Supplemental information titles and legends:

ALDH1A1^{+} cells are endometrial progenitors contributing to regeneration and endometriosis

Suni Tang[1,2]; Anna Catherine Unser[1,2]; Peixin Jiang[1]; Sydney E. Parks[1,2]; Genesis J. Herrera[1,2]; Ting Geng[1,2]; Linda Alpuing Radilla[3]; Brooke A. Thigpen[3], Xiaoming Guan[3]; Diana Monsivais[1,2]

1. Department of Pathology & Immunology, Baylor College of Medicine, Houston, Texas, USA
2. Center for Drug Discovery, Baylor College of Medicine, Houston, Texas, USA
3. Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, Texas, USA

Table S1. Differentially expressed genes between clusters 4 and 25 versus other epithelial cell clusters in the scRNAseq dataset.

Table S2. Differentially expressed genes and gene enrichment analysis of ALDH^{HI} vs. ALDH^{LO} mouse endometrial epithelial organoids.

Table S3. Differentially expressed genes and gene enrichment analysis of ALDH^{HI} vs. ALDH^{LO} human endometrial epithelial organoids.

Table S4. Transcriptomic analysis of ALDH^{HI} eutopic vs. ALDH^{HI} ectopic human endometrial epithelial organoids.

Table S5. Primer sequences and antibody information.

Table S6. Patient sample information.

Figure S1. Estrogen and progesterone impact the expression and localization of ALDH1A1 in the WT mouse uterus. A) Description of the experimental scheme used to ovariectomize and administer hormonal treatments to 6-8-week-old WT mice. B-E') ALDH1A1 immunohistochemistry in the uterine cross-sections of ovariectomized mice treated with vehicle (B-B'), 1mg P4 (C-C'), 50ng E2 (D-D'), or 1mg P4 + 50ng E2 (E-E'). ALDH1A1 immunohistochemistry was also performed in the uterine cross-sections of mice that were ovariectomized and implanted with a placebo or E2 pellet for 90 days (F-F'). H-I) *Aldh1a1* gene expression was quantified in the uterine tissues of the ovariectomized mice treated with Vehicle, P4, E2 or E2 + P4 (H). *Aldh1a1* was also quantified in the uterus of 6-8-week-old WT mice collected at different times during the estrous cycle. Experiments were repeated in more than three mice per group. Data in H-I are displayed as mean ± SEM analyzed by a One-Way ANOVA test with a Tukey's post-hoc test. *, $p<0.05$; **, $P<0.01$; ***, $p<0.001$.

Figure S2. Identification and characterization of epithelial stem cells in the adult WT mouse uterus at estrus and diestrus. A) Dot plots showing the expression of Aldh1a1, Lgr5, and Axin2 across the epithelial cell subclusters from the estrus and diestrus phases. B) Estrus and diestrus phase expression of genes identifying epithelial, luminal, and epithelial stem cell clusters across each cluster of epithelial cells. C) Dual feature plot showing the overlapping and unique expression patterns of *Aldh1a1/Lgr5*, *Aldh1a1/Axin2*, and *Lgr5/Axin2* in the epithelial cell clusters in diestrus. D) Heatmap (row for gene, column for individual cell) comparing epithelial cell subclusters from our dataset to DEGs identified from EpSC DEGs of cluster 12 (Padilla-Banks et al., 2023). E-F) UMAP of epithelial cell subclusters from mouse samples obtained during the estrus (E) or diestrus phases (F). G-M) Lineage trajectories using pseudotime analysis of the epithelial cell types in estrus (G,H) and diestrus (I-M).

Figure S3. A-B) Lineage tracing of *Aldh1a1* reporter mice starting at postnatal days 7 and 8 using 4-hydroxytamoxifen (0.5µg/g B.W.) and tracing for one (PND7 to PND8) or 6 (PND8 to PND14) days. C-D) RFP immunohistochemistry in uterine cross-sections (C-C') and longitudinal sections (D-D') of mice traced from PND7 to PND8. Red arrows indicate positive cells. E-F') RFP immunohistochemistry in uterine cross-sections (E-E') or longitudinal (F-F') sections of mice traced from PND8 to PND14. G-H) Quantification of total RFP positive cells in the cross sections of mice traced from PND14 to PND28 or from PND14 to PND56. Data are presented as mean ± SEM analyzed. *PND,* postnatal day; *4-OHT*, 4-hydroxytamoxifen.

Figure S4. Analysis of uterine tissues with ALDH1A1 cell ablation. Immunohistochemistry of the uterus of mice collected 4 days after DT administration in control (A, C, E) was performed using antibodies for FOXA2 (A-B'), (C-D'), (E-F'). G) Clustering analysis of nuclear hormone receptors in the mouse epithelial cells established from $ALDH^{HI}$ or $ALDH^{LO}$ organoids.

Figure S5. Comparison of $ALDH^{HI}$ and $ALDH^{LO}$ formation in the organoids from human eutopic endometrial epithelium. A-D) Images (A-D) and quantification (E) comparing the growth of $ALDH^{LO}$ (A,C) and $ALDH^{HI}$ (B,D) organoids at passage 11 (A,D) or passage 16 (B-D). F-H) Phase contrast confocal images and quantification of an organoid formation assay comparing $ALDH^{HI}$ and $ALDH^{LO}$ cells. I-J) $ALDH^{LO}$ (I) and $ALDH^{HI}$ (J) eutopic epithelial organoids were established and expanded over several weeks. Data in E,G are displayed as mean ± SD analyzed by a two-tailed t-test. *, $p<0.05$; **, $P<0.01$; ***, $p<0.001$.

