

SUPPLEMENTARY INFORMATION

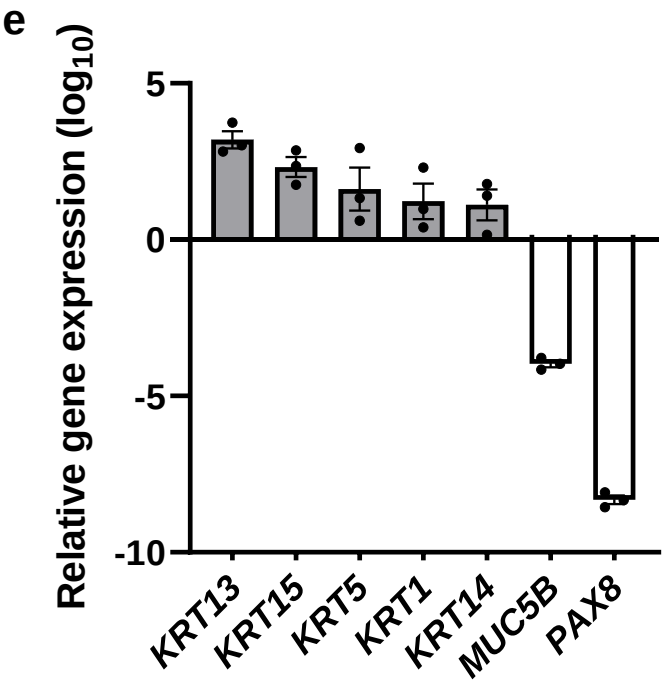
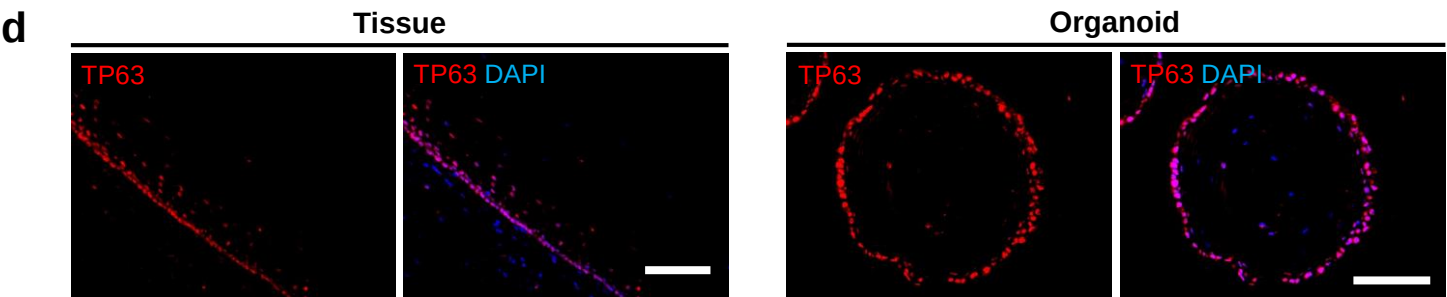
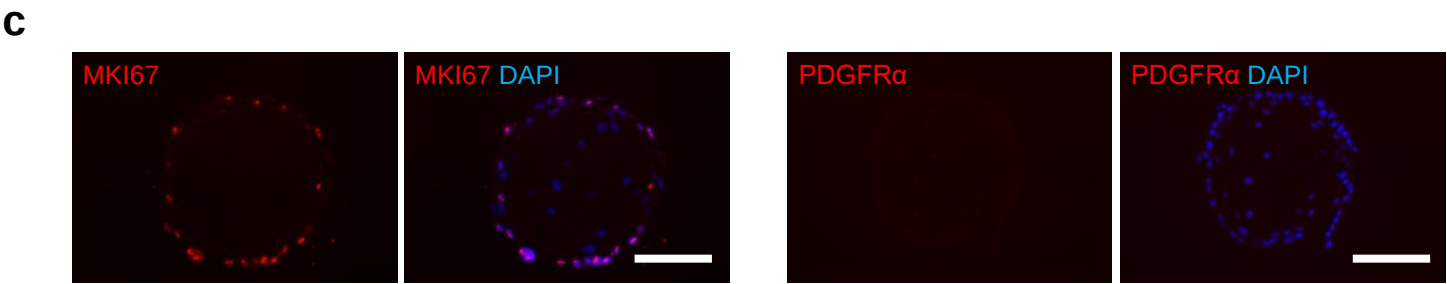
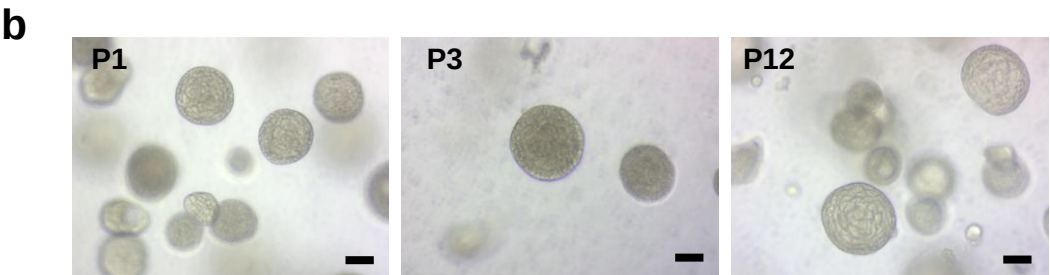
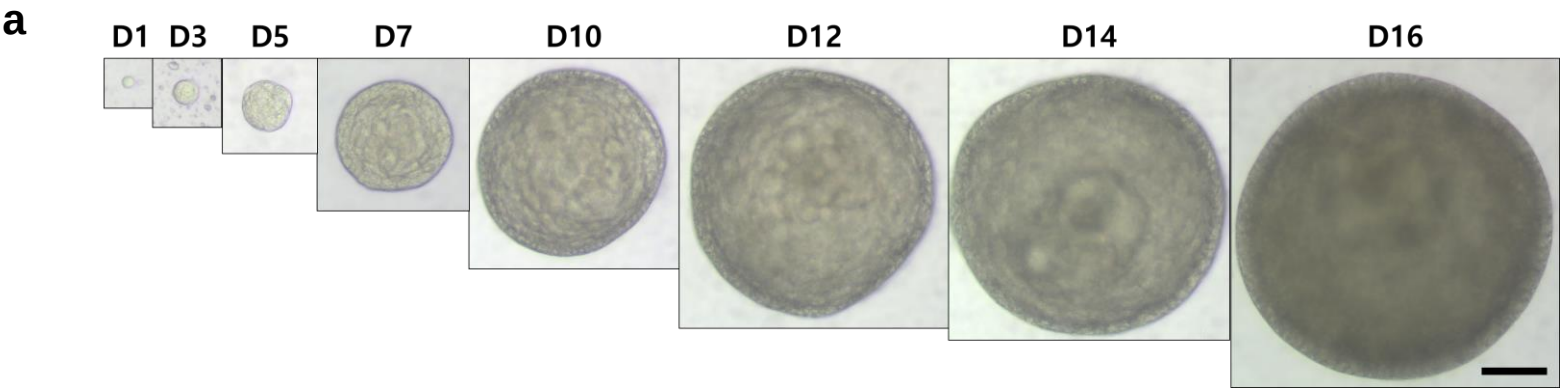
Microbial metabolites control self-renewal and precancerous progression of human cervical stem cells

Supplementary Figure 1-9

Supplementary Table 1-3

Supplementary Figures

Supplementary Figure 1



Supplementary Figure 1. The establishment of long-term, three-dimensional cultures of human ectocervical organoids.

a, Time course images of human ectocervical organoids over 16 days.

b, Representative bright-field images of long-term passaged cultures of human ectocervical organoids, with samples from passages 1, 3, and 12 shown.

c, Immunofluorescence staining of MKI67 in human ectocervical organoids with isotype-matched control (PDGFR α).

d, Immunofluorescence staining of human ectocervix and ectocervical organoids. TP63 was used as the marker for ectocervical basal cells.

