coverage plots of top sex-biased genes in M-WT, F-WT, and M-KO PT-S2/3: *Slco1a1, Cyp2j13, and Atp11a* (M-biased; F-H); *Hao2, Gsta4* (F-biased; I-K). Gene expression is also shown by violin plots adjacent to the tracks, with gene accessibility score ψ listed above each track. All peaks called in the region are shown in gray boxes, where DARs are highlighted in red. Peaks with potential AR binding site are indicated by red arrows.

Figure S6

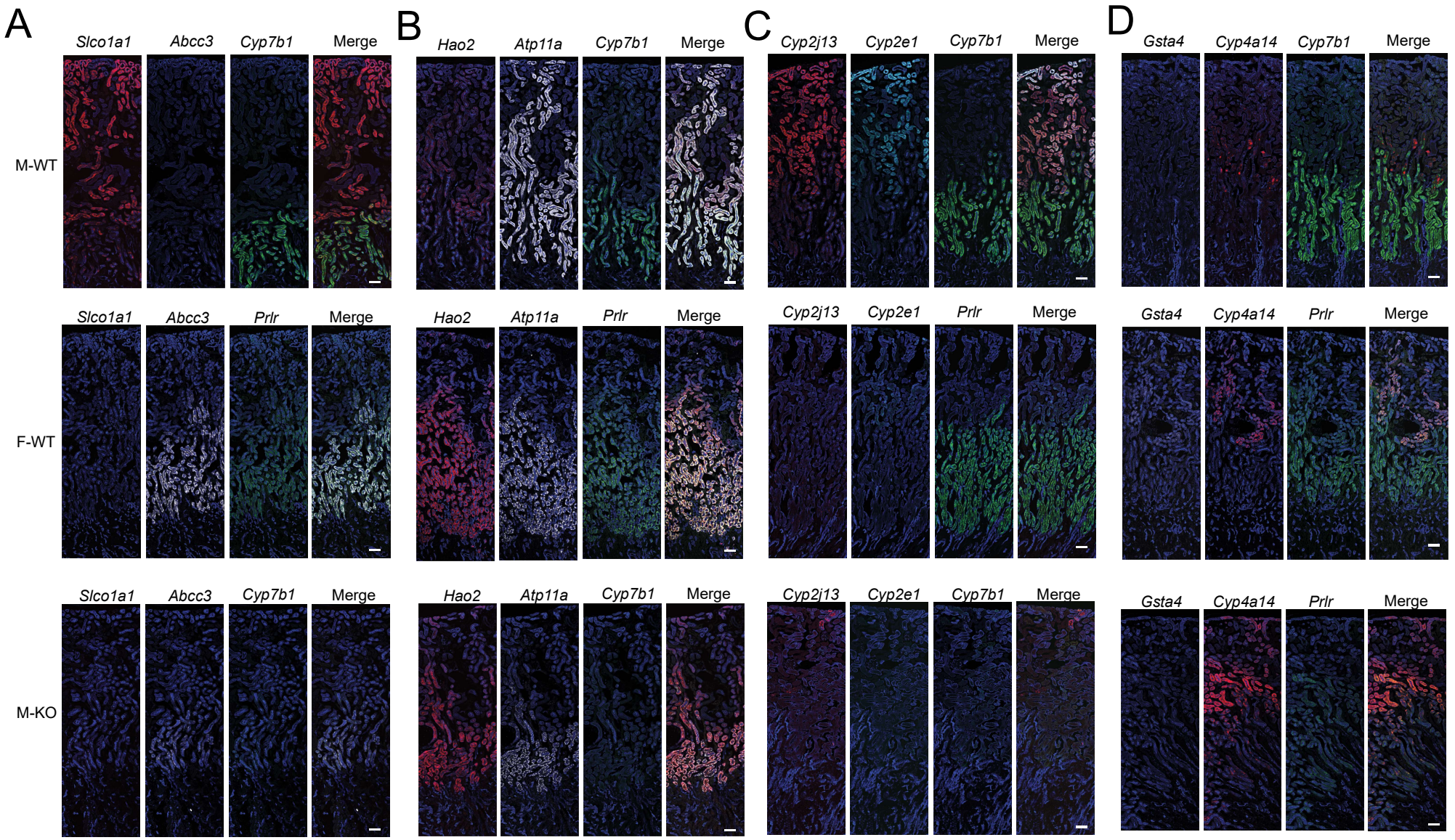

A
Slco1a1
Abcc3
Cyp7b1
Merge
M-WT
Slco1a1
Abcc3
Prlr
Merge
F-WT
Slco1a1
Abcc3
Cyp7b1
Merge
M-KO
B
Hao2
Atp11a
Cyp7b1
Merge
Hao2
Atp11a
Prlr
Merge
Hao2
Atp11a
Cyp7b1
Merge
C
Cyp2j13
Cyp2e1
Cyp7b1
Merge
Cyp2j13
Cyp2e1
Prlr
Merge
Cyp2j13
Cyp2e1
Cyp7b1
Merge
D
Gsta4
Cyp4a14
Cyp7b1
Merge
Gsta4
Cyp4a14
Prlr
Merge
Gsta4
Cyp4a14
Prlr
Merge

Figure S6. Fluorescent RNA in situ hybridization by RNAscope validates dimorphic gene expression in proximal tubule. Related to Fig. 6.

RNAscope assay directly visualized the expression levels of sex-biased genes in M-WT, F-WT, and M-KO across the renal cortex. Scale bars = 100 μm.

(A) *Slco1a1* (red)*, Abcc3* (white)*, Cyp7b1* (green) co-stained with antibody against Aqp1 (blue).

(B) *Hao2* (red)*, Atp11a* (white)*, Cyp7b1* (green) co-stained with antibody against Aqp1 (blue).

(C) *Cyp2j13* (red)*, Cyp2e1* (Cyan)*, Cyp7b1* (green) co-stained with antibody against Aqp1 (blue).

(D) *Gsta4* (white)*, Cyp4a14* (red)*, Cyp7b1* (green) co-stained with antibody against Aqp1 (blue).