Figure S6. Gene expression pathways are differentially regulated between $ALDH^{HI}$ and $ALDH^{LO}$ cells in the eutopic and ectopic endometrial tissues. A-B) Clustering of genes associated with proliferation (A) and nuclear hormone receptors (B) in the $ALDH^{HI}$ eutopic, $ALDH^{LO}$ eutopic, and $ALDH^{HI}$ ectopic organoids. C) Heat map displaying relative levels of keratin-related genes in the organoids from $ALDH^{HI}$ eutopic, $ALDH^{LO}$ eutopic and $ALDH^{HI}$ ectopic cells. D) Gene ontology chord plot showing the gene-network interactions of differentially expressed genes in $ALDH^{HI}$ eutopic vs $ALDH^{LO}$ organoids. E-F) Clustering of the top 100 differentially expressed genes in $ALDH^{HI}$ eutopic vs $ALDH^{LO}$ organoids (E) or $ALDH^{HI}$ eutopic vs $ALDH^{HI}$ ectopic organoids

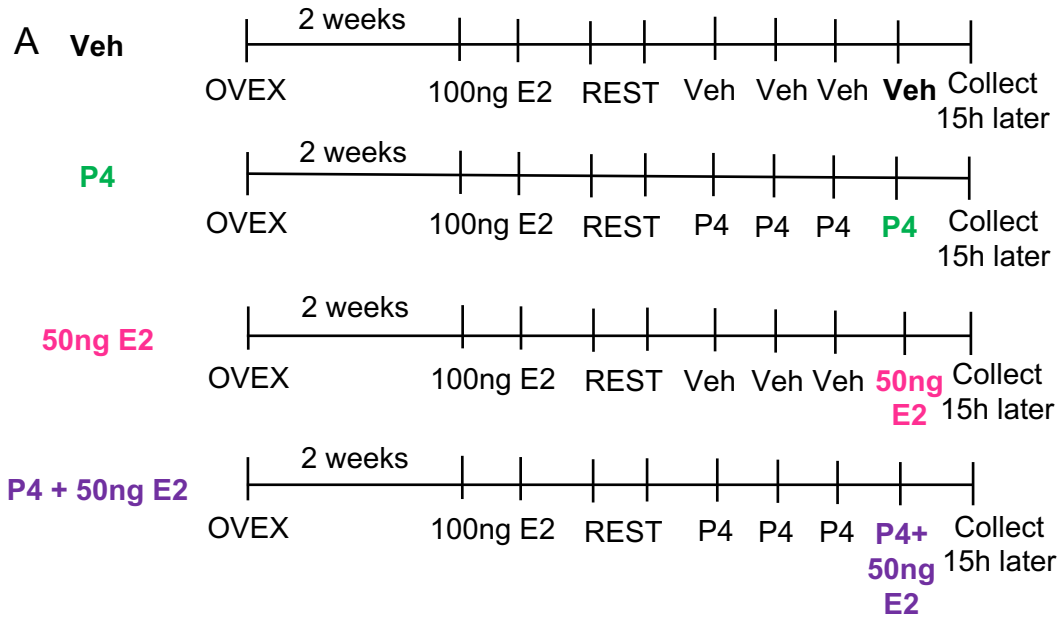
(F). G) Gene enrichment analysis of differentially expressed genes in $ALDH^{HI}$ eutopic vs $ALDH^{LO}$ eutopic organoids. H) Volcano plot displaying the total number of differentially expressed genes when comparing $ALDH^{HI}$ eutopic vs $ALDH^{HI}$ ectopic organoids.

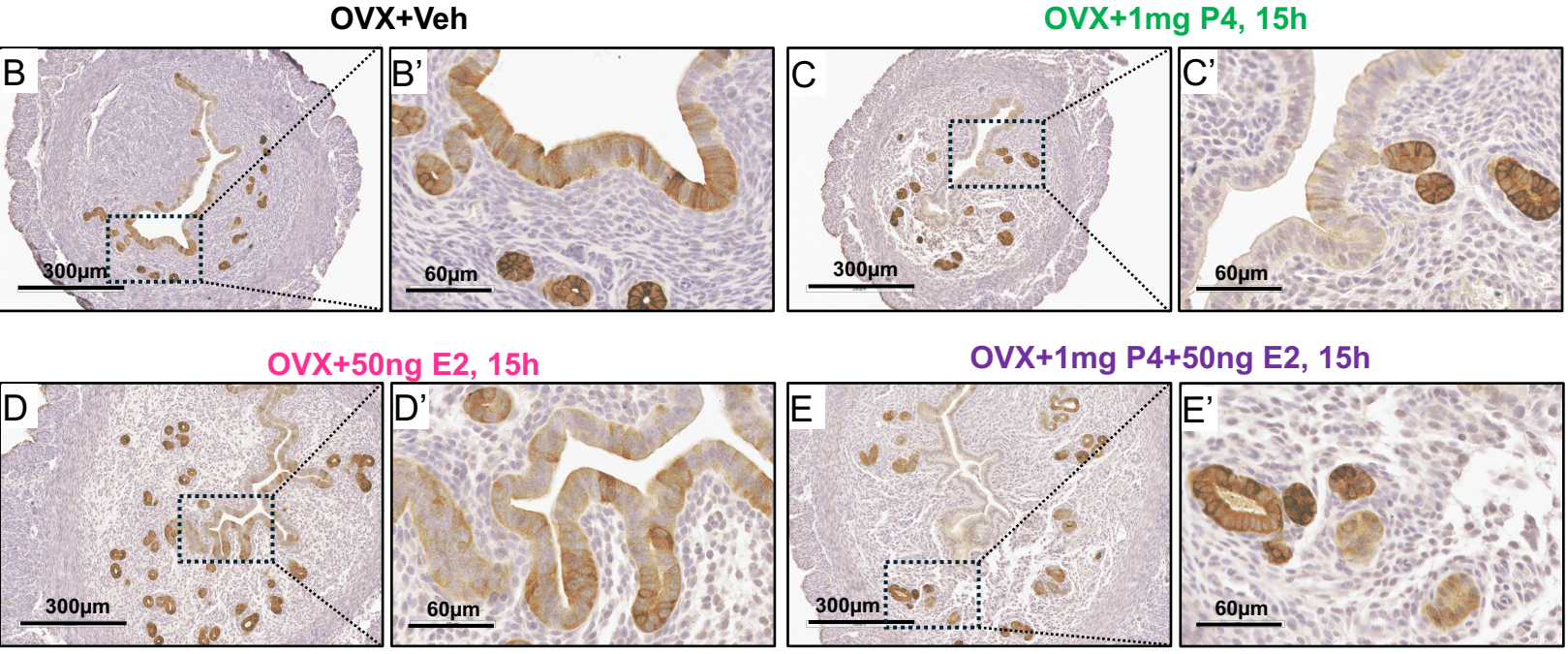
Figure S7. Genotyping strategy for the various mouse lines used in the study. A-C) Gel electrophoresis results showing the PCR results for the Aldh1a1cre/ERT2 allele (A), for the Ai9/Ai9 tdTomato reporter (B) and for the DTR knockin and WT alleles (C).