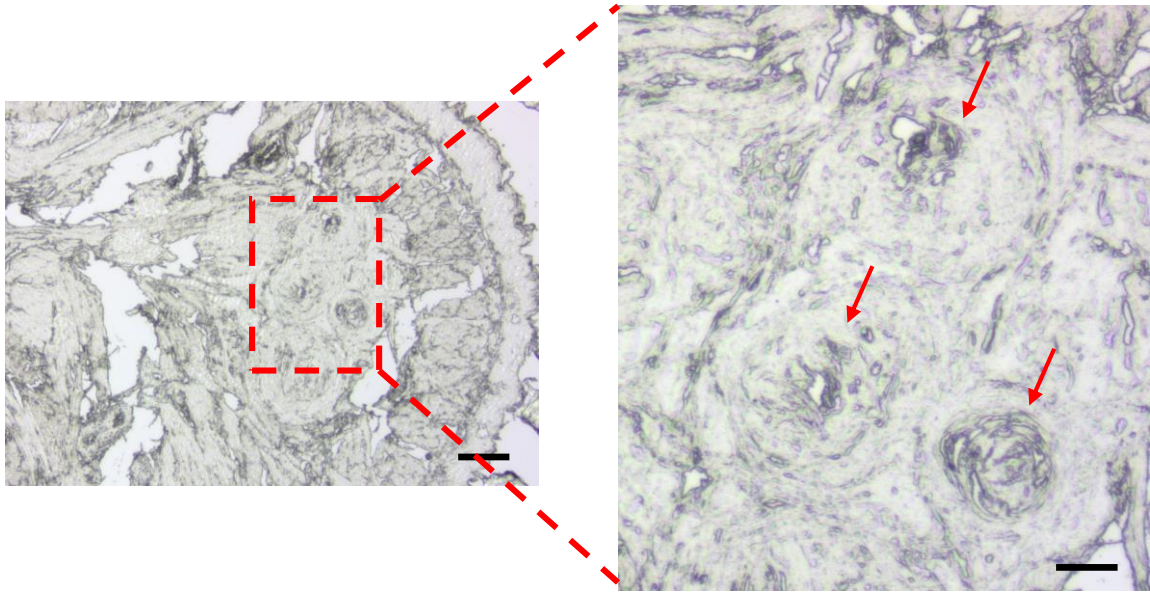
e, Relative gene expression of representative ectocervical and endocervical genes from human ectocervical organoids (grey) compared to endocervical organoids (white). Data are collected from 3 biological replicates and presented as mean $\pm$ S.E.M.

Source data are provided as a Source Data file.

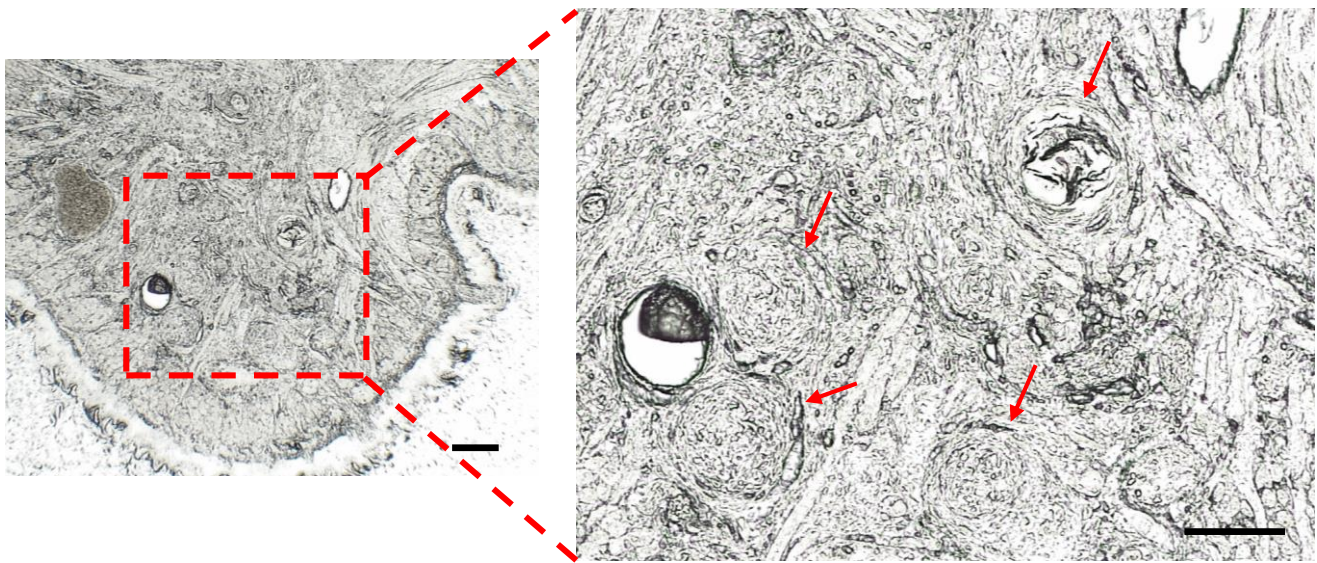
Scale bars: 100 μ m.

Supplementary Figure 2

a



b

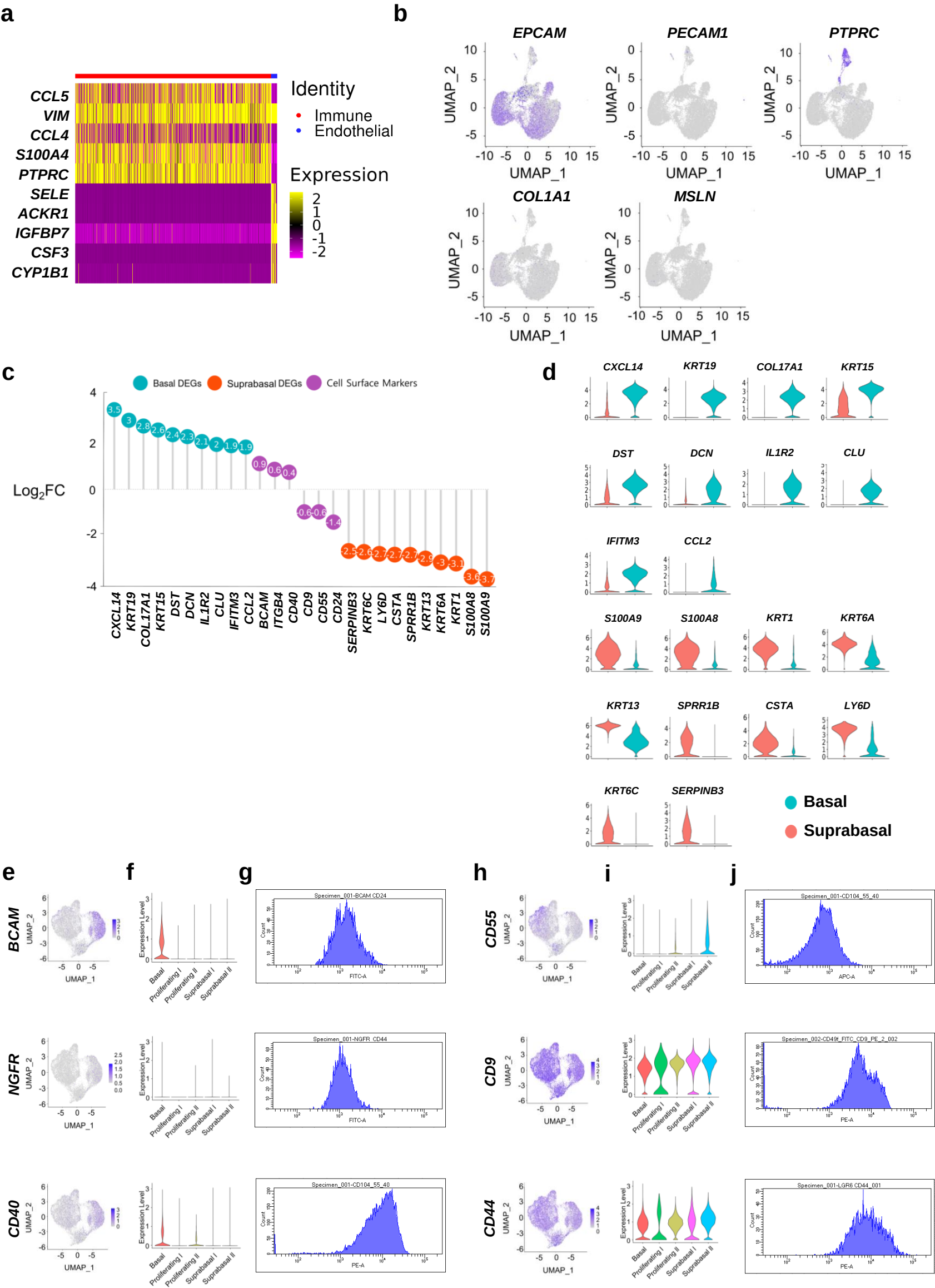


Supplementary Figure 2. Regeneration of stratified-squamous epithelium-like microstructures from human ectocervical organoids or cells.

a, Representative bright-field images of the sectioned murine tongues with the stratified squamous epithelium-like micro-structures regenerated from the transplanted cervical organoids. The tongues were fixed at 4 weeks after organoid transplantation and processed for staining.

b, Representative bright-field images of the sectioned murine tongues with the stratified squamous epithelium-like micro-structures regenerated from the transplanted cervical cells. The tongues were fixed at 3 weeks after the transplantation of single cells and processed for staining.

