



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

Received: 30 September 2024

Accepted: 18 February 2025

Published online: 08 March 2025

 Check for updates

Jihyeon Myeong^{1,14}, Minho Lee^{2,14}, Bawool Lee^{1,3,14}, Joon Hyung Kim², Yeji Nam⁴, Yeseul Choi^{5,6}, Jeongmin Kim¹, Se Young Jeon^{7,8}, Haewon Shim², Da-Ryung Jung⁹, Youngjin Shin², Minsoo Jeong⁹, Byungmoo Oh^{1,3}, Jaehun Jung², Christine S. Kim^{1,3}, Hyung Soo Han^{5,6,8,10}, Jae-Ho Shin⁹, Yoon Hee Lee^{7,8,11}, Nora Jee-Young Park^{8,12,13}, Gun Oh Chong²  & Youngtae Jeong² 

Cervical cancer is the fourth most common female cancer, with the uterine ectocervix being the most commonly affected site. However, cervical stem cells, their differentiation, and their regulation remain poorly understood. Here, we report the isolation of a population enriched for human cervical stem cells and their regulatory mechanisms. Using single-cell RNA sequencing, we characterize the cellular heterogeneity of the human ectocervix and identify cluster-specific cell surface markers. By establishing normal and precancerous cervical organoids and an intralingual transplantation system, we show that ITGB4 and CD24 enable enrichment of human and murine ectocervical stem cells. We discover that *Lactobacilli*-derived lactic acid regulates cervical stem cells' self-renewal and early tumorigenesis through the PI3K-AKT pathway and YAP1. Finally, we show that D-lactic acid suppresses growth of normal and precancerous organoids, while L-lactic acid does not. Our findings reveal roles of human cervical stem cells and microbial metabolites in cervical health and diseases.

With an estimated incidence of more than 600,000 cases per year worldwide^{1,2}, cervical cancer is the fourth most common cancer in women. Infection with high-risk human papillomavirus (HPV) strains such as HPV16 and HPV18 is a well-known inducer of malignant transformation of cervical epithelial cells³, and encouragingly, widespread HPV vaccination has decreased cervical cancer incidence over the past decade. More than 70% of cervical cancers are squamous cell

carcinoma (SCC), most of which develop in the ectocervix and progress through premalignant stages of low grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL) into cervical cancer⁴.

In addition to HPV, the microbiota in the cervicovaginal tract can regulate the integrity of the cervix and its carcinogenesis. Higher levels of *Lactobacilli* seem to be protective against HPV infection⁵, while

¹Department of New Biology, DGIST, Daegu, Korea. ²Department of Life Science, Dongguk University, Gyeonggi-do, Korea. ³New Biology Research Center, DGIST, Daegu, Korea. ⁴School of Undergraduate Studies, DGIST, Daegu, Korea. ⁵Graduate Program, Department of Biomedical Science, Kyungpook National University School of Medicine, Daegu, Korea. ⁶BK21 Four Program, Kyungpook National University School of Medicine, Daegu, Korea. ⁷Department of Obstetrics and Gynecology, Kyungpook National University Chilgok Hospital, Daegu, Korea. ⁸Clinical Omics Institute, Kyungpook National University, Daegu, Korea. ⁹Department of Applied Biosciences, Kyungpook National University, Daegu, Korea. ¹⁰Department of Physiology, Kyungpook National University School of Medicine, Daegu, Korea. ¹¹Department of Obstetrics and Gynecology, Kyungpook National University School of Medicine, Daegu, Korea. ¹²Department of Pathology, Kyungpook National University Chilgok Hospital, Daegu, Korea. ¹³Department of Pathology, Kyungpook National University School of Medicine, Daegu, Korea. ¹⁴These authors contributed equally: Jihyeon Myeong, Minho Lee, Bawool Lee. ✉e-mail: gochong@knu.ac.kr; jyt@dgist.ac.kr

higher vaginal microbial diversity is associated with HPV infection and cervical cancer^{6,7}. Previously, we found that *Lactobacilli* were more abundant in the vaginas of healthy control women than in the pre-malignant cervix; in contrast, *Streptococcus* species were more abundant in the vaginas of women with cervical cancers than those with premalignant lesions⁸. Experimental studies showed that *Lactobacillus* supernatant inhibits the cell proliferation of diverse cancer cells^{9,10}. Tissue stem cells often serve as the cell of origin for cancers^{11–13}, and microbiota regulate tissue stem cells^{14,15}. Therefore, microbiota may control cancer development by regulating tissue stem cells, especially in a tissue like the cervix where the microbiota is in close physical contact with the tissue. However, the regulatory roles of the cervicovaginal microbiota in the self-renewal of cervical stem cells are not well understood.

The human ectocervical epithelium is composed of stratified squamous epithelium^{16,17}. In the ectocervical epithelium, basal cells are attached to the basal lamina and differentiate through parabasal cells into superficial cells. These differentiated cells form the protective epithelial barrier that directly interacts with microbial flora and environmental toxins. Although basal cells are thought to include cervical stem cells, the identity and cell surface markers to isolate human cervical stem cells have not been reported. Recent studies performed single-cell transcriptomic analyses of mouse cervical epithelium and identified murine putative stem cell clusters^{18,19}. However, it remains unclear whether these findings in mice translate to humans. Also, other studies have recently established human cervical organoids for the study of cervical stem cells, but did not identify the human cervical stem cell population, and the efficiency of organoid generation still needs to be improved^{18,20}. In addition to the role of cervical stem cells in the tissue maintenance and regeneration, it is thought that cervical stem cells are the target of HPV and that HPV infection induces uncontrolled expansion of cervical stem cells, leading to cervical cancer^{17,21}. Therefore, the identification of human cervical stem cells and a deeper understanding of their regulatory mechanisms and intercellular networks in the human cervix are crucial to build a picture of how the cervical epithelium maintains its integrity and guards against cervical cancer formation. Furthermore, the identification of human cervical stem cells will help to uncover the role of microbial commensals in supporting or altering cervical stem cell activity.

Here, we report the identification of a population enriched for human ectocervical stem cells and their regulatory mechanisms. We performed single-cell RNA sequencing of the human ectocervical epithelium and established human cervical normal and precancerous organoids and a murine intralingual transplantation system as new model systems to study human cervical tissue in vitro and in vivo, which enabled us to enrich human ectocervical stem cells and identify their differentiation trajectory. We further show that the PI3K-AKT signaling pathway, YAP1, and lactic acid (LA) isomers produced by commensal *Lactobacilli* differentially regulate the self-renewal and differentiation of human ectocervical stem cells. Finally, we show the distinct role of LA isomers in cervical stem cells and precancerous progression.

Results

Establishment of human cervical organoids

In order to identify human cervical stem cells and their differentiation trajectory, we first aimed to establish a human cervical organoid system and an intralingual transplantation system, which can be used to evaluate the self-renewal capability of cervical stem cells (Fig. 1a). Organoids are stem cell-derived 3-dimensional structures resembling the histology of their tissue of origin^{22,23}. Organoid passaging mimics the self-renewal of tissue stem cells. Previous studies, including ours, employed organoids to identify human tissue stem cells including tracheal and esophageal stem cells^{24,25}. Moreover, organoids

can provide important model systems to study tissue (patho) physiology in vitro and in high throughput without having to rely on animal models that often struggle to recapitulate human-specific biology.