Figure S7

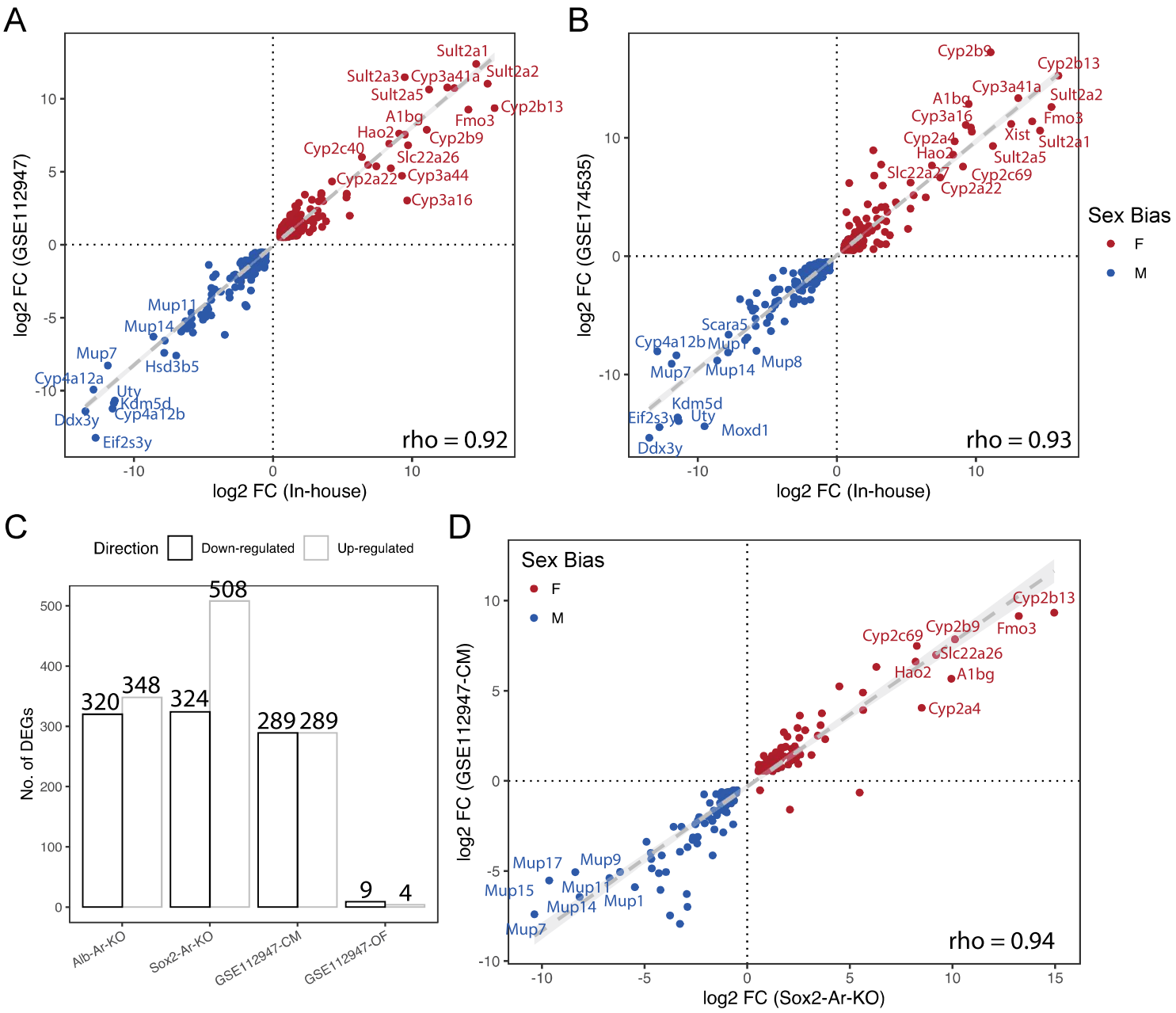

A
log2 FC (GSE112947)
log2 FC (In-house)
rho = 0.92
Sult2a1
Sult2a3
Cyp3a41a
Sult2a2
Sult2a5
Cyp2b13
A1bg
Fmo3
Hao2
Cyp2b9
Cyp2c40
Slc22a26
Cyp2a22
Cyp3a44
Cyp3a16
Mup11
Mup14
Mup7
Hsd3b5
Cyp4a12a
Uty
Kdm5d
Ddx3y
Cyp4a12b
Eif2s3y
B
log2 FC (GSE174535)
log2 FC (In-house)
rho = 0.93
Sex Bias
F
M
Cyp2b9
Cyp2b13
Cyp3a41a
A1bg
Sult2a2
Cyp3a16
Fmo3
Cyp2a4
Xist
Sult2a1
Hao2
Sult2a5
Slc22a26
Cyp2c69
Cyp2a22
Scara5
Cyp4a12b
Mup1
Mup7
Mup14
Mup8
Kdm5d
Eif2s3y
Uty
Moxd1
Ddx3y
C
Direction
Down-regulated
Up-regulated
No. of DEGs
320
348
324
508
289
289
9
4
Alb-Ar-KO
Sox2-Ar-KO
GSE112947-CM
GSE112947-OF
D
Sex Bias
F
M
log2 FC (GSE112947-CM)
log2 FC (Sox2-Ar-KO)
rho = 0.94
Cyp2b13
Cyp2c69
Cyp2b9
Fmo3
Slc22a26
Hao2
A1bg
Cyp2a4
Mup17
Mup9
Mup15
Mup11
Mup14
Mup1
Mup7

Figure S7. Comparative study on kidney chromatin and hepatic gene expression. Related to Fig. 7.

(A-B) Scatter plot compares changes in the expression of liver sex-biased genes that are in common between the in-house list and those identified from GSE112947 (A) and GSE174535 (B).

(C) Bar plot compares the number of differentially expressed genes identified in different treatment experiments.

(D) Scatter plot compares changes in the expression of liver sex-biased genes that are commonly perturbed by whole-body AR removal and castration in male mice.