Supplementary Figure 3



Supplementary Figure 3. scRNA-seq profiling of human ectocervical epithelium.

a, Heatmap of top highly-expressed genes in endothelial cells and immune cells. The color scale represents the z-score distribution -2 (purple) to 2 (yellow), and black means 0.

b, Feature plots of five major cell type-specific marker genes of *EPCAM* (epithelial cells), *PECAM1* (endothelial cells), *PTPRC* (immune cells), *COL1A1* (fibroblasts), and *MSLN* (mesothelial cells).

c, Dot chart showing the top 20 highly expressed genes and purple-colored representative cell surface markers in each cell type.

d, Violin plots of top differentially expressed cell type-specific marker genes of basal and suprabasal cells.

e, Feature plots of three cell surface markers differentially expressed in the basal cell clusters.

f, Violin plots of three cell surface markers differentially expressed in the basal cell clusters.

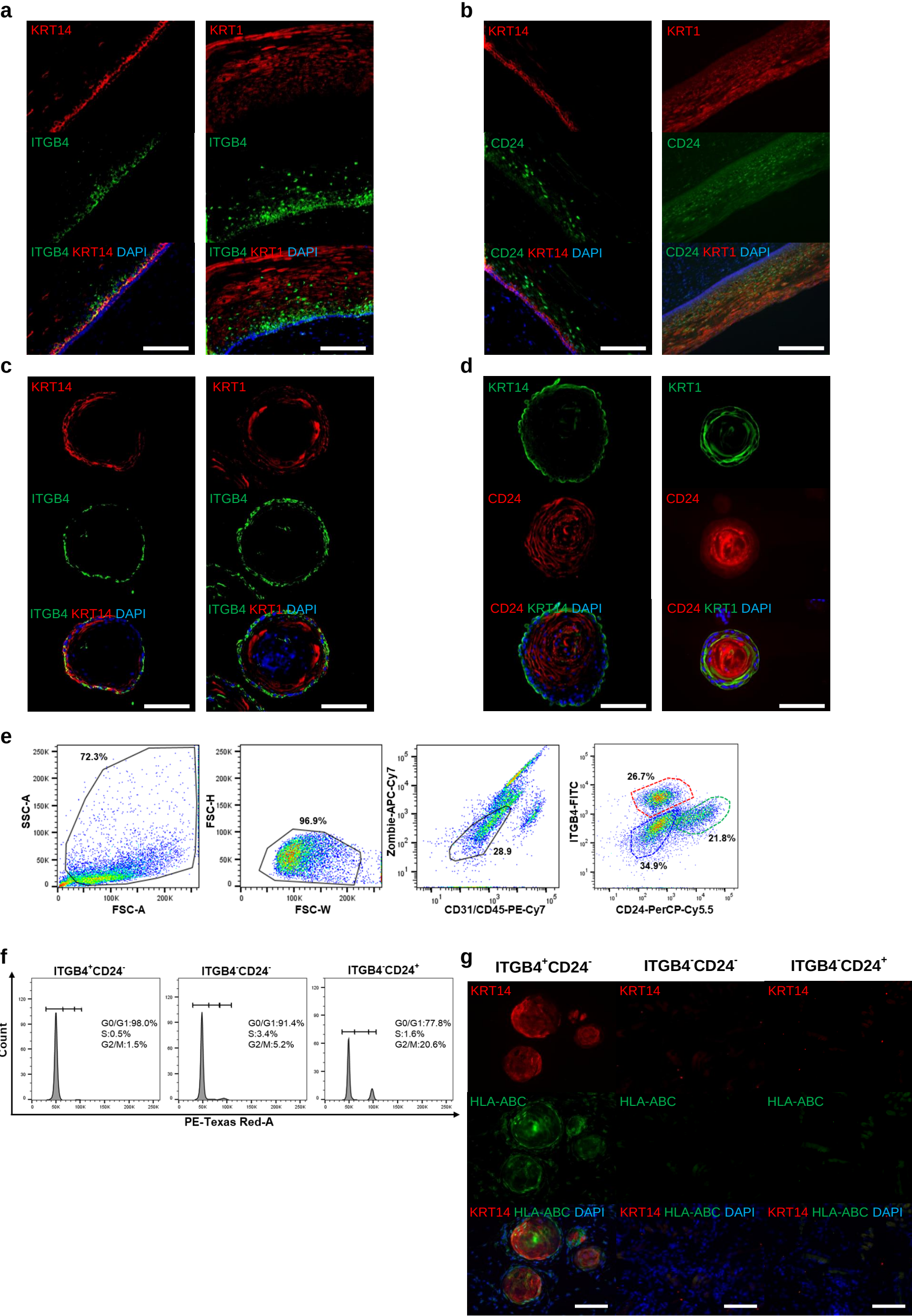
g, Histogram of flow cytometric analyses of three cell surface markers differentially expressed in the basal cell clusters.

h, Feature plots of three cell surface markers differentially expressed in the suprabasal cell clusters.

i, Violin plots of three cell surface markers differentially expressed in the suprabasal cell clusters.

j, Histogram of flow cytometric analyses of three cell surface markers differentially expressed in the suprabasal cell clusters.

Supplementary Figure 4



Supplementary Figure 4. Differential expression of ITGB4 and CD24 in the basal and suprabasal layers.

a, Immunofluorescence staining of ITGB4 and basal (KRT14) or suprabasal (KRT1) markers in the human cervical epithelium.

b, Immunofluorescence staining of CD24 and basal (KRT14) or suprabasal (KRT1) markers in the human cervical epithelium.

c, Immunofluorescence staining of ITGB4 and basal (KRT14) or suprabasal (KRT1) markers in the human cervical organoids.

d, Immunofluorescence staining of CD24 and basal (KRT14) or suprabasal (KRT1) markers in the human cervical organoids.

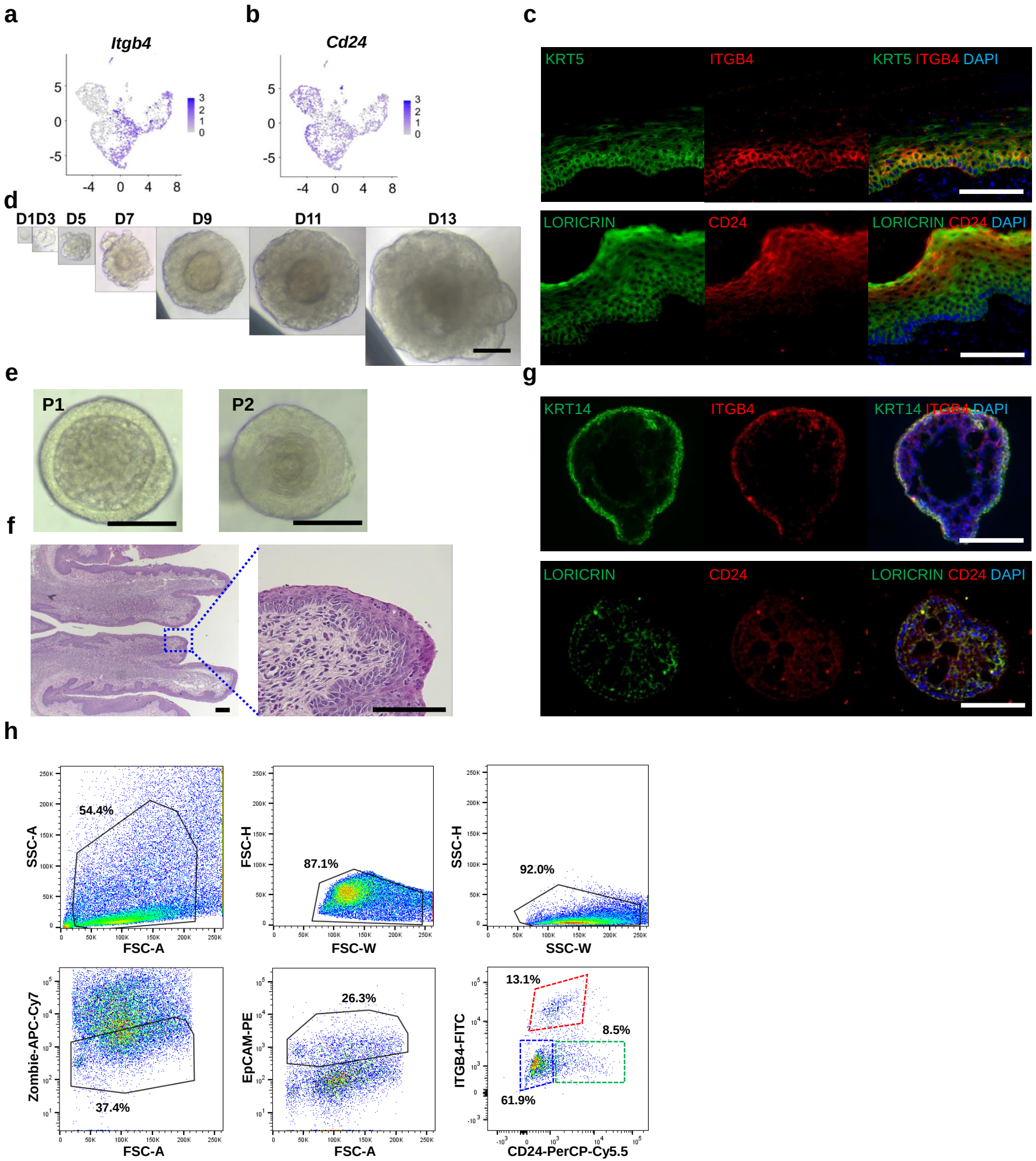
e, Detailed flow cytometric analysis scheme of primary human ectocervical epithelial cells by ITGB4 and CD24.

f, Representative figures out of 5 biological replicates of flow cytometry analysis of cell cycle.

g, Representative immunostaining of the human stratified squamous epithelial tissues regenerated from each subpopulation of human ectocervical cells transplanted inside the immunocompromised murine tongues.