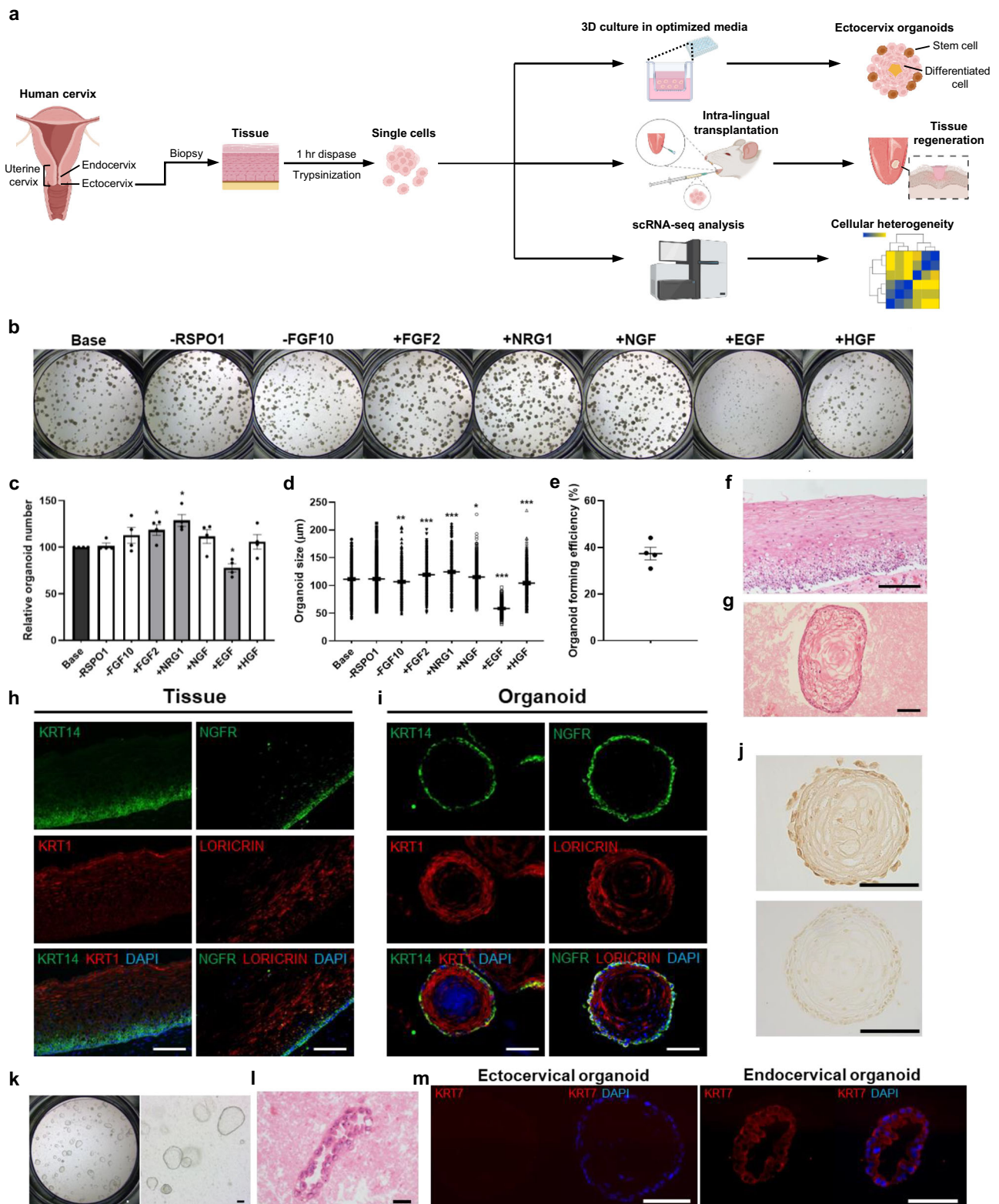
To establish a human ectocervical organoid system with high efficiency from dissociated single cells, we obtained tissue samples from HPV-uninfected, thirteen patients with uterine myoma and nine patients with adenomyosis who underwent hysterectomy. We isolated the normal ectocervical epithelia and dissociated them into single cells. Then, we tested a panel of media conditions to find an optimal composition containing the minimal components necessary to efficiently generate ectocervical organoids displaying a stratified squamous epithelium-like histology. Addition or removal of diverse factors starting from our basal squamous organoid culture media led us to customize the ectocervical organoid culture media to maximize organoid-forming efficiency both in number and size (Fig. 1b–d). Addition of fibroblast growth factor 2 (FGF2), neuregulin 1 (NRG1), nerve growth factor (NGF), and hepatocyte growth factor (HGF) increased either the number or size of human ectocervical organoids, while addition of epidermal growth factor (EGF) decreased the number and size of organoids. On the other hand, R-spondin 1 or FGF10 did not appear to have any significant effect. Under these optimized media conditions, human ectocervical organoids grow and reach a diameter of up to 500 μ m within 10–14 days (Supplementary Fig. 1a). The organoid-forming efficiency reached about 40% (Fig. 1e), and we achieved long-term culture and passaging over at least 12 passages and 6 months (Supplementary Fig. 1b). Recapitulating the stratified squamous epithelium of the human ectocervix, our ectocervical organoids exhibited a tissue organization resembling the concentric rings of a tree (Fig. 1f, g). Our ectocervical organoids expressed the basal markers KRT14 and NGFR in the outer rim, while expressing the differentiated parabasal and suprabasal markers KRT1 and Loricrin in the internal layers (Fig. 1h, i). Of note, Ki67 was detected in the outer rim of organoids, suggesting robust cell proliferation at the periphery (basal layer) of the ectocervical organoids (Fig. 1j and Supplementary Fig. 1c).

We also established human endocervical organoids (Fig. 1k, l). In contrast to the human ectocervix, the human endocervix is lined by the simple columnar epithelium. Thus, the endocervical organoids displayed a cystic shape like other columnar epithelial organoids such as bronchial organoids, while the ectocervical organoids displayed an internally solid shape (Fig. 1k and Supplementary Fig. 1a). Ectocervical and endocervical organoids preferentially expressed ectocervix- or endocervix-specific genes TP63 and KRT7, respectively, at the protein level and *KRT1*, *KRT5*, *KRT13*, *KRT14*, *KRT15* (ectocervix), and *PAX8*, *MUC5B* (endocervix) at the transcript level (Fig. 1m and Supplementary Fig. 1d, e).

These data indicate the successful establishment of highly efficient human ectocervical organoids that recapitulate the human ectocervical epithelium and reflect cervical stem cells' activity.

Establishment of an intralingual transplantation system to examine tissue reconstitution by stem cells and organoids

Next, we aimed to establish a transplantation system to demonstrate in vivo self-renewal capability of stem cells and organoids and employ this method to identify human cervical stem cells. Currently, subrenal capsule transplantation in immunocompromised mice is widely used to demonstrate tissue reconstitution by organoids or stem cells when orthotopic implantation is not feasible. However, the kidney capsule is fragile, and the invasive procedure requires substantial surgical skills. Thus, we aimed to develop an easier and more broadly accessible system for stem cell and organoid transplantation. We focused our efforts on the mouse tongue, as it has sufficient blood flow and potential to hold the implanted cells within a limited space, simplifying subsequent identification.



Therefore, we first implanted human cervical organoids into the tongues of immunocompromised mice to test whether they can regenerate ectocervical stratified squamous epithelium *in vivo*. We removed Matrigel from the cervical organoids when their size reached 50–70 μm , mixed organoid-containing media with growth factor-reduced Matrigel at a 1:1 ratio, and injected 50 μl into the muscular portion of the tongue through the right antero-lateral side of the tongues of 1–12 months old NOD/SCID/IL2R $^{\text{null}}$ (NSG) mice (Fig. 2a, b).

Inflation of the tongues was obvious upon implantation. After 4 weeks, regenerated tissues at the implanted sites were still visible (Fig. 2c), and we identified stratified squamous epithelium-like structures within the glossal muscles (Fig. 2d and Supplementary Fig. 2a). Regenerated tissues express KRT14 and HLA-ABC, confirming that the stratified squamous epithelium-like structures were derived from human ectocervical organoids (Fig. 2e). We also asked whether ectocervical cells in single-cell suspension could generate stratified squamous epithelium-

Fig. 1 | The establishment and optimization of long-term, three-dimensional cultures of human ectocervical and endocervical organoids. **a** Schematic representation of experimental design. Created in BioRender. Myeong, J. (2025) <https://BioRender.com/s27e866>. **b–d** Optimization of organoid culture medium. Representative bright-field organoid images of each indicated culture condition (**b**). Relative numbers (**c**) and size (**d**) of organoids in each culture condition. All data in **c** and **d** are collected from four biological replicates and presented as mean ± SEM ($*p < 0.05$; $**p < 0.01$; $***p < 0.001$). *P*-values were calculated by two-tailed unpaired Student's *t*-test with the control. Source data and exact *p*-values are provided as a Source Data file. **e** Organoid forming efficiency in our cervical organoid culture

medium. Mean ± SEM of four biological replicates are shown. **f** H&E staining of human ectocervical epithelium. **g** H&E staining of human ectocervical organoid. **h, i** Immunofluorescence staining of human ectocervical epithelium (**h**) and organoids (**i**). KRT14 and NGFR were used as markers for basal cells, and KRT1 and Loricrin were used as markers for suprabasal cells. **j** Immunohistochemistry staining of human ectocervical organoids for Ki67 (up) and isotype matched control (down). **k** Representative bright-field images of endocervical organoids. **l** H&E staining of human endocervical organoid. **m**, Immunofluorescence staining of human ectocervical and endocervical organoids. KRT7 was used as the marker for endocervical columnar cells. Scale bars: 100 μm.