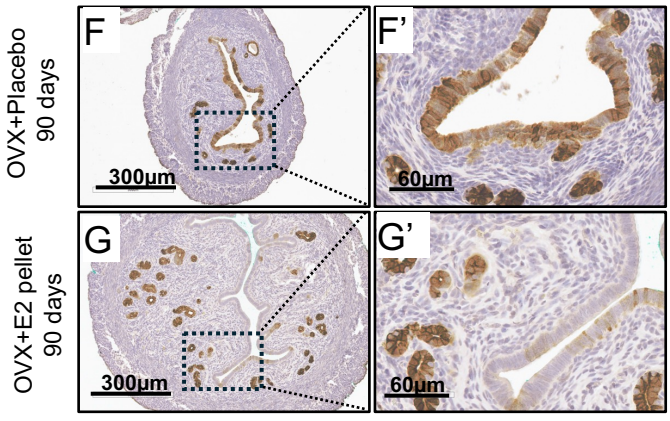
Supplementary Table S5. List of primers and antibodies

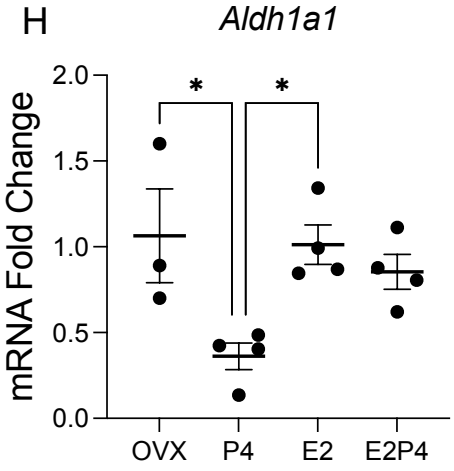
Genotyping Primers	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**
Aldh1a1$^{cre/ERT2}$ (promoter)	AGTCTGCCCATCCAATCATATC	
Aldh1a1$^{cre/ERT2}$ (*Aldh1a1*)	TCTATGGTTCACTGCACTCTGG	
Aldh1a1$^{cre/ERT2}$ (Cre)	GTTCTTGCGAACCTCATCACTC	
Ai9 tdTomato (oIMR9020)	AAGGGAGCTGCAGTGGAGTA	
Ai9 tdTomato (oIMR9021)	CCGAAAATCTGTGGGAAGTC	
Ai9 tdTomato (oIMR9103)	GGCATTAAAGCAGCGTATCC	
Ai9 tdTomato (oIMR9105)	CTGTTCCTGTACGGCATG G	
DTR (mutant)		GCGAAGAGTTTGTCCTCAACC
DTR (common)	AAAGTCGCTCTGAGTTGTTAT	
DTR (WT)		GGAGCGGGAGAAATGGATATG
PCR Primers	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**
Aldh1a1	ATACTTGTCGGATTTAGGAGGCT	GGGCCTATCTTCCAAATGAACA
Gapdh	CAATGTGTCCGTCGTGGATCT	GCCTGCTTCACCACCTTCTT
Rpl17	ACCGCACTGAGATTCGGATG	GAACCTTACGAACCATTTGGGC
Hprt	GTGATTAGCGATGATGAACCA	GCAAGTCTTTCAGTCCTGTC
Antibody Name	**Vendor and Catalog#**	**Dilution used**
ALDH1A1	Abcam, ab52492	1:50 (IHC)
Ki-67	BD Pharmingen, 550609	1:500 (IHC)
cleaved caspase-3	Cell Signaling, 9661	1:400 (IHC)
RFP	Rockland, 600-401-379	1:500 (IHC)
RFP	Abcam, ab125244	1:200 (IF)
FOXA2	Abcam, ab108422	1:200 (IHC), 1:50 (IF)
Cytokeratin 8	DSHB, TROMA-1	1:50 (IF)
Vimentin	Cell Signaling, 5741	1:200 (IF)
Phospho Histone H3	Abcam, ab5176	1:200 (IF)
Acetyl-α-Tubulin	Cell Signaling, 5335	1:200 (IF)
Anti-Rabbit IgG Antibody	Vector Laboratories, BA-1000	1:200 (IHC)
Anti-Mouse IgG Antibody	Vector Laboratories, BA-9200	1:200 (IHC)
Anti-Rat IgG (H+L), AF488	Invitrogen, A-21208	1:250 (IF)
Anti-Rabbit IgG (H+L), AF594	Invitrogen, A-21207	1:250 (IF)
Anti-Mouse IgG (H+L), AF488	Invitrogen, A-21202	1:250 (IF)
Anti-Rabbit IgG (H+L), AF568	Invitrogen, A-10042	1:250 (IF)
Anti-Rat IgG (H+L), AF647	Invitrogen, A78947	1:250 (IF)

Figure S1


A
Veh
2 weeks
OVEX
100ng E2
REST
Veh
Veh
Veh
Veh
Collect 15h later
P4
2 weeks
OVEX
100ng E2
REST
P4
P4
P4
P4
Collect 15h later
50ng E2
2 weeks
OVEX
100ng E2
REST
Veh
Veh
Veh
50ng E2
Collect 15h later
P4 + 50ng E2
2 weeks
OVEX
100ng E2
REST
P4
P4
P4
P4+ 50ng E2
Collect 15h later