Scale bars: 100 μ m.

Supplementary Figure 5



Supplementary Figure 5. Demonstration of ITGB4 and CD24 as the cell surface markers for murine ectocervical stem cells.

a, Feature plot of *Itgb4* on each cluster.

b, Feature plot of *Cd24* on each cluster.

c, Immunofluorescence staining of ITGB4 and CD24 in murine cervical epithelium.

d, Time course images of murine ectocervical organoids over 2 weeks.

e, Representative bright-field images of murine ectocervical organoids of passages 1 and 2.

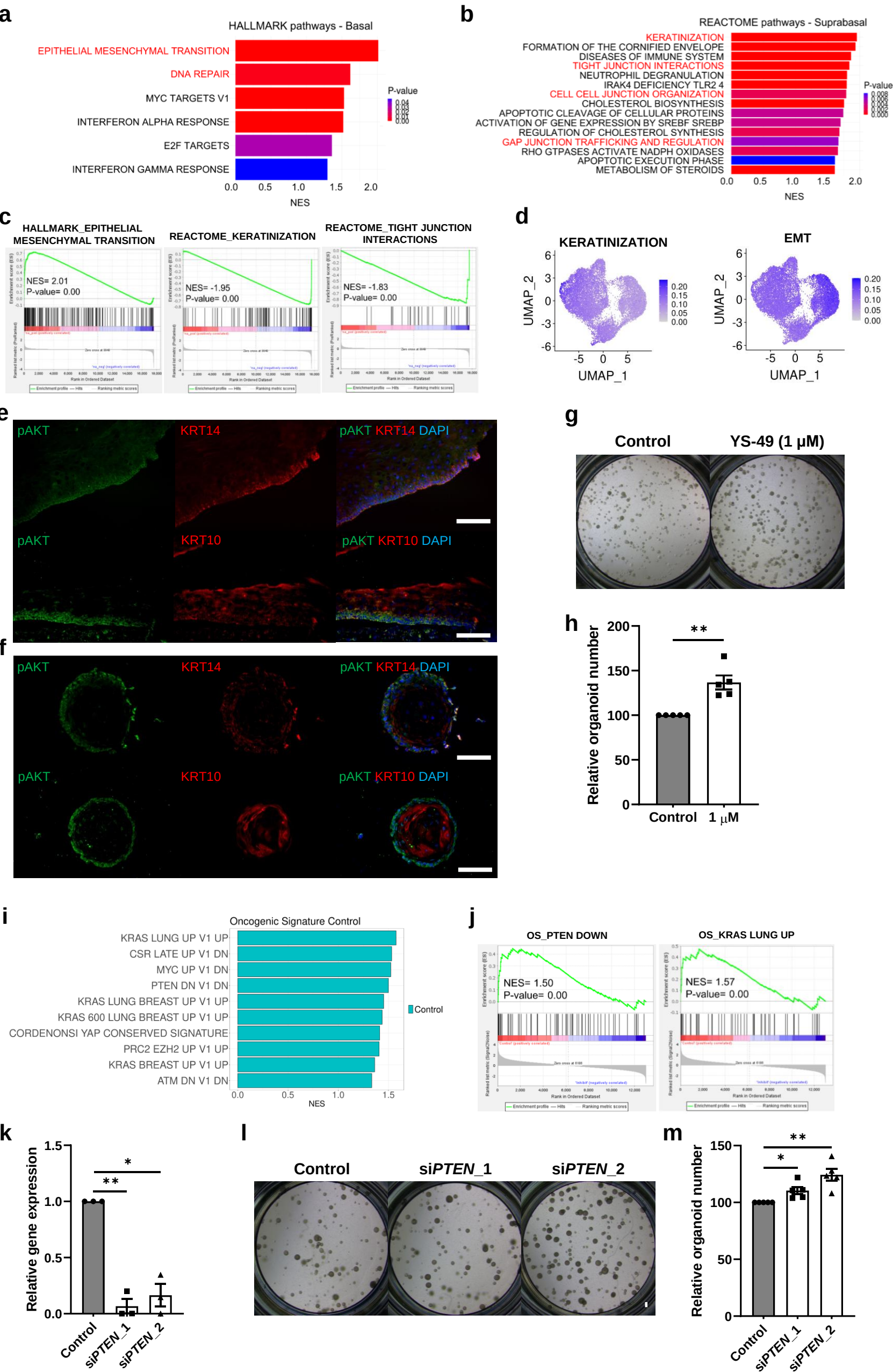
f, H&E staining of the murine cervical epithelium.

g, Immunofluorescence staining of ITGB4 and CD24 in murine cervical organoids.

h, Detailed flow cytometric analysis scheme of primary murine ectocervical epithelial cells using ITGB4 and CD24.

Scale bars: 100 μ m.

Supplementary Figure 6



Supplementary Figure 6. PI3K-AKT pathway regulates the self-renewal of human cervical stem cells.

a,b, Bar graphs representing GSEA of the top upregulated hallmark pathway gene sets for basal cells (**a**) and reactome pathway gene sets for suprabasal cells (**b**). The color scale represents the p-value, with red indicating low values and blue indicating high values. Two-sided Fisher's exact test was employed using negative binomial dispersions with a p-value cutoff 0.05 used to determine statistical significance.

c, GSEA enrichment plots of representative differential gene sets from bulk RNA-seq data.

d, Feature plots of two representative differential gene sets (Keratinization and Epithelial mesenchymal transition) in each epithelial cell type.

e, Immunofluorescence staining of phosphoAKT1 in human ectocervical epithelium. Scale bars: 100 μ m.

f, Immunofluorescence staining of phosphoAKT1 in human ectocervical organoids.

g,h, The effect of a PI3K-AKT pathway activator, YS-49, on the cervical organoid formation. Representative bright field images (**g**) and bar graph for relative numbers (**h**) of the cervical organoids generated ($n = 5$).

i, Bar graph representing GSEA of the top upregulated oncogenic signature gene sets for control cells. Two-sided Fisher's exact test was employed using negative binomial dispersions with p-value cutoff of 0.05 used to determine statistical significance.

j, Additional enriched oncogenic signature gene set in LY294002-treated organoids compared to control organoids ($p < 0.05$).