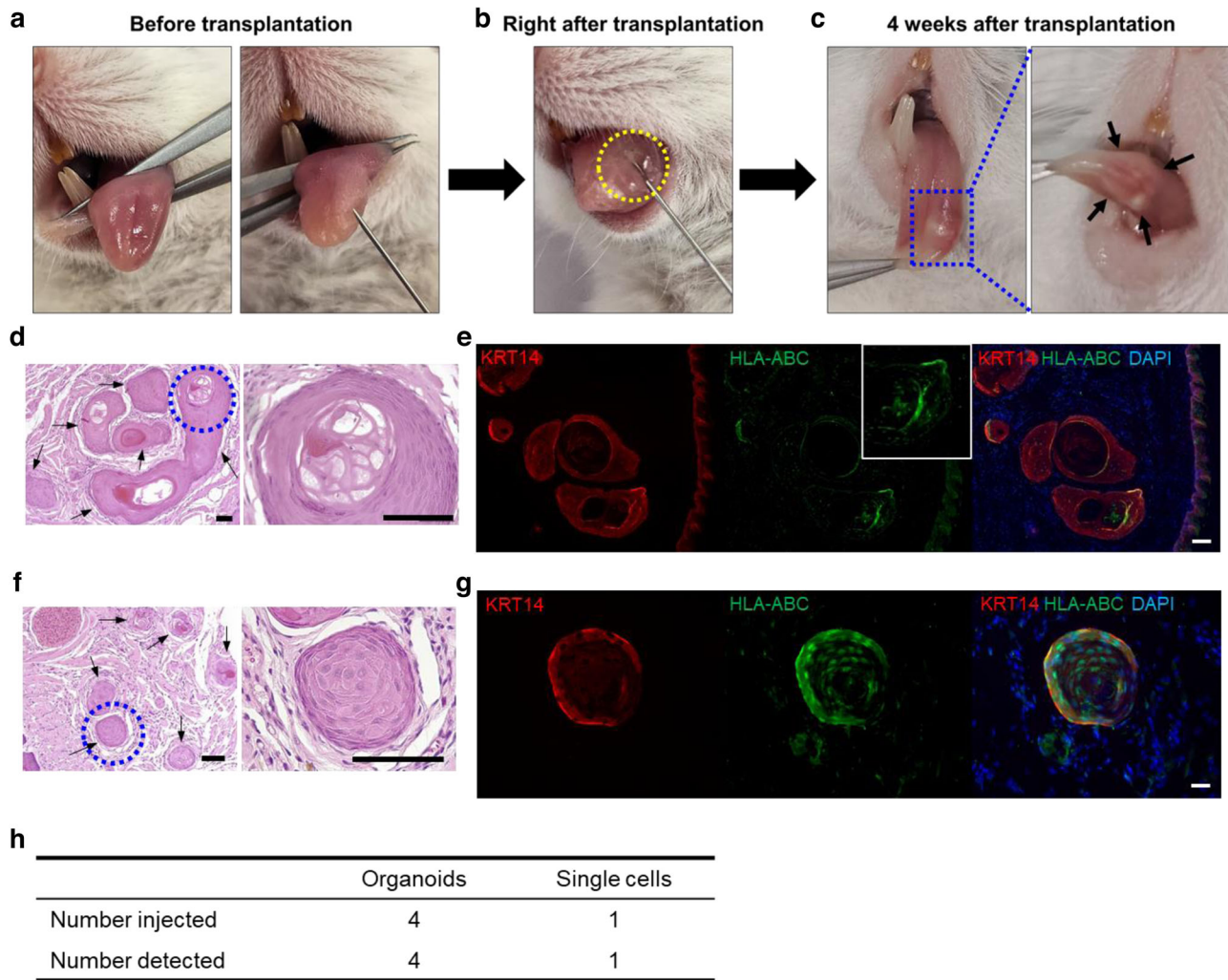


Fig. 2 | Development of an intralingual transplantation approach for assessing tissue reconstitution by organoids and stem cells. **a** Representative images showing the implantation of human ectocervical organoids in murine tongue tissue. **b** Representative gross appearance of the organoid-transplanted sites in the murine tongues just after the implantation. **c** Representative gross appearance of the organoid-transplanted sites in the murine tongues at 4 weeks after implantation. **d** H&E staining of the human tissues regenerated from the cervical organoids transplanted in the murine tongues displaying a stratified squamous epithelium histology. **e** Immunofluorescence staining of the human tissues regenerated from cervical organoids with anti-HLA-ABC and anti-KRT14 antibodies. Note the human-

derived (HLA-ABC-positive), epithelial (KRT14-positive) tissues formed between the murine tongue musculature. **f** H&E staining of the human tissues regenerated from ectocervical single cells transplanted in murine tongues, displaying a stratified squamous epithelium histology. **g** Immunofluorescence staining of the human tissues regenerated from ectocervical single cells with anti-HLA-ABC and anti-KRT14 antibodies. **h** Frequency of human ectocervical tissue regeneration from human ectocervical organoids or single cells when transplanted inside immunocompromised murine tongues. Tongues were harvested at 3–4 weeks after transplantation. Scale bars: 100 μm.

like structures upon intralingual transplantation and found that indeed KRT14-positive and HLA-ABC-positive stratified squamous epithelium-like structures were formed amidst the murine tongue musculature upon injection of cells in single-cell suspension (Fig. 2f,g and Supplementary Fig. 2b). Collectively, we detected stratified squamous

epithelium-like structures within the mouse glossal muscles from all 5 trials (Fig. 2h). These data indicate that our intralingual transplantation system can be used to test the tissue reconstitution capability of organoids and individual stem cells, providing a method that could be generalized to other types of stem cells and organoids in future work.

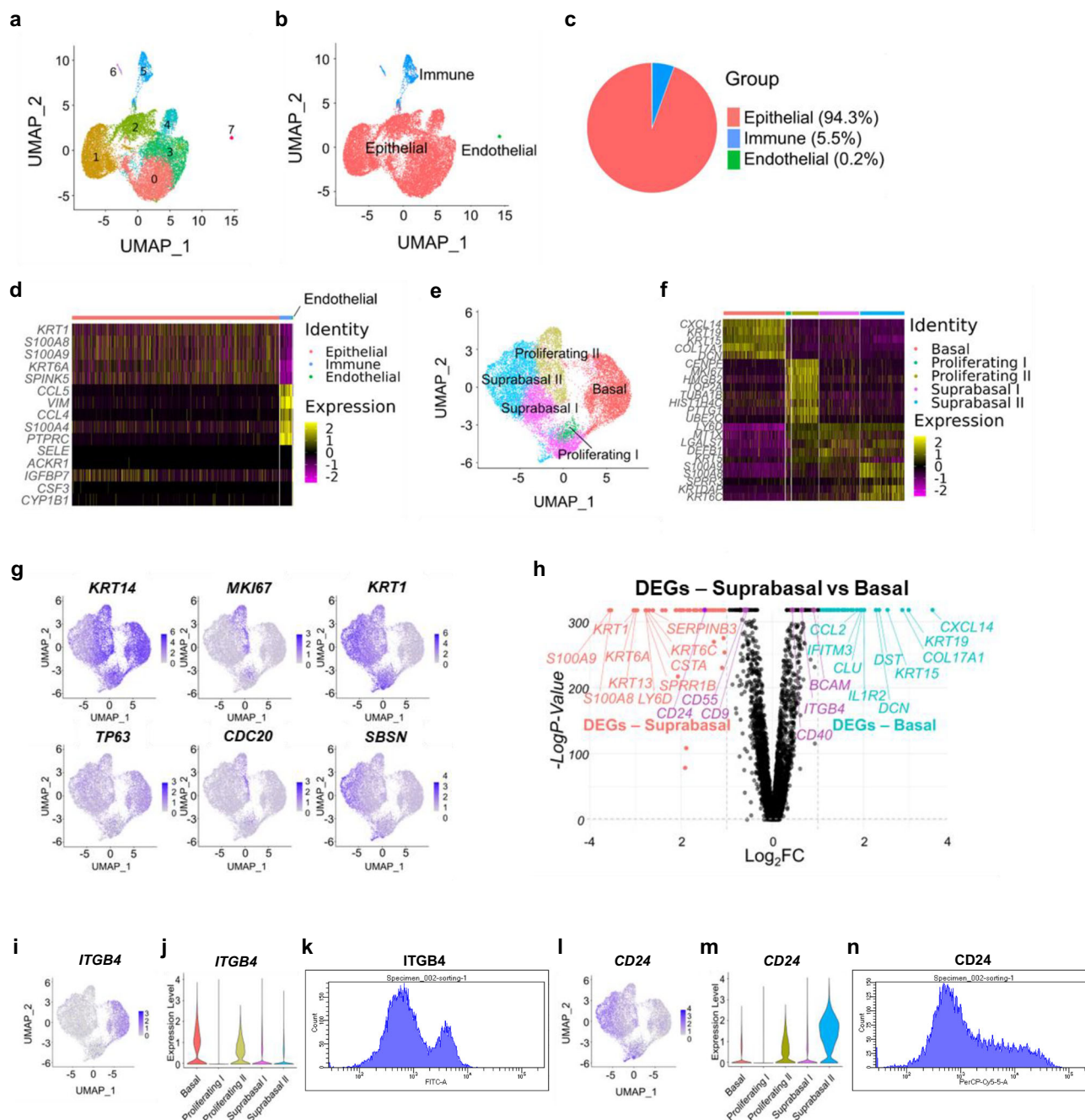


Fig. 3 | scRNA-seq analysis of the human ectocervical epithelium. a UMAP of 19,172 ectocervix tissue cells in 8 clusters. **b** Cell type annotation with red for epithelial cells, green for endothelial cells, and blue for immune cells. **c** Pie chart indicating the proportion of cells belonging to each cluster. **d** Heatmap of top highly-expressed genes in each cell type. The color scale represents distribution of z-score -2 (purple) to 2 (yellow) with black denoting 0. **e** UMAP representation of the epithelial compartment, 18,083 cells from the analysis in **b**. 5 epithelial cell types are denoted in different colors. **f** Heatmap of top highly-expressed genes in each epithelial cell type. The color scale is the same as in **d**. **g** Feature plots of six representative marker genes (*KRT14* and *TP63* for basal cells, *MKI67* and *CDC20* for proliferative basal cells, and *KRT1* and *SBSN* for suprabasal cells) in each epithelial

cell type. **h** Volcano plot showing differentially-expressed genes between basal cells and suprabasal cells. Genes upregulated in basal cells and suprabasal cells are colored by blue and red, respectively. FC stands for fold change. Two-sided Fisher's exact test was employed using negative binomial dispersions. **i** Feature plot showing that *ITGB4* is expressed mainly in the basal cluster. **j** Violin plot comparing the expression of *ITGB4* in each cell cluster. **k** Histogram of flow cytometric analyses of *ITGB4* differentially expressed in the basal cell clusters. **l** Feature plot showing that *CD24* is expressed mainly in the suprabasal cluster. **m** Violin plot comparing the expression of *CD24* in each cell cluster. **n** Histogram of flow cytometric analyses of *CD24* differentially expressed in the basal cell clusters.

Single-cell transcriptome profiling to characterize the cellular heterogeneity of the human ectocervical epithelium

Building on our success establishing human cervical organoid and intraligular transplantation systems to evaluate the self-renewal of cervical stem cells, we sought to delineate the cellular heterogeneity of the human ectocervical epithelium and identify the potential stem cell

subpopulation. After isolating the normal ectocervical epithelia and dissociated them into single cells, we performed single-cell mRNA sequencing (scRNA-seq) (Fig. 1a). After stringent cell filtration and analysis, a total of 19,172 single cells were retained, classified into 8 clusters on Uniform Manifold Approximation and Projection (UMAP), and annotated with 3 cell types based on the expression matrix of cell

type-specific canonical marker genes (Fig. 3a, b). Epithelial cells represented 94.32% of cells (18,083 cells) (Fig. 3c). Immune cells and endothelial cells constituted 5.52% (1058 cells) and 0.16% (30 cells), respectively. Analysis of the differentially expressed genes (DEGs), including upregulation of cell type-specific markers such as *KRT1* and *S100A8* (epithelial cells), *CCL5*, and *IL32* (immune cells), and *SELE* and *CSF3* (endothelial cells), revealed a clear separation of each cell type (Fig. 3d and Supplementary Fig. 3a). UMAP distribution of cell type specific markers such as *EPCAM* (epithelial cells), *PTPRC* (immune cells), *PECAMI* (endothelial cells), *COL1A1* (fibroblasts), and *MSLN* (mesothelial cells) further supported the cell type annotation (Supplementary Fig. 3b). Interestingly, we were not able to identify clusters for fibroblasts or mesothelial cells, likely because we peeled off the ectocervical epithelium from the left-over underlying tissues during the isolation steps.