OVX+Veh
OVX+1mg P4, 15h
B
B'
C
C'
300µm
60µm
OVX+50ng E2, 15h
OVX+1mg P4+50ng E2, 15h
D
D'
E
E'
300µm
60µm


OVX+Placebo 90 days
OVX+E2 pellet 90 days
F
F'
G
G'
300µm
60µm


H
Aldh1a1
mRNA Fold Change
2.0
1.5
1.0
0.5
0.0
*
*
OVX
P4
E2
E2P4

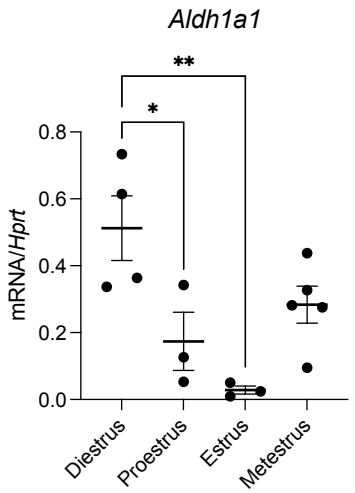

I
Aldh1a1
mRNA/Hprt
0.8
0.6
0.4
0.2
0.0
**
*
Diestrus
Proestrus
Estrus
Metestrus

Figure S2

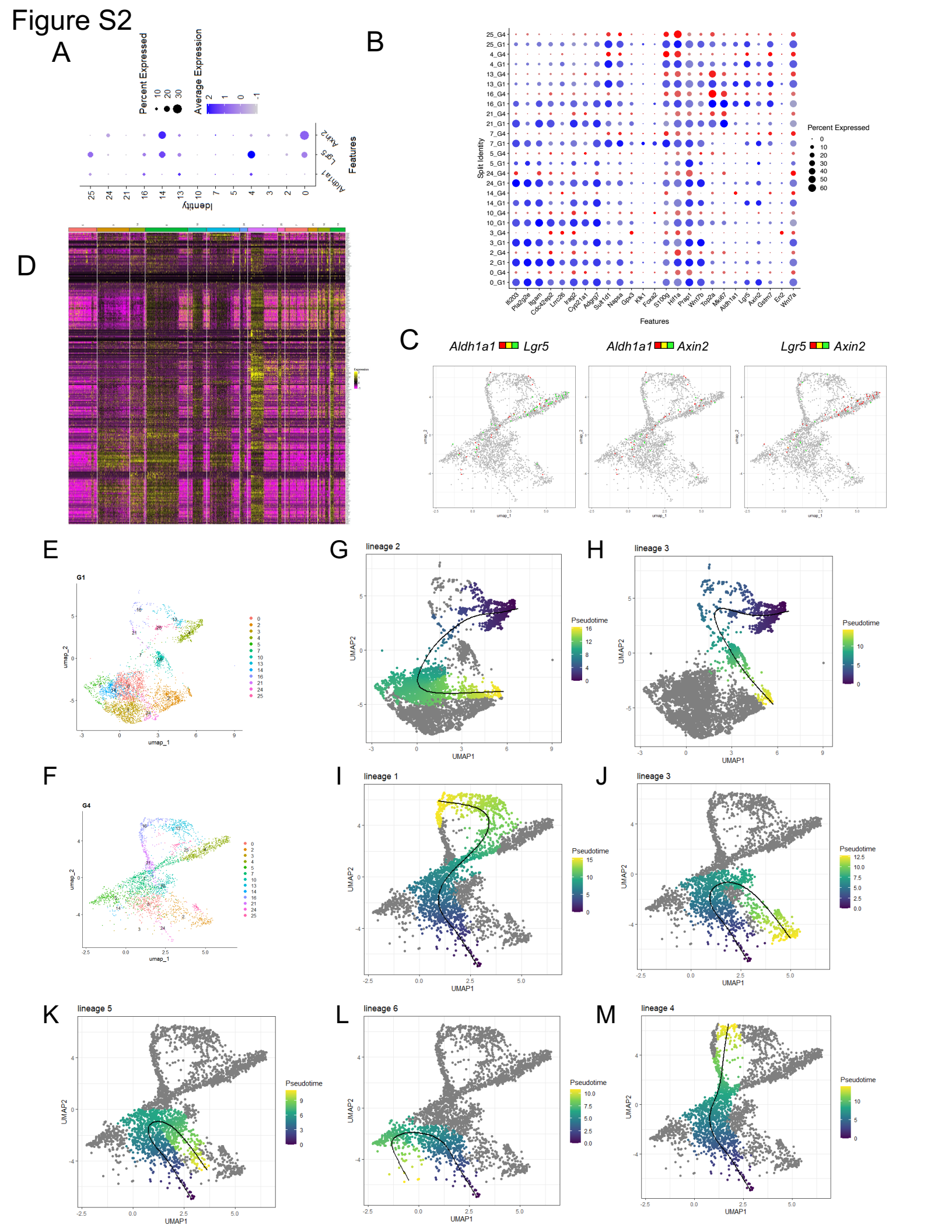

A
Percent Expressed
Average Expression
Features
Identity
B
Split Identity
Features
Percent Expressed
C
Aldh1a1 Lgr5
Aldh1a1 Axin2
Lgr5 Axin2
D
E
G1
F
G4
G
lineage 2
H
lineage 3
I
lineage 1
J
lineage 3
K
lineage 5
L
lineage 6
M
lineage 4
Pseudotime
UMAP1
UMAP2