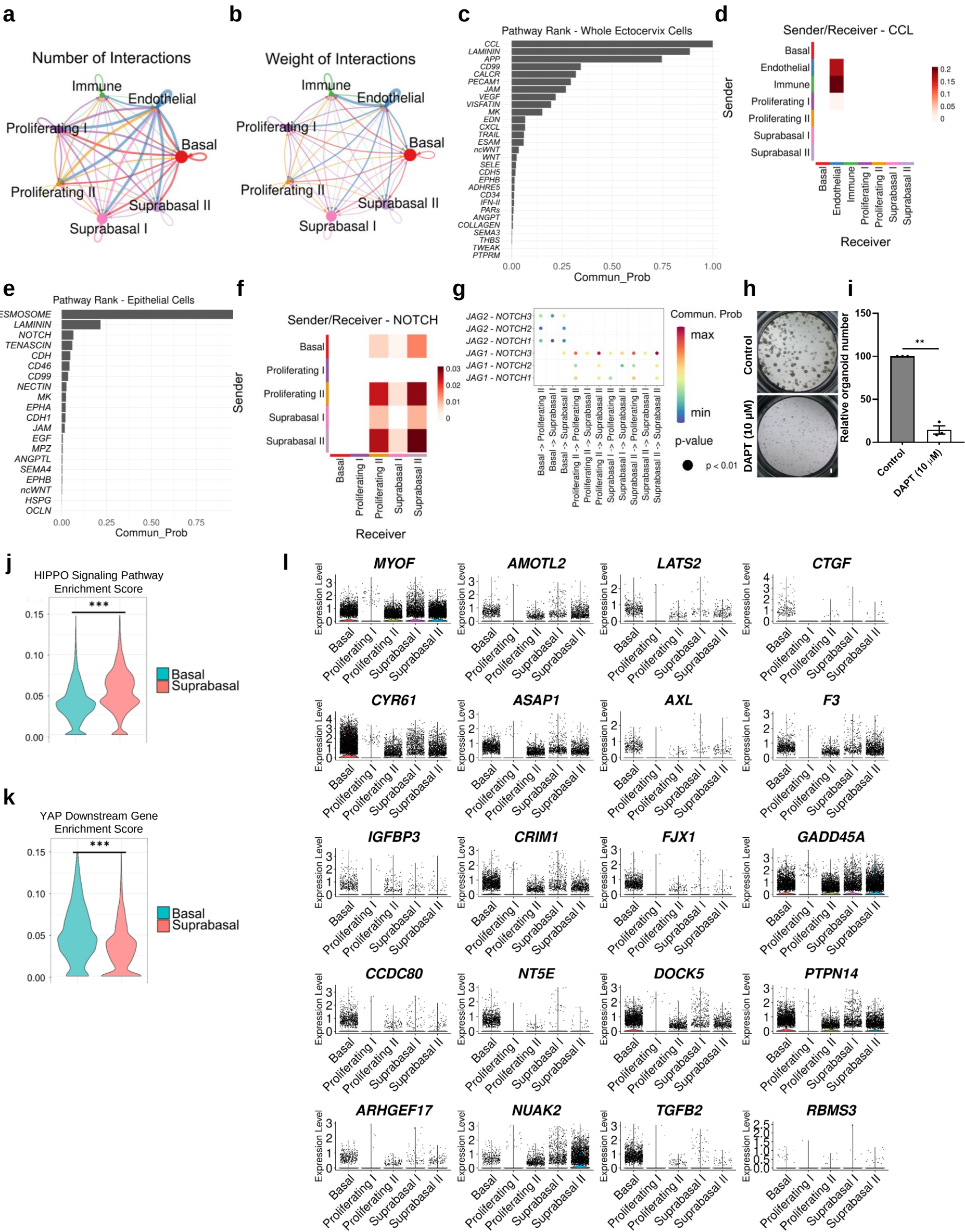
k, Expression of *PTEN* from control or *PTEN* siRNA transfected cervical cells assessed by qPCR ($n = 3$).

l,m, The effect of *PTEN* silencing on the cervical organoid formation. Representative bright field images (**l**) and bar graph for relative numbers (**m**) of the cervical organoids ($n = 5$).

All data are collected from indicated biological replicates and presented as mean $\pm$ S.E.M (*,

$p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). P-values were calculated by two-tailed unpaired Student's t-test (**h**) or one-way ANOVA with Tukey's multiple comparison test (**k,m**). Source data and exact p-values are provided as a Source Data file.

Supplementary Figure 7



Supplementary Figure 7. Network analysis among different cell types in the human ectocervix.

a,b, Total number (**a**) or weights (**b**) of interactions (ligand-receptor pairs) among different cell types in the human ectocervix analyzed using the CellChat R package. Edge width is proportional to the number or weights of ligand–receptor pairs that contribute to the communication between two interacting cell populations.

c, Top-ranked signaling pathways based on their communication probability in whole ectocervix cells.

d, Cell-cell interactions of CCL signals among different cell types in the ectocervix. Colors indicate signal intensity.

e, Top-ranked signaling pathways based on their communication probability within ectocervical epithelial cells.

f, Cell-cell interactions of Notch signals between different epithelial cell types. Colors indicate signal intensity.

g, Detailed ligand-receptor pairs of Notch and their directionality between different epithelial cell types.

h,i, Bright-field images of control and DAPT-treated cervical organoids (**h**) and bar graph for relative organoid numbers (**i**). Data are collected from 3 biological replicates and presented as mean $\pm$ S.E.M (**, $p < 0.01$). P-values were calculated by two-tailed unpaired Student's t-test. Source data and exact p-values are provided as a Source Data file.

j, Violin plot showing enrichment scores of the Hippo signaling pathway gene sets in the basal and suprabasal clusters from our scRNA-seq data.

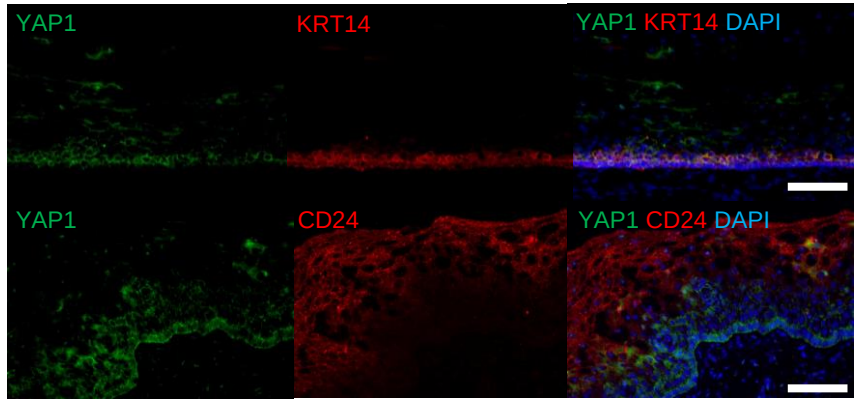
k, Violin plot showing enrichment scores of YAP downstream genes in the basal and suprabasal clusters from our scRNA-seq data. Data are presented as mean $\pm$ S.E.M (***, $p < 0.001$).

l, Feature plots of 20 representative YAP1 target genes.

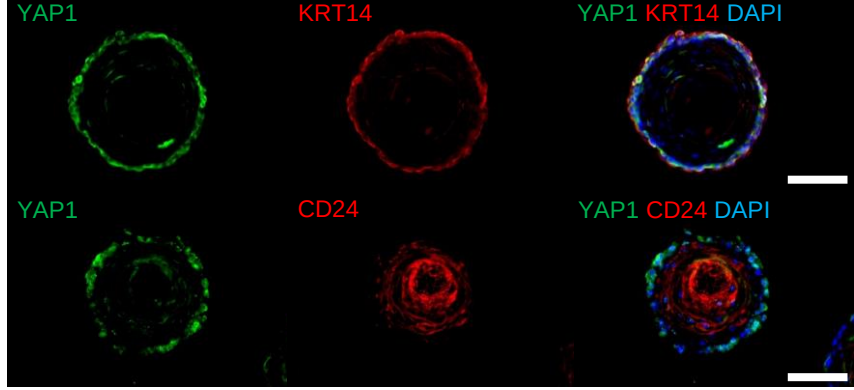
Source data are provided as a Source Data file.