We performed an independent analysis of the epithelial cell compartment to further characterize its heterogeneity. Re-clustering using UMAP designated the epithelial cells into 5 transcriptionally distinct epithelial subclusters, which were separately grouped in the heat map based on the expression matrix of top ranked DEGs: basal cells (*KRT15*, *KRT19*, and *CXCL14*), proliferating cells I and II (*TOP2A*, *CENPE*, and *MKI67*), suprabasal cells I (*MTIX*), and suprabasal cells II (*S100A8*, *S100A9*, and *KRTDAP*) (Fig. 3e, f and Supplementary Data 1). This annotation was supported by UMAP plots showing the distribution of representative cell type-specific genes of basal (*KRT14* and *TP63*), proliferating (*MKI67* and *CDC20*), and suprabasal cells (*KRT1* and *SBSN*) (Fig. 3g)^{26–29}.

In the stratified squamous epithelium, tissue stem cells stay within or close to the basal layer and differentiate upward to the suprabasal layers^{30–32}. To characterize distinct genetic programs between basal cells and suprabasal cells, we compared the gene expression profiles of the basal cell cluster with the suprabasal cell cluster. DEGs include *CXCL14*, *KRT19*, *COL17A1*, *KRT15*, and *DST*, which are upregulated in basal cells, and *S100A9*, *S100A8*, *KRT1*, *KRT6A*, and *KRT13*, which are upregulated in suprabasal cells (Fig. 3h and Supplementary Fig. 3c, d). We further analyzed cluster-specific cell surface markers that could be potentially used to isolate the cervical stem cells and differentiated cells. While *ITGB4*, *BCAM*, *NGFR*, and *CD40* were upregulated in the basal cluster, *CD24*, *CD9*, *CD40*, and *CD55* were upregulated in the suprabasal cluster (Fig. 3h–j, l, m and Supplementary Fig. 3e, f, h, i). Among these cell surface markers, each of *ITGB4* and *CD24* most clearly separated cervical epithelial cells into two distinct subpopulations when analyzed by flow cytometry (Fig. 3k, n and Supplementary Fig. 3g, j).

This analysis catalogs the cellular heterogeneity of the human ectocervical epithelium and suggests potential cell surface markers that could be used to identify and purify ectocervical stem and differentiated cells.

Isolation of a population enriched for human cervical stem cells

We next aimed to assess whether *ITGB4* and *CD24*, the two top candidate cell surface markers differentiating the basal and suprabasal cell populations from our scRNA-seq analysis, can be used as specific cell surface markers to enrich human cervical stem cells. To this end, we first examined their in situ expression patterns in patient tissue samples. Immunostaining revealed overlay of *ITGB4* with the basal marker *KRT14* but distinct localization from the suprabasal marker *KRT1* (Fig. 4a and Supplementary Fig. 4a). In contrast, *CD24* stained the *KRT1*-positive suprabasal layer but did not stain the *KRT14*-positive basal layer (Fig. 4a and Supplementary Fig. 4b). Consistently, *ITGB4* and *CD24* exclusively stained the *KRT14*-positive outer and *KRT1*-positive inner layers of human cervical organoids, respectively, suggesting that these two markers can be employed to separate human cervical stem cells and differentiated cells (Fig. 4b and Supplementary Fig. 4c, d).

To test this, we dissociated human ectocervical epithelium into single cells and stained for *ITGB4* and *CD24* prior to analyzing the cells via flow cytometry. This analysis yielded three distinct subpopulations depending on *ITGB4* and *CD24* expression: *ITGB4*⁺*CD24*[−], *ITGB4*[−]*CD24*[−], and *ITGB4*⁺*CD24*⁺ cells (Fig. 4c and Supplementary Fig. 4e). Thus, we sorted these subpopulations and performed diverse tests to identify which subpopulations are enriched for stem cells or differentiated cells. First, we compared the gene expression of representative basal and suprabasal markers. In other contexts, tissue stem cells in the stratified squamous epithelium have been shown to express higher levels of basal markers but lower levels of suprabasal markers than differentiated cells^{18,24}. In our gene expression analyses, *ITGB4*⁺*CD24*[−] cells express the highest levels of basal markers *TP63* and *KRT14* but the lowest levels of suprabasal markers *KRT1* and *KRT13* and cycling genes *CDC20* and *MKI67*, while *ITGB4*[−]*CD24*[−] cells express the highest levels of *KRT1* and *KRT13* but the lowest levels of *TP63* and *KRT14* (Fig. 4d–f).

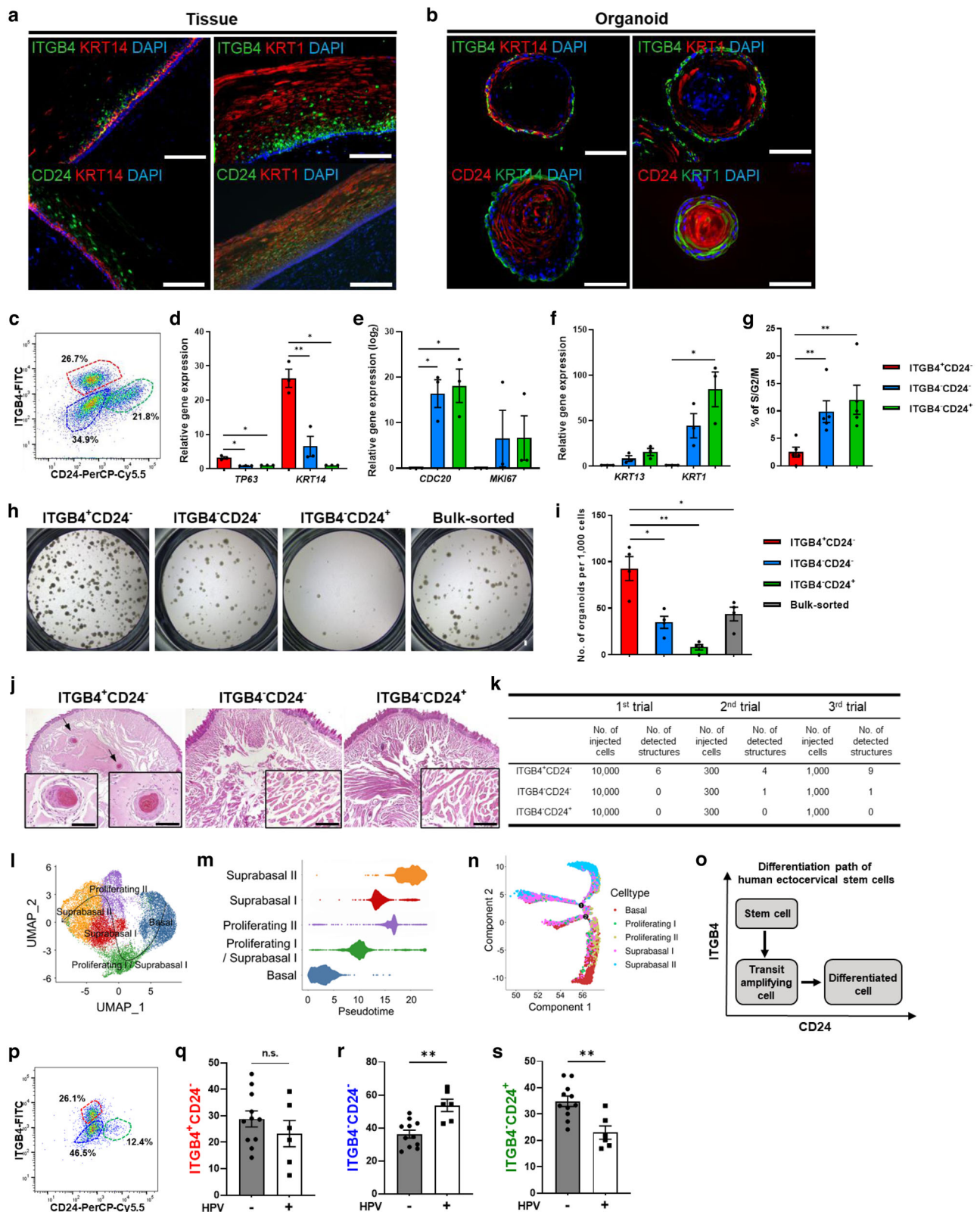
Next, we performed cell cycle analysis. Most tissue stem cells have a characteristically slow cell cycle, although the stem cells of some tissues, like the intestine, are fast cycling^{25,33}. When we measured the ratio of cells in S/G2/M phases, the percentage of *ITGB4*[−]*CD24*[−] cells in S/G2/M phases was ~2%, while ~10% in *ITGB4*⁺*CD24*[−] and *ITGB4*⁺*CD24*⁺ subpopulations, suggesting that *ITGB4*⁺*CD24*[−] cells are slow cycling cells in the human ectocervical epithelium (Fig. 4g and Supplementary Fig. 4f). We further compared the organoid-forming capability of each subpopulation. The *ITGB4*[−]*CD24*[−] subpopulation generated organoids with the highest efficiency (~10% from sorted primary cells), 11-fold more than the *ITGB4*⁺*CD24*⁺ subpopulation that had minimal organoid-generating capacity (Fig. 4h, i). Bulk-sorted cells and the *ITGB4*[−]*CD24*[−] subpopulation generated organoids at an intermediate level.