Figure S3

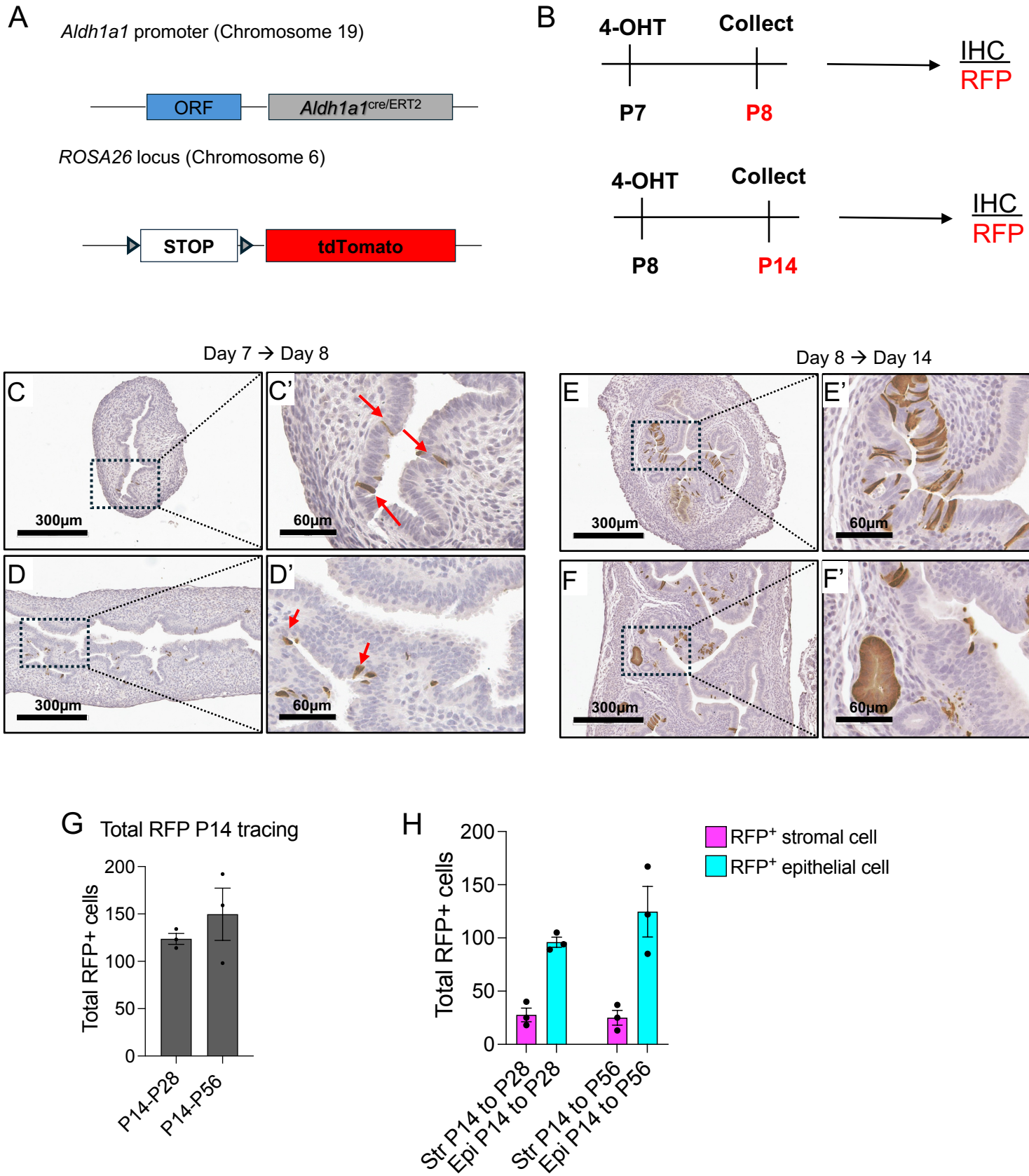

A
Aldh1a1 promoter (Chromosome 19)
ORF
Aldh1a1cre/ERT2
ROSA26 locus (Chromosome 6)
STOP
tdTomato
B
4-OHT
Collect
P7
P8
IHC
RFP
4-OHT
Collect
P8
P14
IHC
RFP
Day 7 → Day 8
Day 8 → Day 14
C
C'
D
D'
E
E'
F
F'
300μm
60μm
G
Total RFP P14 tracing
Total RFP+ cells
200
150
100
50
0
P14-P28
P14-P56
H
Total RFP+ cells
200
150
100
50
0
RFP+ stromal cell
RFP+ epithelial cell
Str P14 to P28
Epi P14 to P28
Str P14 to P56
Epi P14 to P56