Supplementary Figure 8

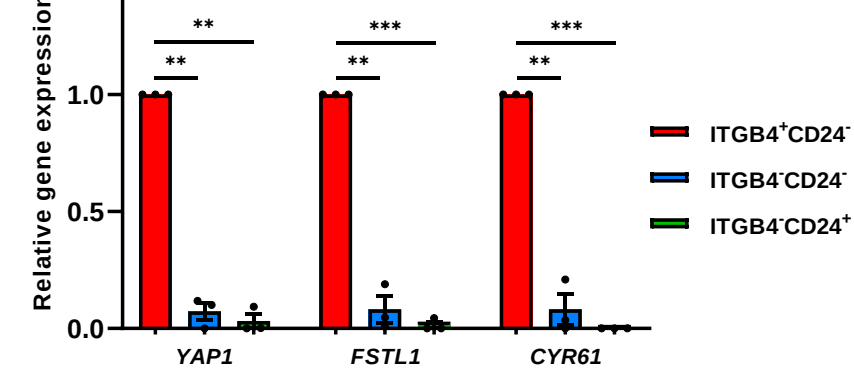
a



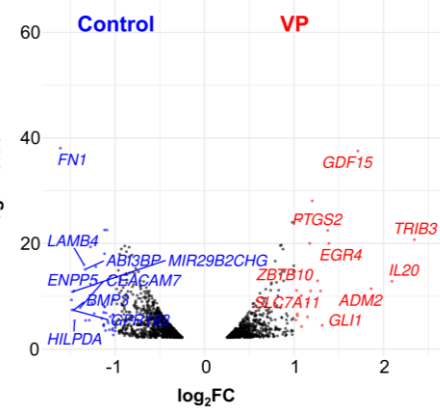
b



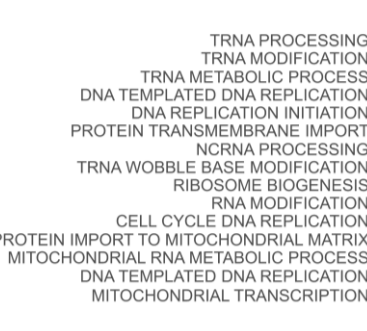
c



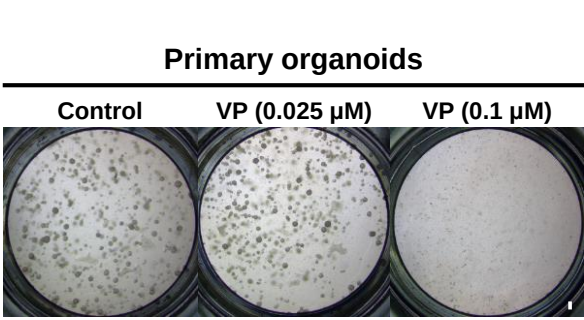
j



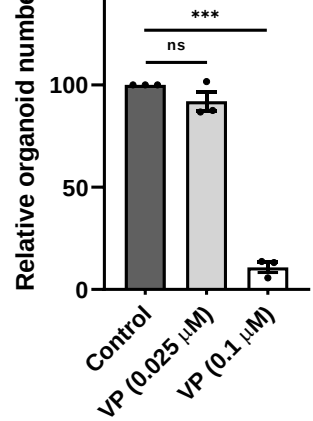
k



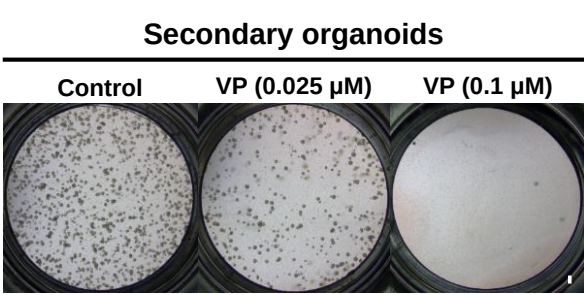
d



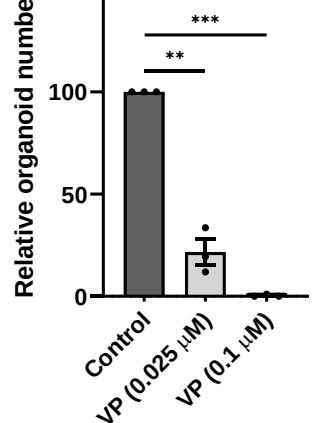
e



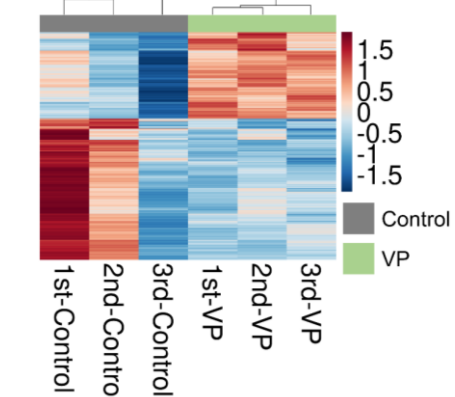
f



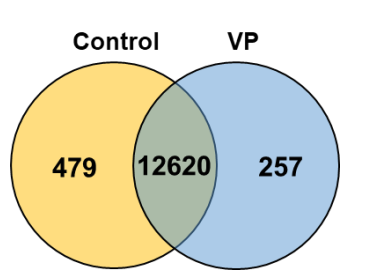
g



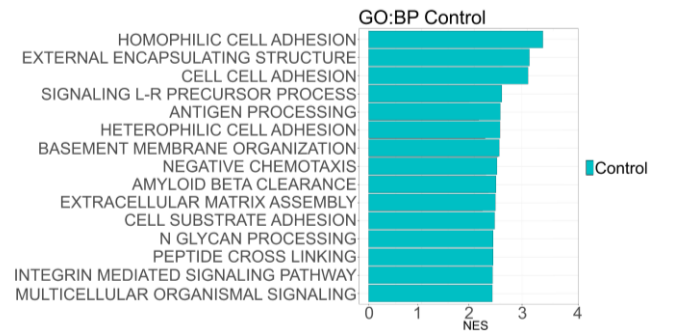
h



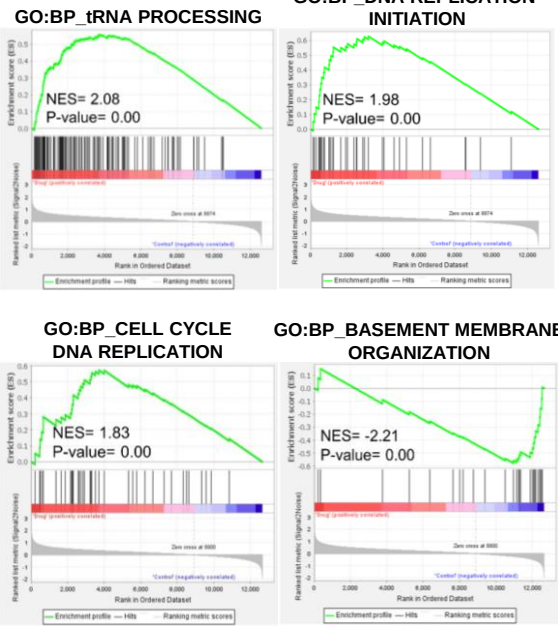
i



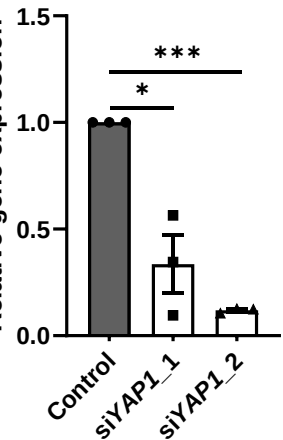
l



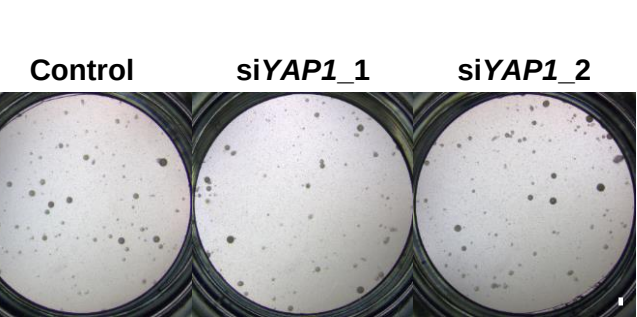
m



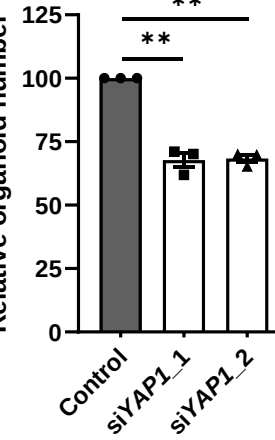
n



o



p



Supplementary Figure 8. YAP1 regulates the self-renewal of human cervical stem cells.

a, Immunofluorescence staining of YAP1 in the human cervical epithelium. Scale bars: 100 μm .

b, Immunofluorescence staining of YAP1 in human ectocervical organoids.

c, Relative gene expression of YAP1 and its target genes ($n = 3$).

d,e, The effect of a YAP1 inhibitor, verteporfin (VP), on primary organoid formation. Representative bright field images (**d**) and bar graph for relative numbers (**e**) ($n = 3$).

f,g, The effect of pharmacological YAP1 inhibition by VP on secondary organoid formation. Representative bright field images (**f**) and bar graph for relative numbers (**g**) ($n = 3$).

h, Heat map of control and VP-treated human ectocervical organoids

i, Venn diagram of DEGs following VP treatment

j, Volcano plot showing DEGs in the control and VP-treated organoids. Running sum statistics with permutation test ($n=1,000$) was employed.

k,l, GSEA bar graph representing the top-ranked GO:BP gene sets enriched in VP-treated (**k**) or control (**l**) organoids. Two-sided Fisher's exact test was employed using negative binomial dispersions ($p < 0.05$).