Finally, we examined which subpopulations of cells have tissue reconstitution capability in vivo, employing our intralingual transplantation system. The *ITGB4*[−]*CD24*[−] subpopulation consistently regenerated multiple stratified squamous epithelium-like tissues from as few as 300 cells, despite the mechanical stress from fluorescence-activated cell sorting (FACS) and the different tissue microenvironment, when transplanted into the tongues of NSG mice (Fig. 4j, k and Supplementary Fig. 4g). We conclude that the *ITGB4*[−]*CD24*[−] cells enrich for human ectocervical stem cells, whereas *ITGB4*⁺*CD24*⁺ cells represent mostly differentiated cells.

Our successful identification of a population enriched for human ectocervical stem cells enabled us to explore their differentiation trajectory. To this end, we employed two different analytical tools, Slingshot and Monocle. Slingshot predicted the transition states of individual cells and suggested that the basal cluster gives rise to suprabasal cluster II through proliferating clusters (Fig. 4l). Consistently, the density plot showed the basal cluster at the earliest time point, followed in order by proliferating cluster I and suprabasal cluster I, proliferating cluster II, and suprabasal cluster II (Fig. 4m). Analysis using Monocle indicated that the pseudotemporal trajectory starts with the basal cluster, proceeds through proliferating I and proliferating II clusters, and ends with suprabasal I and suprabasal II clusters although it divulges into three branches (Fig. 4n). Both analyses support a model in which basal cells (cluster 0) are the stem cells, generating proliferating cells and suprabasal cells. Collectively, these data show that *ITGB4*[−]*CD24*[−] cells identify and enrich human ectocervical stem cells, which differentiate through *ITGB4*[−]*CD24*[−] transit amplifying cells to *ITGB4*[−]*CD24*⁺ differentiated suprabasal cells (Fig. 4o).

HPV infection expands the transit amplifying cell subpopulation

High-risk HPV infection is a major oncogenic event leading to cervical cancers. Although the role of viral genes such as E6 and E7 and the cervical transformation process through LSIL and HSIL into cervical cancers are well-known, the early effects of viral infection



on the normal cervical epithelium are not well understood. Leveraging our identification of cell surface markers for cervical stem cells, transit amplifying cells, and differentiated cells, we aimed to investigate how HPV infection affects the cellular composition of cervical epithelium. Therefore, we used flow cytometry to analyze the cellular composition of cervical epithelia with apparently normal appearance from patients diagnosed with HSIL. Remarkably,

HPV-infected normal cervical epithelia contained higher proportions of ITGB4⁺CD24⁻ transit amplifying cells and lower proportions of ITGB4⁺CD24⁺ differentiated cells than HPV-uninfected normal cervical epithelia, with no difference in the proportions of ITGB4⁺CD24⁻ stem cells (Fig. 4p–s). These data reveal the early stage effect of HPV infection in cervical epithelia, which causes expansion of transit amplifying cells.

Fig. 4 | Identification of the cervical cell subpopulation enriched for human cervical stem cells. **a** Immunofluorescence staining of ITGB4 and CD24 in human cervical epithelium. Scale bars: 100 μ m. **b** Immunofluorescence staining of ITGB4 and CD24 in human cervical organoids. **c** Flow cytometry analysis of primary human cervical epithelial cells using ITGB4 and CD24. Expression of basal markers (**d**), cell cycle genes (**e**), and suprabasal markers (**f**) in each subpopulation assessed by qPCR ($n = 3$). **g** Cell cycle analysis of each subpopulation. The percentage of cells in S/G2/M phases in each subpopulation was measured ($n = 5$). Cervical organoids generated from each subpopulation and bulk sorted cells. Representative bright-field images (**h**) and number (**i**) of organoids ($n = 4$). **j** Representative H&E staining of the human stratified squamous epithelial tissues regenerated from each subpopulation of human ectocervical cells transplanted inside immunocompromised murine tongues. **k** Frequency of human ectocervical tissue regeneration from each subpopulation of cells. Transplanted murine tongues were harvested at 3–4 weeks

after transplantation. **l** Pseudo-temporal trajectory plot showing the order of cell transitions at the cluster level. The black line starts at basal cluster and ends at suprabasal cluster II. **m** Density dot plot showing the number of cells in each cluster over time. **n** Single-cell trajectory reconstructed by Monocle for epithelial cells. **o** Schematic model of human cervical stem cell differentiation. **p** Representative flow cytometry analysis of primary HPV-infected human cervical epithelial cells using ITGB4 and CD24. Comparison of the ratios of ITGB4⁺CD24⁺ (**q**), ITGB4⁺CD24⁺ (**r**), ITGB4⁺CD24⁺ (**s**) cells between six HPV-infected and eleven HPV-uninfected normal cervical epithelia. All data are collected from indicated biological replicates and presented as mean \pm SEM (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). P -values were calculated by one-way ANOVA with Tukey's multiple comparison test (**d–g**, **i**) or two-tailed unpaired Student's t -test (**q–s**). Source data and exact p -values are provided as a Source Data file. Scale bars: 100 μ m.

ITGB4 and CD24 also mark the cervical stem cell population in mice

In some tissues such as tracheal epithelium, the same basal stem cell markers are applicable to both human and mouse tissue²⁵, while human and murine esophageal stem cells have different cell surface markers²⁴. Therefore, we explored whether ITGB4 and CD24 can also distinguish murine ectocervical stem cells and differentiated cells. We first re-analyzed a publicly available murine cervical scRNA-seq dataset (NCBI GEO GSE128987). Subsetting and an independent analysis of the epithelial cells resulted in 4 clusters of murine ectocervical epithelial cells: basal, proliferating, early suprabasal, and late suprabasal cell clusters (Fig. 5a, b). This cell type annotation is supported by UMAP plots showing the distribution of representative cell type specific genes of basal (*Krt15* and *Trp63*), proliferating (*Mki67* and *Cdc20*), and suprabasal cells (*Krt1* and *Sbsn*) (Fig. 5c). Of note, *Itgb4* is highly expressed in the basal cell cluster, while *Cd24* is highly expressed in the early suprabasal cell cluster (Fig. 5d, e and Supplementary Fig. 5a, b). Furthermore, in murine cervical epithelium, ITGB4 is exclusively expressed in the KRT5-positive basal layer while CD24 is expressed in the LORICRIN-positive suprabasal layer (Fig. 5f and Supplementary Fig. 5c), suggesting that murine ectocervical stem and differentiated cells can be purified by the same cell surface markers as human ectocervical stem and differentiated cells: ITGB4 and CD24.

We next established murine ectocervical organoids. Ectocervical cells were dissociated into single cells from 1–6 months old C57BL/6 female mice for organoid culture. Addition of R-spondin 1 and removal of FGF2, NRG1, and NGF from our human ectocervical organoid culture media allowed the culture of murine ectocervical organoids. Our murine ectocervical organoids grew up to 400 μ m in diameter over 2 weeks (Fig. 5g and Supplementary Fig. 5d) and could be passaged at least twice (Supplementary Fig. 5e). These organoids displayed stratified squamous epithelium-like histology similar to in vivo murine cervical epithelium (Fig. 5h, i and Supplementary Fig. 5f) and expressed ITGB4 and CD24 genes in the same pattern as we observed in the in vivo murine cervical epithelium (Supplementary Fig. 5g).

Thus, we analyzed murine ectocervical epithelial cells by flow cytometry using antibodies for ITGB4 and CD24 to determine whether they separate the murine cervical epithelial cells into distinct subpopulations. Indeed, ITGB4 and CD24 antibody staining divided murine ectocervical epithelial cells into three subpopulations, just like the human ectocervical epithelial cells (Fig. 5j and Supplementary Fig. 5h). Subsequent sorting and organoid culture led to organoid formation from ITGB4⁺CD24⁺ cells with the highest efficiency, 7-fold more than ITGB4⁺CD24⁺ cells (Fig. 5k, l), mirroring what we previously observed with the analogous human cell subpopulations. Taken together, these data show that ITGB4 and CD24 enrich for ectocervical stem cells in both humans and mice, suggesting broader species applicability.

PI3K-AKT pathway regulates the self-renewal of the human cervical stem cells

To explore signaling pathways that regulate cervical stem cell function, we analyzed our scRNA-seq data to identify pathways or genes enriched in ectocervical basal cells compared to suprabasal cells. Gene set enrichment analysis (GSEA) revealed unique transcriptional features and enriched pathways relevant to differentiation status (Fig. 6a, b and Supplementary Fig. 6a–d). For example, gene sets for epithelial mesenchymal transition and DNA repair were significantly enriched in the basal cluster, while gene sets for keratinization, tight junction interactions, cell-cell junction organization, and gap junction trafficking and regulation were enriched in the suprabasal clusters. Of note, GSEA of gene ontology (GO) biological process (BP) showed that the phosphoinositide 3-kinase (PI3K) signaling pathway is a gene set enriched in basal cells (Fig. 6a, b). The PI3K-AKT signaling pathway has been shown to regulate the quiescence and activation of various epithelial stem cells, including intestinal, prostate, and bronchioalveolar stem cells^{34–36}.

Therefore, we further investigated whether the PI3K-AKT pathway also regulates the self-renewal of cervical stem cells. We first examined expression of phospho-AKT1 (pAKT1) in human cervical epithelium and organoids. We detected pAKT1 expression in the basal and lower parabasal layers but not in the suprabasal layers (Fig. 6c and Supplementary Fig. 6e). Consistently, pAKT1 was expressed in the outer layer of cervical organoids where basal cells are located (Fig. 6d and Supplementary Fig. 6f). In addition, expression of the PI3K-AKT pathway genes was significantly higher in the ITGB4⁺CD24⁺ human stem cell subpopulation than ITGB4⁺CD24⁺ transit amplifying cell and ITGB4⁺CD24⁺ differentiated cell subpopulations (Fig. 6e). These data indicate that the activity of the PI3K-AKT pathway is more robust in basal cells than in other cell subpopulations, suggesting that the PI3K-AKT signaling pathway regulates self-renewal of human cervical stem cells.