Figure S4

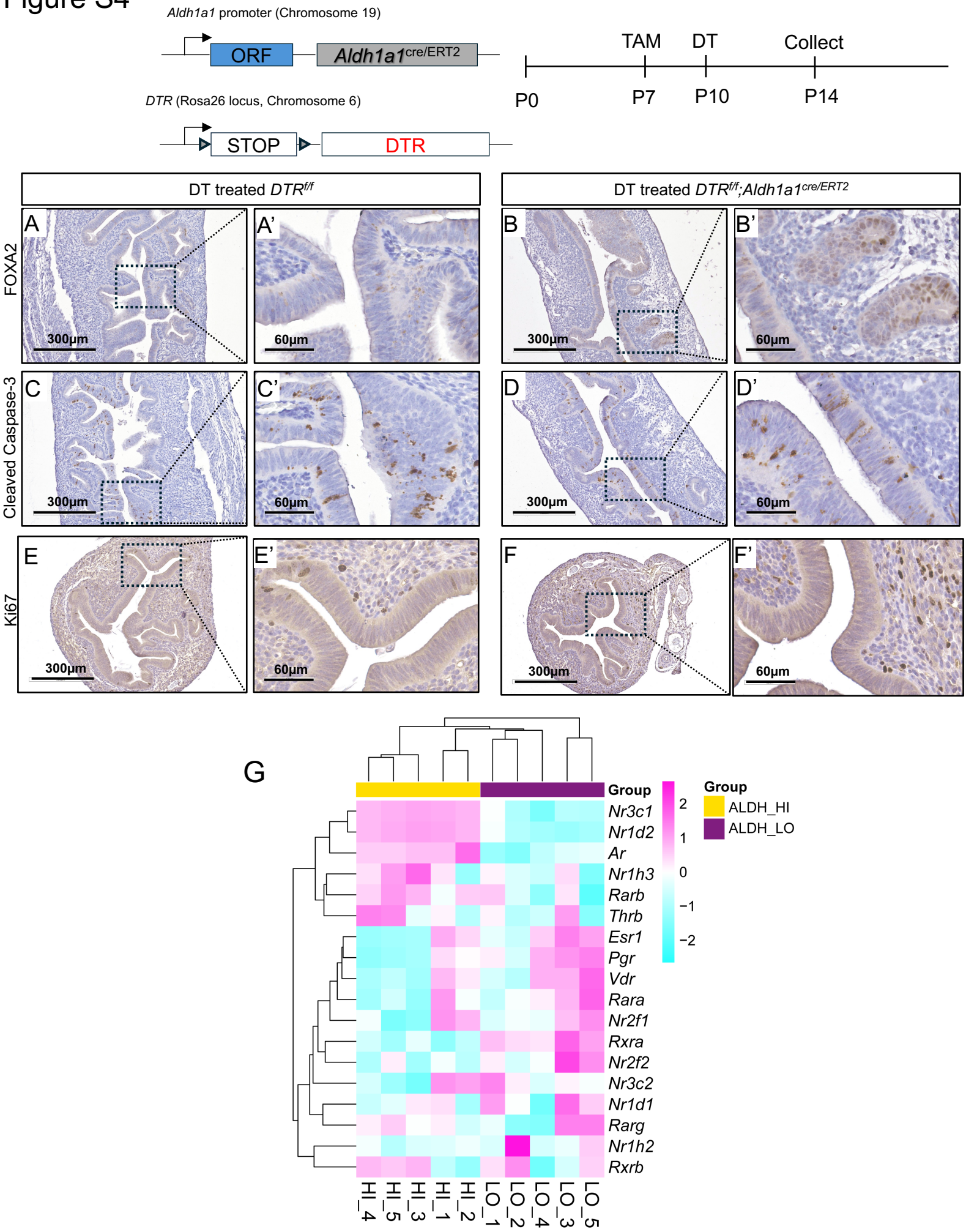
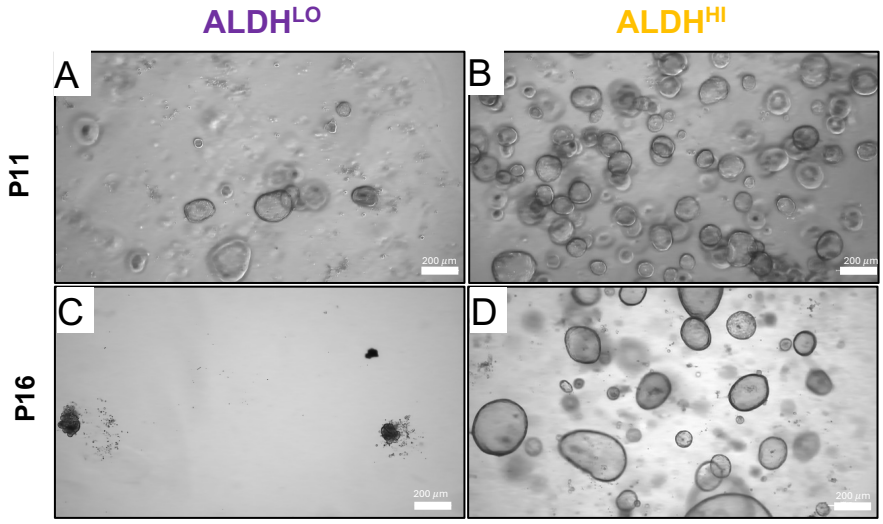
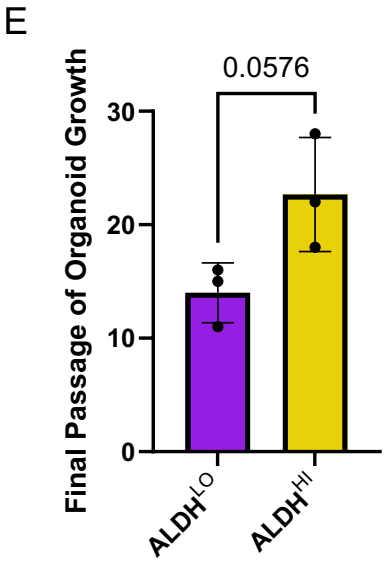
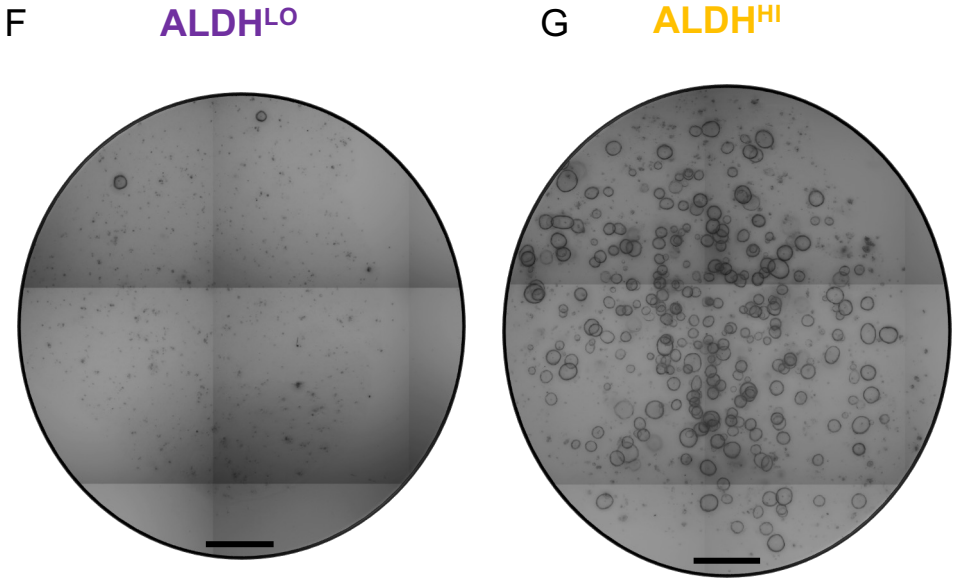
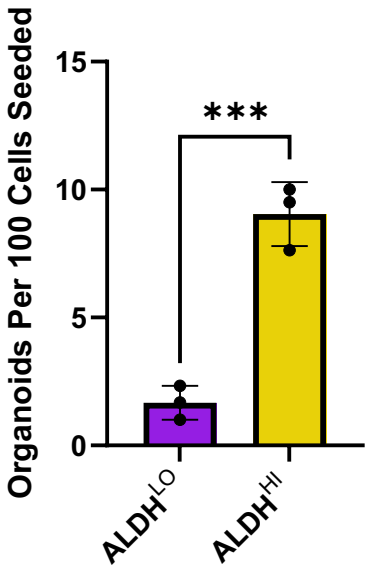