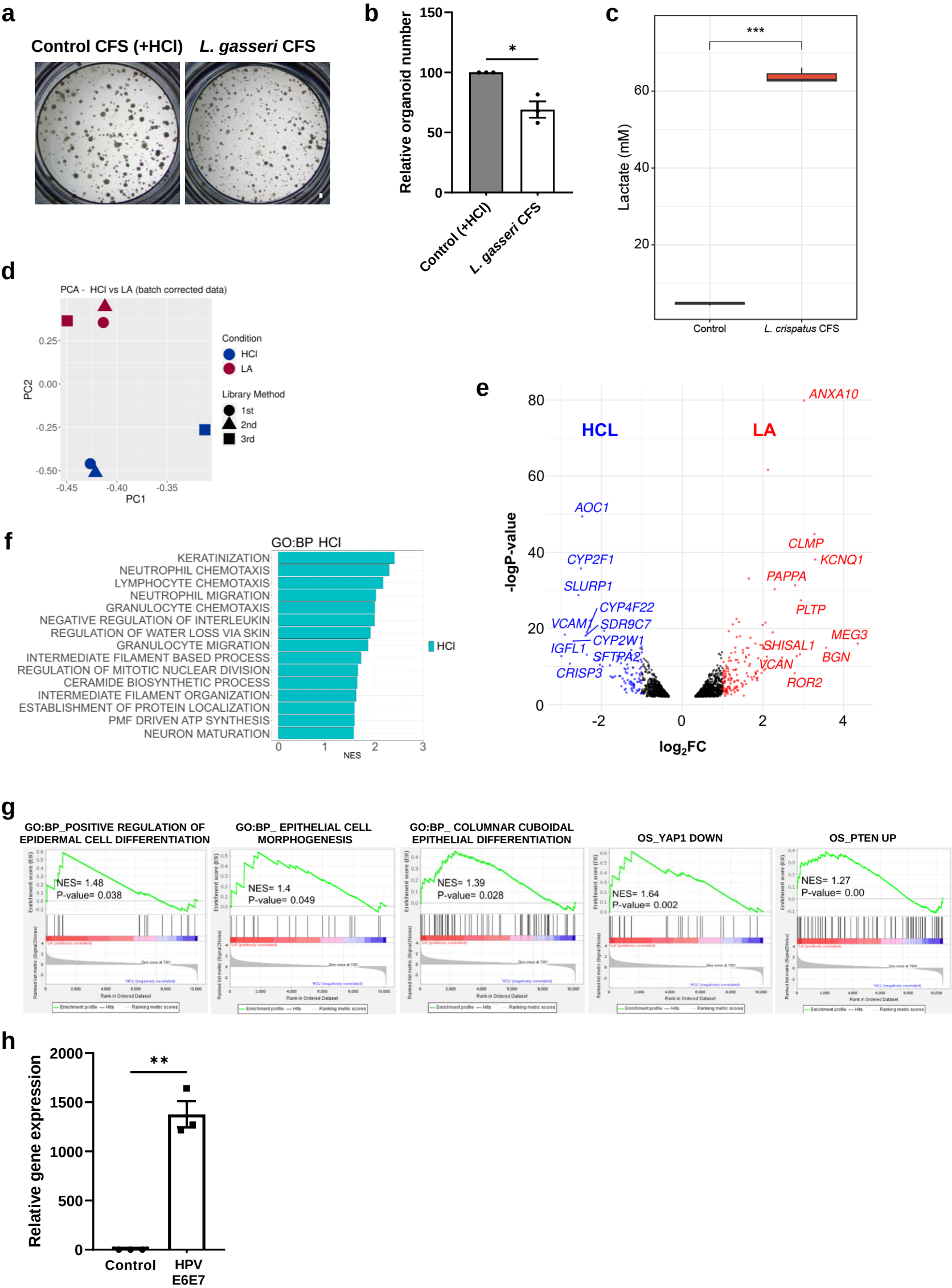
m, GSEA enrichment plots of representative differential gene sets from RNA-seq data.

n, Expression of *YAP1* from control or YAP1 siRNA-transfected cervical cells assessed by qPCR ($n = 3$).

o,p, The effect of YAP1 genetic silencing by siRNA on primary organoid formation. Representative bright-field images (**o**) and bar graph for relative numbers (**p**) ($n = 3$).

All data are collected from indicated biological replicates and presented as mean $\pm$ S.E.M (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). P-values were calculated by one-way ANOVA with Tukey's multiple comparison test (**e,g,n,p**). Source data and exact p-values are provided as a Source Data file.

Supplementary Figure 9



Supplementary Figure 9. The effect of lactic acid on human cervical stem cells.

a,b, The effect of *L. gasseri* CFS on human ectocervical organoid formation. Representative bright field images (**a**) and bar graph for relative numbers (**b**) of organoids ($n = 3$).

c, Measurement of lactate concentration in control media and *L. crispatus* CFS (Wilcoxon-rank sum test, *** $p < 0.001$).

d, Principal component analysis of differential gene expression from HCl- or LA-treated human ectocervical organoids.

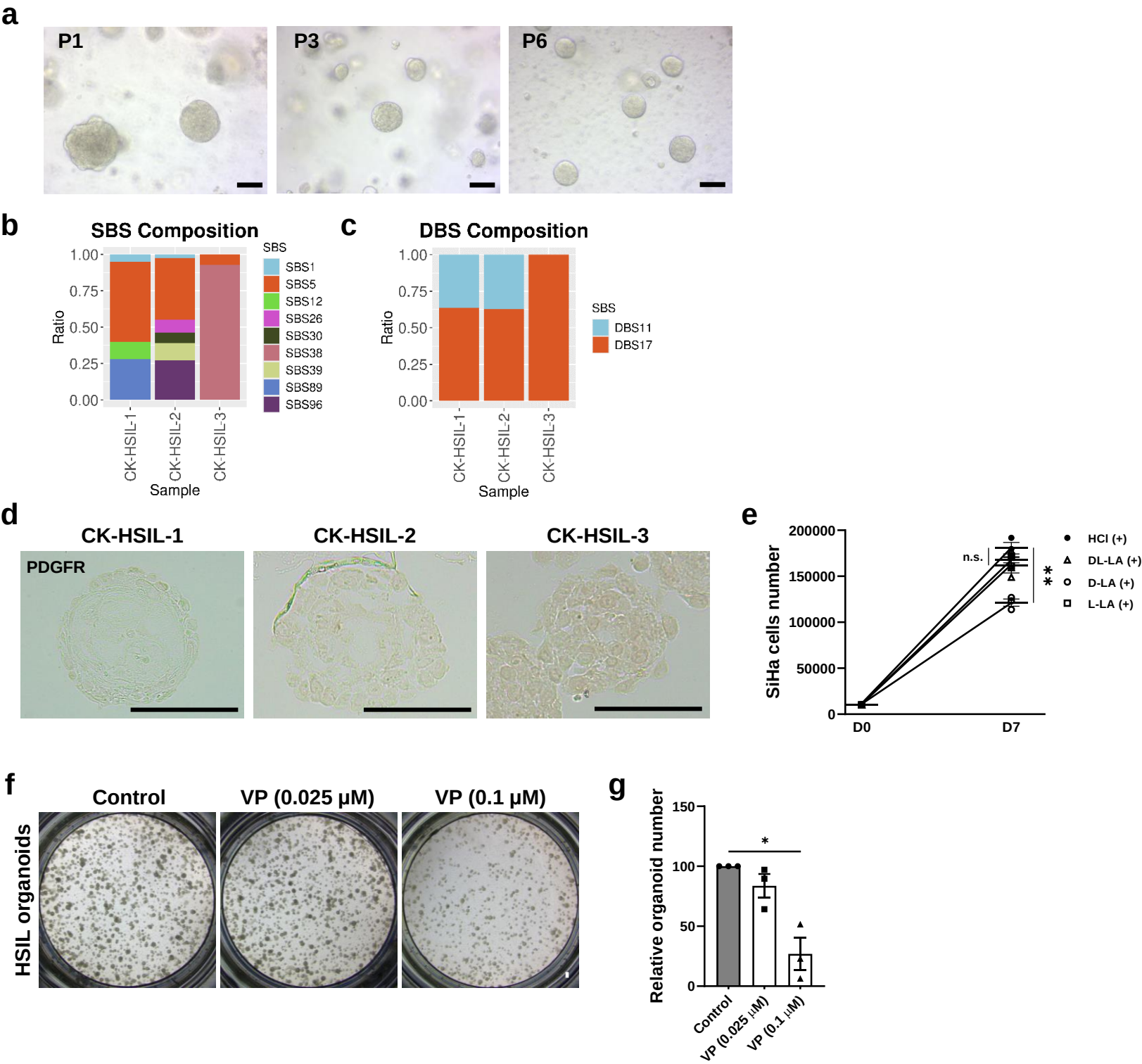
e, Volcano plot of DEGs in HCl- or LA-treated human ectocervical organoids. Genes upregulated in HCl- or LA-treated human ectocervical organoids (p -adjusted value < 0.05 with a $\log_2$ (Fold Change) > 1) are colored by blue and red, respectively. The top 10 DEGs for each direction are denoted by gene name.

f, Most enriched GO:BP gene sets in HCl-treated organoids compared to LA-treated organoids. Two-sided Fisher's exact test was employed using negative binomial dispersions ($p < 0.05$).

g, Representative enrichment plots for gene sets highly enriched in LA-treated organoids. Gene sets related to positive regulation of epidermal cell differentiation, epithelial cell morphogenesis, columnar cuboidal epithelial cell differentiation, YAP, and PTEN are shown.

h, Relative mRNA expression of HPV16 E6 genes analyzed by qRT-PCR ($n = 3$). Data are collected from 3 biological replicates and presented as mean $\pm$ S.E.M (**, $p < 0.01$). P-values were calculated by two-tailed unpaired Student's t-test (**b,h**). Source data and exact p-values are provided as a Source Data file.