To directly test this possibility, we interrogated whether modulation of the PI3K-AKT pathway affects cervical organoid formation. Indeed, treatment with a PI3K-AKT signaling pathway inhibitor, LY294002, reduced cervical organoid formation (Fig. 6f, g). Conversely, a PI3K-AKT pathway activator, YS-49, increased cervical organoid formation (Supplementary Fig. 6g, h). These data indicate that the PI3K-AKT pathway regulates cervical stem cell self-renewal.

To gain insight into the global transcriptional response elicited by the inhibition of the PI3K-AKT pathway, we performed bulk RNA sequencing on control and LY294002-treated cervical cells. Unsupervised hierarchical clustering showed clear separation between control and LY294002-treated cervical cells (Fig. 6h). 774 out of 10,826 genes in our dataset were differentially regulated by LY294002 treatment (Fig. 6i, j). GSEA for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways revealed that LY294002 treatment significantly depleted gene sets involved in metabolism and cell-extracellular matrix interaction, such as PPAR signaling pathway and focal adhesion

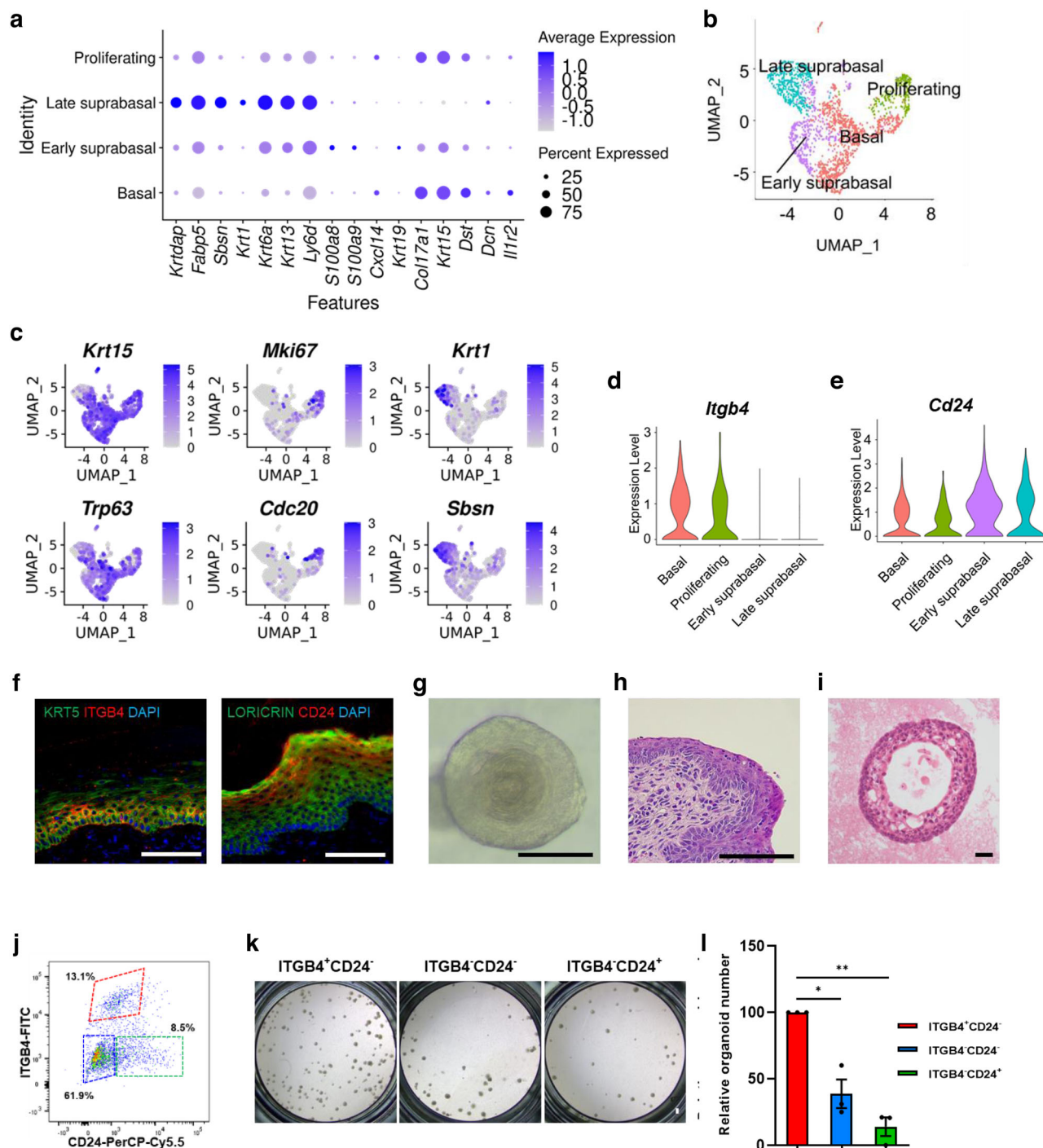


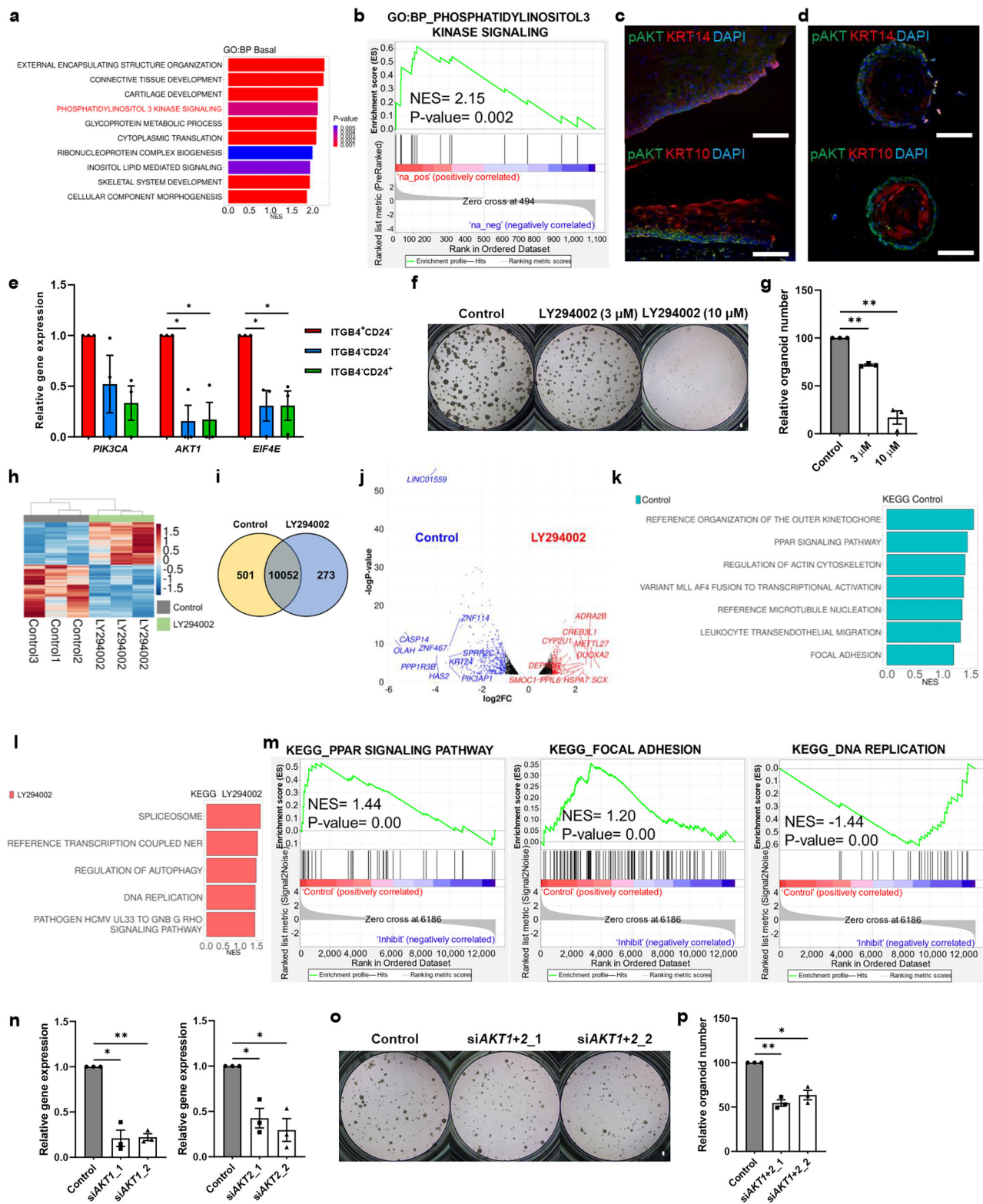
Fig. 5 | ITGB4 and CD24 can be used across species to purify mouse ectocervical stem cells. **a** Dot plot showing the expression of the representative basal and suprabasal genes from previously published murine cervical scRNA-seq dataset (NCBI GEO GSE128987). **b** UMAP of epithelial cell clusters from murine cervical scRNA-seq dataset. **c** Feature plots showing six representative marker genes (*Krt15* and *Trp63* for basal cells, *Mki67* and *Cdc20* for proliferative basal cells, and *Krt1* and *Sbsn* for suprabasal cells) in each epithelial cell type. **d** Violin plot comparing the expression of *Itgb4* in each cell cluster. **e** Violin plot comparing the expression of *Cd24* in each cell cluster. **f** Immunofluorescence staining of ITGB4 and CD24 in murine cervical epithelium. **g** Representative bright-field image of

murine ectocervical organoids. **h** H&E staining of murine cervical epithelium. **i** H&E staining of murine ectocervical organoids. **j** Flow cytometric analysis of murine ectocervical epithelium by CD104 and CD24. **k**, **l** Cervical organoids generated from each subpopulation. Representative bright-field images (**k**) and relative numbers (**l**) of organoids when organoid numbers of ITGB4⁺CD24⁻ cells were set to 100. Data are collected from 3 biological replicates and presented as mean ± SEM (**p* < 0.05; ***p* < 0.01). *P*-values were calculated by one-way ANOVA with Tukey's multiple comparison test. Source data and exact *p*-values are provided as a Source Data file. Scale bars: 100 μm.