Aldh1a1 promoter (Chromosome 19)
ORF
Aldh1a1cre/ERT2
DTR (Rosa26 locus, Chromosome 6)
STOP
DTR
TAM
DT
Collect
P0
P7
P10
P14
DT treated DTRf/f
DT treated DTRf/f;Aldh1a1cre/ERT2
FOXA2
Cleaved Caspase-3
Ki67
A
A'
B
B'
C
C'
D
D'
E
E'
F
F'
300μm
60μm
G
Group
ALDH_HI
ALDH_LO
Nr3c1
Nr1d2
Ar
Nr1h3
Rarb
Thrb
Esr1
Pgr
Vdr
Rara
Nr2f1
Rxra
Nr2f2
Nr3c2
Nr1d1
Rarg
Nr1h2
Rxrb
HI_4
HI_5
HI_3
HI_1
HI_2
LO_1
LO_2
LO_4
LO_3
LO_5
2
1
0
−1
−2

Figure S5


ALDH^LO
ALDH^HI
A
B
C
D
P11
P16
200 μm


E
0.0576
Final Passage of Organoid Growth
30
20
10
0
ALDH^LO
ALDH^HI


F
ALDH^LO
G
ALDH^HI


H
***
Organoids Per 100 Cells Seeded
15
10
5
0
ALDH^LO
ALDH^HI

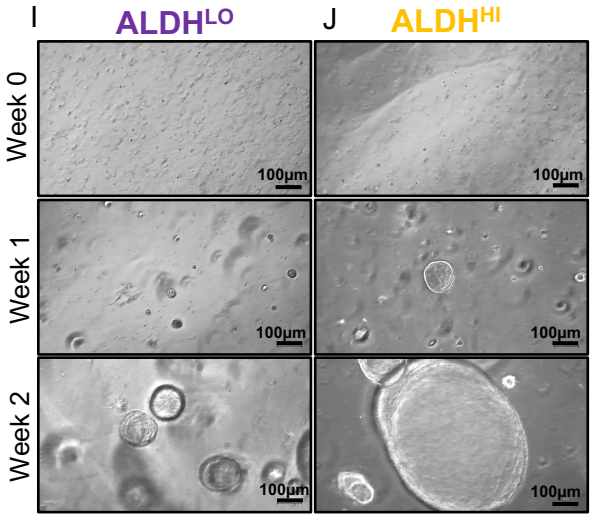

I
ALDH^LO
J
ALDH^HI
Week 0
Week 1
Week 2
100μm

Figure S6

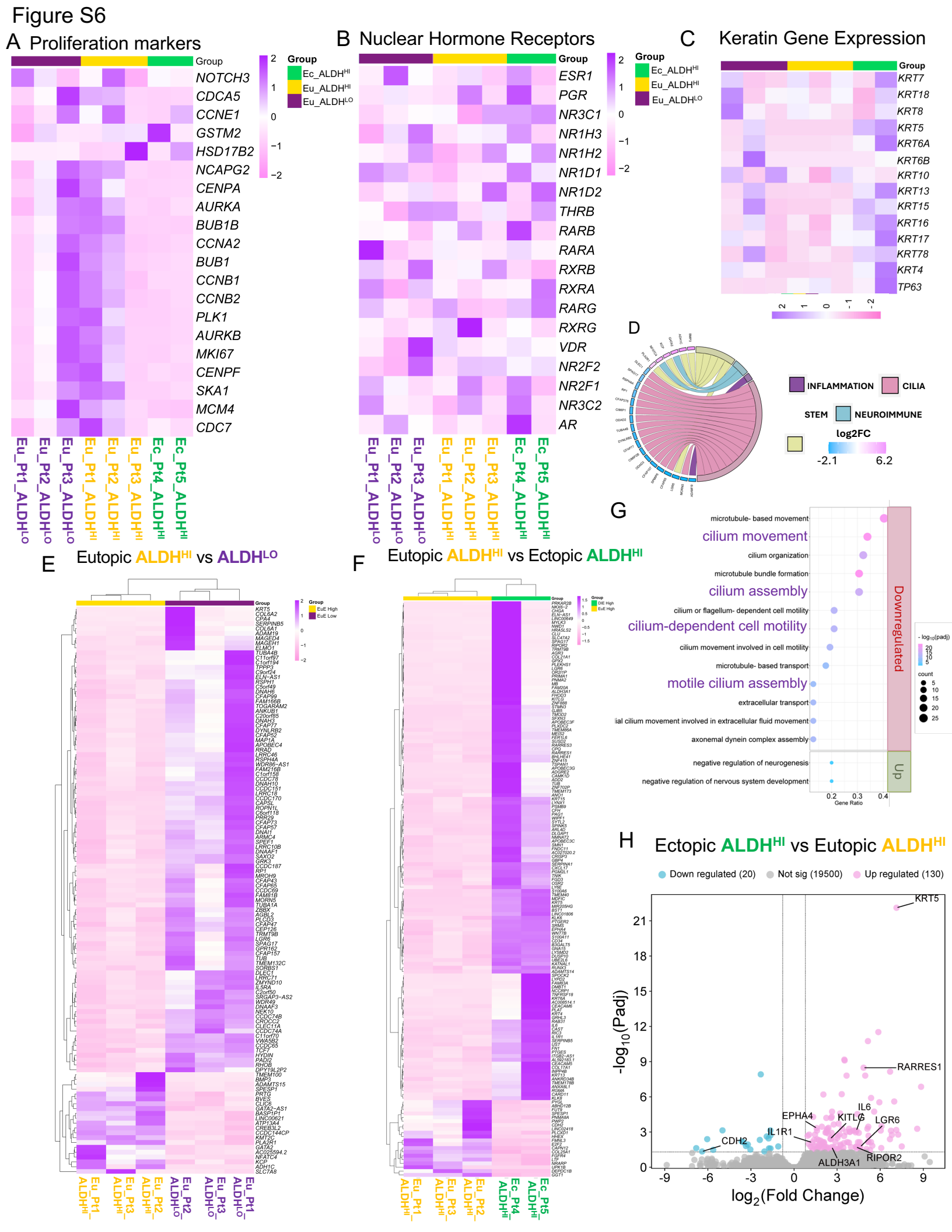
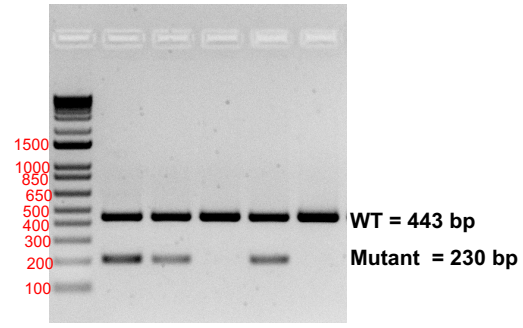

A Proliferation markers
Group
Ec_ALDH^HI
Eu_ALDH^HI
Eu_ALDH^LO
NOTCH3
CDCA5
CCNE1
GSTM2
HSD17B2
NCAPG2
CENPA
AURKA
BUB1B
CCNA2
BUB1
CCNB1
CCNB2
PLK1
AURKB
MKI67
CENPF
SKA1
MCM4
CDC7
Eu_Pt1_ALDH^LO
Eu_Pt2_ALDH^LO
Eu_Pt3_ALDH^LO
Eu_Pt1_ALDH^HI
Eu_Pt2_ALDH^HI
Eu_Pt3_ALDH^HI
Ec_Pt4_ALDH^HI
Ec_Pt5_ALDH^HI
B Nuclear Hormone Receptors
ESR1
PGR
NR3C1
NR1H3
NR1H2
NR1D1
NR1D2
THRB
RARB
RARA
RXRB
RXRA
RARG
RXRG
VDR
NR2F2
NR2F1
NR3C2
AR
C Keratin Gene Expression
KRT7
KRT18
KRT8
KRT5
KRT6A
KRT6B
KRT10
KRT13
KRT15
KRT16
KRT17
KRT78
KRT4
TP63
D
INFLAMMATION
CILIA
STEM
NEUROIMMUNE