Supplementary Figure 10



Supplementary Figure 10. The effect of lactic acid on human cervical cancer cells.

a, Representative bright-field images of long-term passaged cultures of human ectocervical organoids, with samples from passages 1, 3, and 6 shown.

b,c, Ratios of SBS (**b**) and DBS (**c**) mutation signatures defined by overall variants of two cervical HSIL organoids (CK-HSIL-1/2) and one HSIL tissue (CK-HSIL-3) using COSMIC database. The individual SBS / DBS numbers are associated with a correlated aetiology.

d, Immunohistochemistry staining of human HSIL organoids with isotype-matched control.

e, The effect of lactate isomers on cervical cancer cell proliferation. 20,000 SiHa cells were treated with DL-LA, D-lactate, and L-lactate, and cell numbers were counted after 7 days ($n = 3$).

f,g, The effect of VP on HSIL organoid formation. Representative bright field images (**f**) and bar graph for relative numbers (**g**) of HSIL organoids ($n = 3$).

All data are collected from indicated biological replicates and presented as mean $\pm$ S.E.M (**, $p < 0.01$; ***, $p < 0.001$; n.s., not significant). P-values were calculated by one-way ANOVA with Tukey's multiple comparison test (**e,g**). Source data and exact p-values are provided as a Source Data file.

Supplementary Tables

Supplementary Table 1. Clinical characteristics of tissue donors.

Patient ID	Sex	Tissue
1	Female	Ectocervix
2	Female	Ectocervix
3	Female	Ectocervix
4	Female	Ectocervix
5	Female	Ectocervix
6	Female	Ectocervix
7	Female	Ectocervix
8	Female	Ectocervix
9	Female	Ectocervix
10	Female	Ectocervix, Endocervix
11	Female	Ectocervix, Endocervix
12	Female	Ectocervix, Endocervix
13	Female	Ectocervix
14	Female	Ectocervix
15	Female	Ectocervix
16	Female	Ectocervix
17	Female	Ectocervix
18	Female	Ectocervix
19	Female	Ectocervix
20	Female	Ectocervix
21	Female	Ectocervix
22	Female	Ectocervix
23	Female	Ectocervix
24	Female	Ectocervix
25	Female	Ectocervix
26	Female	Ectocervix
27	Female	Ectocervix
28	Female	Ectocervix
29	Female	Ectocervix

30	Female	Ectocervix
31	Female	Ectocervix

Supplementary Table 2. List of quantitative PCR primers.

Name	Sequence
Human <i>KRT14</i> fw	5'-AAG TGA AGA TCC GTG ACT GGT-3'
Human <i>KRT14</i> rv	5'-ATT GTC CAC TGT GGC TGT GAG-3'
Human <i>TP63</i> fw	5'-GAC AGG AAG GCG GAT AG-3'
Human <i>TP63</i> rv	5'-TGT TTC TGA AGT AAG TGC TGG TGC-3'
Human <i>KRT13</i> fw	5'-GAC CGC CAC CAT TGA AAA CAA-3'
Human <i>KRT13</i> rv	5'-CCC ACC ATC AGG AGA GAG TC-3'
Human <i>KRT1</i> fw	5'-CCG AAG GAG AGT GGA CCA AC-3'
Human <i>KRT1</i> rv	5'-CTC TGC ATT TGT CCG CTT GT-3'
Human <i>KRT15</i> fw	5'-AGAAAT CTG AAT TCC TAT TGC AGG-3'
Human <i>KRT15</i> rv	5'-CCC TGAAAG CTT AGA CCG AGG GAC -3'
Human <i>MUC5B</i> fw	5'-CAG AAC CAG GCT GAC GAC TT -3'
Human <i>MUC5B</i> rv	5'-ATG CAG TTC GAG TGG AAG GG-3'
Human <i>PAX8</i> fw	5'-AAG GTG GTG GAG AAG ATT GG-3'
Human <i>PAX8</i> rv	5'-AGG GAG GTT GAA TGG TTG C-3'
Human <i>YAP1</i> fw	5'-GCAAAT TCT CCA AAA TGT CAG G-3'
Human <i>YAP1</i> rv	5'-CGG GAG AAG ACA CTG GAT TT-3'
Human <i>FSTL1</i> fw	5'-GCC ATG ACC TGT GAC GGAA A-3'
Human <i>FSTL1</i> rv	5'-CAG CGC TGA AGT GGA GAA GA-3'
Human <i>MKI67</i> fw	5'-TCC TTT GGT GGG CAC CTAAGA CCT G-3'
Human <i>MKI67</i> rv	5'-TGA TGG TTG AGG TCG TTC CTT GAT G-3'
Human <i>CDC20</i> fw	5'-GAC CAC TCC TAG CAA ACC TGG-3'
Human <i>CDC20</i> rv	5'-GGG CGT CTG GCT GTT TTC A-3'
Human <i>KRT5</i> fw	5'-ATC GCC ACT TAC CGC AAG C-3'
Human <i>KRT5</i> rv	5'-CCA CTG CCA TAT CCA GAG GAA A-3'
Human <i>CYR61</i> fw	5'-TGA TGA TCC AGT CCT GCA AA-3'
Human <i>CYR61</i> rv	5'-GTC TAG GTG TGC CCT GGAA A-3'
Human <i>CTGF</i> fw	5'-GTT TGG CCC AGA CCC AAC TA-3'
Human <i>CTGF</i> rv	5'-GGC TCT GCT TCT CTA GCC TG-3'
Human <i>ACTB</i> fw	5'-CCC TGG ACT TCG AGC AAG AG-3'

Human <i>ACTB</i> rv	5'-ACT CCA TGC CCA GGA AGG AA-3'
Human <i>AKT1</i> fw	5'-ACC TGA CCA AGA TGA CAG CA-3'
Human <i>AKT1</i> rv	5'-CCA TCC CTC CAA GCT ATC GT-3'
Human <i>AKT2</i> fw	5'-ATG TGT GTT TGT GTG AGG GC-3'
Human <i>AKT2</i> rv	5'-GCA GCT CAT GGA TCA CAG TG-3'
Human <i>PIK3CA</i> fw	5'-CCT ATT GTC GTG CAT GTG GG-3'
Human <i>PIK3CA</i> rv	5'-AAT CTG GTC GCC TCA TTT GC-3'
Human <i>EIF4E</i> fw	5'-CCC TAC AGA ACA GAT GGG CA-3'
Human <i>EIF4E</i> rv	5'-AGT AGT CAC AGC CAG GCA TT-3'
Human <i>PTEN</i> fw	5'-CAG AAA GAC TTG AAG GCG TAT-3'
Human <i>PTEN</i> rv	5'-AAC GGC TGA GGG AAC TC-3'
HPV16 E6 fw	5'-ACA AAC CGT TGT GTG ATT TGT T-3'
HPV16 E6 rv	5'-CAG TGG CTT TTG ACA GTT AAT ACA-3'

Supplementary Table 3. List of siRNAs. Related to STAR Methods.

Name	Sequence
Negative control	SN-1003 (Bioneer, Republic of Korea)
Human <i>YAP1</i> _1 fw	5'-CAG AAG AUC AAA GCU ACU U-3'
Human <i>YAP1</i> _1 rv	5'-AAG UAG CUU UGA UCU UCU G-3'
Human <i>YAP1</i> _2 fw	5'-AGA ACC GUU UCC CAG ACU A-3'
Human <i>YAP1</i> _2 rv	5'-UAG UCU GGG AAA CGG UUC U-3'
Human <i>AKT1</i> _1 fw	5'-GAC AAC CGC CAU CCA GAC U-3'
Human <i>AKT1</i> _1 rv	5'-AGU CUG GAU GGC GGU UGU C-3'
Human <i>AKT1</i> _2 fw	5'-CCU UUU CGA CGC UUA ACC U-3'
Human <i>AKT1</i> _2 rv	5'-AGG UUA AGC GUC GAA AAG G-3'
Human <i>AKT2</i> _1 fw	5'-CAG AAU GCC AGC UGA UGA A-3'
Human <i>AKT2</i> _1 rv	5'-UUC AUC AGC UGG CAU UCU G-3'
Human <i>AKT2</i> _2 fw	5'-ACA CAA GGU ACU UCG AUG A-3'
Human <i>AKT2</i> _2 rv	5'-UCA UCG AAG UAC CUU GUG U-3'
Human <i>PTEN</i> _1 fw	5'-GAC ACG UUU UCC AUA CCU U-3'
Human <i>PTEN</i> _1 rv	5'-AAG GUA UGG AAA ACG UGU C-3'
Human <i>PTEN</i> _2 fw	5'-CUG CUA ACA UGA AAC UCU U-3'
Human <i>PTEN</i> _2 rv	5'-AAG AGU UUC AUG UUA GCA G-3'