(Fig. 6k, m). Conversely, LY294002 treatment significantly enriched gene sets involved in the cell cycle such as DNA replication (Fig. 6l, m). GSEA for oncogenic signature showed a significant enrichment of gene sets such as KRAS lung up and PTEN down in control cells compared to

LY294002-treated cells, confirming the inhibition of the PI3K-AKT pathway by LY294002 treatment (Supplementary Fig. 6i, j).

Finally, as an orthogonal measure, we tested the effect of genetic interruption of the PI3K-AKT pathway on cervical organoid



formation. Silencing of *AKT1* and *AKT2* reduced cervical organoid formation (Fig. 6n–p). In contrast, silencing of phosphatase and tensin homolog (*PTEN*) increased cervical organoid formation (Supplementary Fig. 6k–m), consistent with the data from pharmacologic inhibition and activation of the PI3K-AKT pathway. Taken together, our data indicate that the PI3K-AKT signaling pathway regulates the self-renewal and differentiation of human cervical stem cells.

Network analyses reveal key intercellular interactions in human ectocervix

We next investigated the intercellular network among different cell types in the human ectocervix. To this end, we first analyzed our scRNA-seq data with the CellChat R package, which predicts receptor-ligand interactions based on gene expression profiles across different cell clusters. Intriguingly, we found that endothelial cells exhibited the highest number and strength of interactions with other cell types

Fig. 6 | PI3K-AKT signaling pathway regulates the self-renewal of human cervical stem cells. **a** GSEA bar graph representing the top-ranked GO biological process gene sets enriched in basal cells. Two-sided Fisher's exact test was employed using negative binomial dispersions. **b** GSEA enrichment plot showing that PI3K pathway-related gene sets are enriched in basal cells compared to suprabasal or proliferating cells. **c** Immunofluorescence staining of phosphoAKT1 in human ectocervical epithelium. Scale bars: 100 μ m. **d** Immunofluorescence staining of phosphoAKT1 in human ectocervical organoids. **e** Relative gene expression of genes in the PI3K-AKT pathway ($n = 3$). **f, g** The effect of a PI3K-AKT pathway inhibitor, LY294002, on cervical organoid formation. Representative bright-field images (**f**) and bar graph for relative numbers (**g**) of the cervical organoids generated ($n = 3$). **h** Heat map of control and LY294002-treated human ectocervical organoids. **i** Venn diagram of DEGs between control and LY294002-treated cervical cells. **j** Volcano plot showing DEGs in control and VP-treated

organoids. Running sum statistics with permutation test ($n = 1000$) was employed. **k, l** GSEA bar graph representing the top-ranked KEGG pathway gene sets enriched in control cervical cells (**k**) compared to LY294002-treated cervical cells (**l**) ($p < 0.05$). **m** GSEA enrichment plots of representative differential gene sets from bulk RNA-seq data. Two-sided Fisher's exact test was employed using negative binomial dispersions (**k–m**). **n** Expression of *AKT1* and *AKT2* from control or *AKT1*- or *AKT2*-silenced cervical cells assessed by qPCR ($n = 3$). **o, p** The effect of *AKT* silencing on cervical organoid formation. Representative bright-field images (**o**) and bar graph for relative numbers (**p**) of cervical organoids ($n = 3$). All data are collected from 3 biological replicates and presented as mean \pm SEM ($^*p < 0.05$; $^{**}p < 0.01$). 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 Fig. 7a, b). Pathway rank analysis revealed the top-ranked signals and pathways in the whole human ectocervix, most of which are well-known cell signaling pathways in immune or endothelial cells, such as CCL, PECAM1, and VEGF (Supplementary Fig. 7c). The top-scoring signal was CC chemokine ligands (CCL), which interconnects endothelial cells and immune cells (Supplementary Fig. 3a and 7c, d). Focusing next on interactions within the epithelial cell compartment, we found that Notch signals are highly ranked after basal lamina-related signals such as desmosome and laminin (Supplementary Fig. 7e). More detailed analysis suggested that the Notch signal mediates communication between the proliferating type II cells and suprabasal type II cells; JAG1 and NOTCH3 were predicted to be the strongest ligand-receptor pair between those cell types (Supplementary Fig. 7f, g). We therefore examined whether Notch perturbation elicits changes in cervical organoid formation and thus possibly cervical epithelial integrity. Indeed, treatment with the Notch inhibitor DAPT interferes with cervical organoid formation (Supplementary Fig. 7h, i). Altogether, our analysis highlights key pathways mediating intercellular interactions in the human ectocervix.

YAP1 regulates the self-renewal of the human cervical stem cells

Our network analysis revealed that desmosome and laminin signals are top-ranked intercellular signals within the epithelial compartment (Supplementary Fig. 7e). Desmosomes are major intercellular adhesive junctions between epithelial cells, and major pathways mediating downstream desmosome signals include the RAS/RAF/MAPK pathway, Hippo pathway, PI3K-AKT pathway, and Wnt/ β -catenin pathway^{37,38}. Laminins are one of the major components in the basement membrane and interact with integrins or dystroglycans to relay their signals^{39,40}. One of the major functions of integrin is sensing and mediating the force, rigidity, and ligands of the extracellular matrix, which regulates the nuclear-cytoplasmic distribution of transcription regulators such as Yes-associated protein 1 (YAP1), the effector of the Hippo signaling pathway⁴¹. In addition, the Hippo signaling pathway and YAP1 have been shown to regulate the self-renewal and differentiation of diverse epithelial stem cells, including intestinal, tracheal, esophageal, and epidermal stem cells^{42–45}. This led us to ask whether YAP1 and the Hippo signaling pathway also regulate the self-renewal and differentiation of cervical stem cells. We first investigated whether the genetic programs of the Hippo signaling pathway and YAP1 are differentially expressed between basal cells (containing the stem cell population) and suprabasal cells. In our scRNA-seq data analysis, the Hippo signaling pathway enrichment score was significantly lower in the basal cell cluster than suprabasal cell cluster, while the YAP downstream gene enrichment score was significantly higher in the basal cell cluster (Supplementary Fig. 7j,k). Also, the expression of most of the 20 representative YAP1 target genes⁴⁶ was higher in the basal cell cluster (Supplementary Fig. 7l). Consistently, immunostaining disclosed that nuclear YAP1 expression was detected only at the basal layer and lower parabasal layers (Supplementary Fig. 8a). YAP1 was expressed in the outer layer of ectocervical

organoids where the basal cells are located (Supplementary Fig. 8b). Also, gene expression of *YAP1* and its target genes, *FSTL1* and *CYR61*, was significantly higher in the ITGB4⁺CD24⁺ human stem cell subpopulation than ITGB4⁺CD24⁺ transit amplifying cell and ITGB4⁺CD24⁺ differentiated cell subpopulations (Supplementary Fig. 8c). These data indicate that the expression and activity of YAP1 are higher in stem cells than in transit amplifying cells and differentiated cells, suggesting that YAP1 plays important roles in the self-renewal of human cervical stem cells.

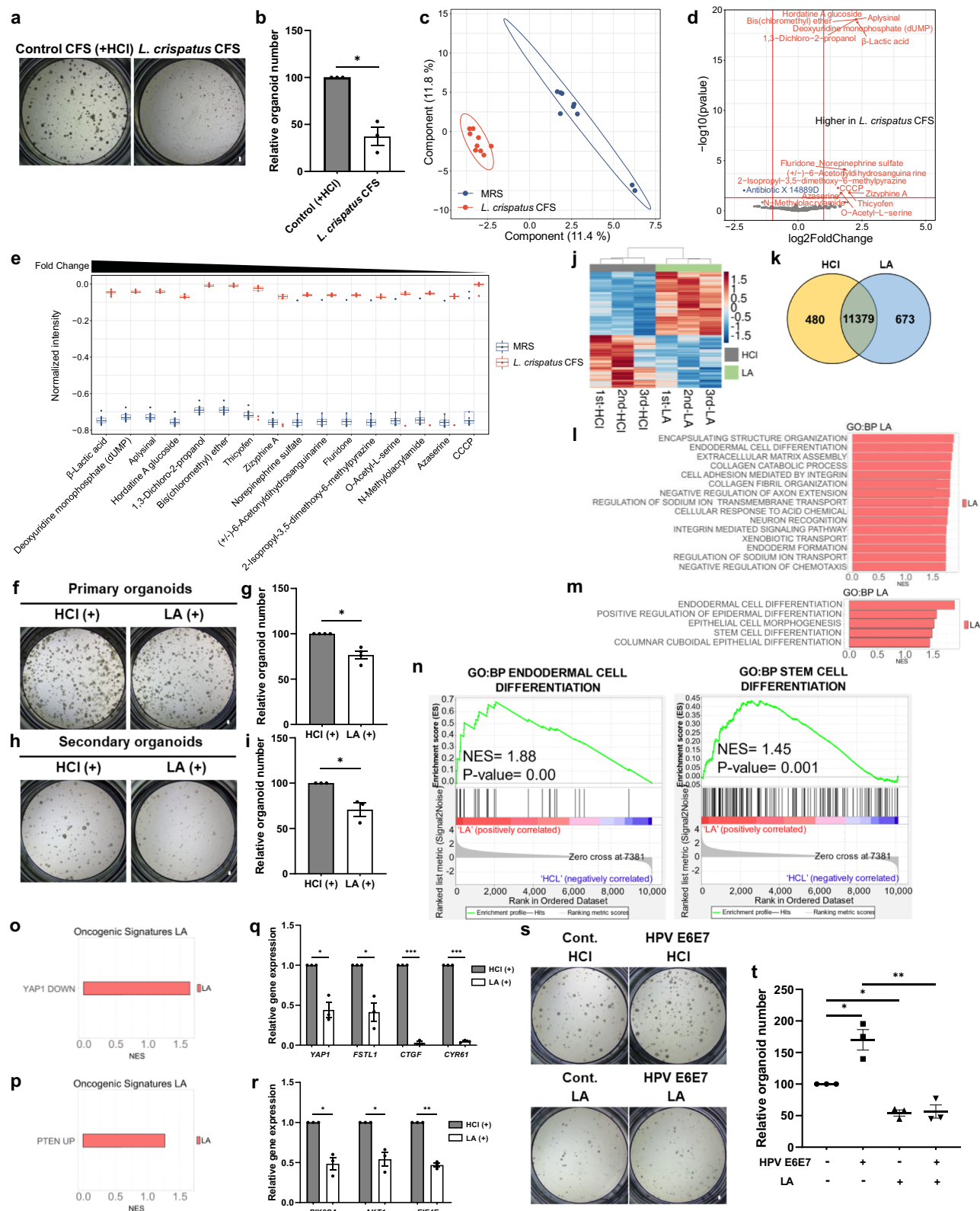
Primary organoid formation and secondary organoid formation after passaging are a readout of the self-renewal capacity of stem cells^{22,23}. Therefore, we aimed to experimentally demonstrate the role of YAP1 in cervical stem cells using organoids. Indeed, treatment with verteporfin (VP), a YAP1 inhibitor, decreased not only primary ectocervical organoid formation but also secondary organoid formation, indicating that YAP1 regulates cervical stem cell self-renewal (Supplementary Fig. 8d–g).