log2FC
-2.1
6.2
E Eutopic ALDH^HI vs ALDH^LO
EuE High
EuE Low
F Eutopic ALDH^HI vs Ectopic ALDH^HI
DiE High
EuE High
G
microtubule- based movement
cilium movement
cilium organization
microtubule bundle formation
cilium assembly
cilium or flagellum- dependent cell motility
cilium-dependent cell motility
cilium movement involved in cell motility
microtubule- based transport
motile cilium assembly
extracellular transport
ial cilium movement involved in extracellular fluid movement
axonemal dynein complex assembly
negative regulation of neurogenesis
negative regulation of nervous system development
Downregulated
Up
Gene Ratio
- log10(padj)
count
H Ectopic ALDH^HI vs Eutopic ALDH^HI
Down regulated (20)
Not sig (19500)
Up regulated (130)
KRT5
RARRES1
IL6
EPHA4
KITLG
LGR6
IL1R1
CDH2
RIPOR2
ALDH3A1
-log10(Padj)
log2(Fold Change)

Figure S7

Genotyping information for *Aldh1a1*$^{cre/ERT2}$, Ai9 tdTomato, and DTR alleles

A

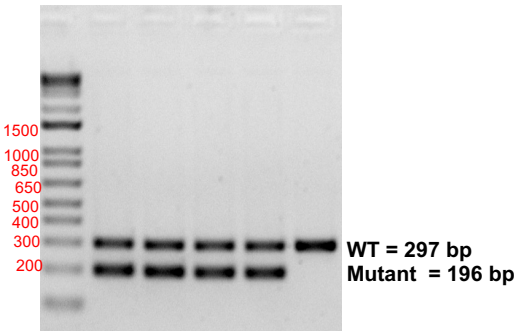


B

Aldh1a1$^{cre/ERT2}$

Step	Temp °C	Time
1	94	4 min
2	94	30 sec
3	60	30 sec
4	72	35 sec
5	Repeat steps 2-4 for 34 cycles	
6	72	5 min
7	16	hold

C



D

Ai9 tdTomato / *DTR*

Step	Temp °C	Time	Note
1	94	2 min	
2	94	20sec	
3	65	15sec	-0.5 C per cycle decrease
4	68	10sec	
5	Repeat steps 2-4 for 10 cycles (Touchdown)		
6	94	15sec	
7	60	15sec	
8	72	10sec	
9	Repeat steps 6-8 for 28 cycles		
10	72	2 min	
11	10	hold	

E