To determine the global transcriptional response induced by YAP1 inhibition, we performed bulk RNA sequencing on these VP-treated and control organoids. As expected, unsupervised hierarchical clustering showed clear separation between control and VP-treated cervical organoids (Supplementary Fig. 8h). 736 out of 13,356 genes in our dataset were differentially regulated by VP treatment (Supplementary Fig. 8i, j). GSEA for GO:BP demonstrated a significant enrichment of the gene sets involved in the cell cycle progression such as tRNA processing, DNA replication initiation, and cell cycle DNA replication in VP-treated group (Supplementary Fig. 8k, m). Conversely, gene sets involved in the interaction between stem cell and stem cell niche such as basement membrane organization and extracellular matrix assembly were significantly depleted in VP-treated group (Supplementary Fig. 8l, m). Given that stem cells detach from the basement membrane and begin cell cycle progression for differentiation in the stratified squamous epithelium, these data further support that YAP1 inhibition induces stem cell differentiation.

Finally, we examined the effect of *YAP1* genetic silencing on cervical organoid formation (Supplementary Fig. 8n). Consistent with the pharmacological inhibition with VP, genetic silencing of *YAP1* using siRNA significantly decreased cervical organoid formation, suggesting a reduction in stem cell activity (Supplementary Fig. 8o, p). Altogether, these data indicate that YAP1 regulates the self-renewal and differentiation of human ectocervical stem cells.

A *Lactobacillus* metabolite regulates the self-renewal of human cervical stem cells

We next aimed to explore the stem cell-extrinsic regulatory factors in the human ectocervix. Normal and pathogenic microflora are known to modulate tissue integrity and disease pathogenesis by direct contact or through microbially-produced metabolites⁴⁷. In the human cervicovaginal tract, the most common bacterial constituent of the local microbiota is *Lactobacillus*, with *L. crispatus* and *L. gasseri* being particularly abundant species⁴⁸. Previous studies including ours suggest a



protective effect of *Lactobacillus* or their isolated metabolites against cervical cancers^{8–10}, but it remains unclear which cervical cell type receives and processes these extrinsic signals to mediate that effect.

We hypothesized that the effect of *Lactobacillus* on cervical cancer could act through cervical stem cells, and so we first examined its role in their self-renewal and differentiation. Since the basal cells reside underneath multiple layers of parabasal and suprabasal cells, we

reasoned that *Lactobacillus* would likely communicate with stem cells mainly through secreted metabolites rather than direct cellular contact. Thus, we examined the effect of *L. crispatus* cell free supernatant (CFS) and *L. gasseri* CFS on cervical organoid formation. Surprisingly, the formation of ectocervical organoids was significantly decreased by the addition of *L. crispatus* CFS or *L. gasseri* CFS compared to control (Fig. 7a, b and Supplementary Fig. 9a, b). In order to identify the

Fig. 7 | A *Lactobacillus* metabolite regulates human cervical stem cell self-renewal. **a, b** The effect of *L. crispatus* CFS on human ectocervical organoid formation. Representative bright-field images (**a**) and bar graph for relative numbers (**b**) ($n = 3$). **c** Partial least squares-discriminant analysis of metabolites. **d** Volcano plot showing differential metabolites between control and *L. crispatus* CFS. **e** Box plot of metabolites more abundant in *L. crispatus* CFS. The center line represents the median with the box spanning the interquartile range (IQR) and the whiskers $1.5 \times$ IQR. Statistically significant metabolites (\log_2 FC > 1.0 , FDR adjusted p -value < 0.05) are labeled, with p -value calculated using the two-tailed unpaired Wilcoxon rank-sum test (**d, e**) ($n = 3$). **f, g** The effect of LA on primary organoid formation. Representative bright-field images (**f**) and bar graph for relative numbers (**g**) ($n = 4$). **h, i** The effect of LA on secondary organoid formation. Representative bright-field images (**h**) and bar graph for relative numbers (**i**) ($n = 3$). **j** Heat map of gene expression in HCl- or LA-treated ectocervical organoids. **k** Venn

diagram of DEGs by LA treatment. Most (**l**) and additionally (**m**) enriched GO:BP gene sets in LA-treated organoids ($p < 0.05$). **n** Representative enrichment plots for gene sets highly enriched in LA-treated organoids. Two-sided Fisher's exact test was employed using negative binomial dispersions (**l–n**). **o** Additional enriched oncogenic signatures of a set of downregulated genes by YAP overexpression⁴⁹ in LA-treated organoids. **p** Additional enriched oncogenic signatures of a set of upregulated genes by PTEN in LA-treated organoids. **q, r** Relative gene expression of YAP1 target genes (**q**) and the PI3K-AKT pathway genes (**r**) of control and LA-treated cervical cells ($n = 3$). The effect of LA treatment and E6E7 overexpression on cervical organoid formation. Representative bright-field images (**s**) and bar graph for relative numbers (**t**) ($n = 3$). Data are collected from indicated biological replicates and presented as mean \pm SEM ($^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$). P -values were calculated by two-tailed unpaired Student's t -test (**b, g, i, q, r, t**). Source data and exact p -values are provided as a Source Data file.

metabolites responsible for this effect, we performed quadrupole time of flight liquid chromatography mass spectrometry (QTOF-LCMS) of control MRS media and *L. crispatus* CFS. LA was the top enriched metabolite in the *L. crispatus* CFS (Fig. 7c–e), with an orthogonal measurement of lactate levels confirming a 60-fold enrichment of lactate in *L. crispatus* CFS (Supplementary Fig. 9c). Strikingly, treatment with LA alone decreased both primary and secondary cervical organoid formation, suggesting that LA derived from local resident *Lactobacilli* is sufficient to suppress the self-renewal of human cervical stem cells (Fig. 7f–i).

To gain insight into the global transcriptomic effect induced by LA in cervical stem cells, we performed bulk RNA sequencing of human ectocervical organoids treated with either control media or LA-supplemented media, both at pH 6.5. Unsupervised hierarchical clustering revealed clear separation between control and LA-treated cervical organoids (Fig. 7j and Supplementary Fig. 9d, e). LA treatment differentially regulated the expression of 1153 out of 12,532 genes (Fig. 7k). In GSEA for GO:BP, gene sets of encapsulating structure organization and endodermal cell differentiation were enriched in LA-treated organoids, while gene sets of keratinization and neutrophil chemotaxis were depleted in LA-treated organoids (Fig. 7l, n and Supplementary Fig. 9f, g). Additional analysis showed that LA treatment induced the enrichment of genes involved in the gene sets of positive regulation of epidermal cell differentiation, epithelial cell morphogenesis, stem cell differentiation, and columnar cuboidal epithelial cell differentiation, further indicating that LA treatment promotes cervical stem cell differentiation (Fig. 7m, n and Supplementary Fig. 9g). Of note, LA treatment led to a significant enrichment of the YAP1 down gene set⁴⁹, consistent with the role of YAP1 in cervical stem cells in Supplementary Fig. 8 (Fig. 7o and Supplementary Fig. 9g). LA treatment also led to significant enrichment of PTEN up gene set, consistent with the role of the PI3K-AKT pathway in cervical stem cells in Fig. 6 and Supplementary Fig. 6 (Fig. 7p and Supplementary Fig. 9g). Indeed, LA treatment reduced the expression of YAP1 target genes and PI3K-AKT pathway genes (Fig. 7q, r), indicating that LA secreted by *Lactobacilli* species regulates human cervical stem cells through YAP1 and the PI3K-AKT pathway.

Aberrant expansion of tissue stem cells often precedes malignant transformation¹¹. Because LA decreases cervical organoid formation and promotes cervical stem cell differentiation, we asked whether LA can interfere with cervical carcinogenesis. A recent study reported that overexpression of HPV16 E6 and E7 proteins increased the number and size of cervical organoids⁵⁰. We therefore examined whether LA could counteract the effect of overexpression of HPV E6 and E7 in cervical organoid formation. Indeed, the increasing effect of cervical organoid number by HPV E6 and E7 overexpression was inhibited by LA treatment (Fig. 7s, t and Supplementary Fig. 9h).

Taken together, our data indicate that LA, a *Lactobacillus* metabolite, inhibits self-renewal and promotes differentiation of human cervical stem cells, suggesting that this is one of the ways that

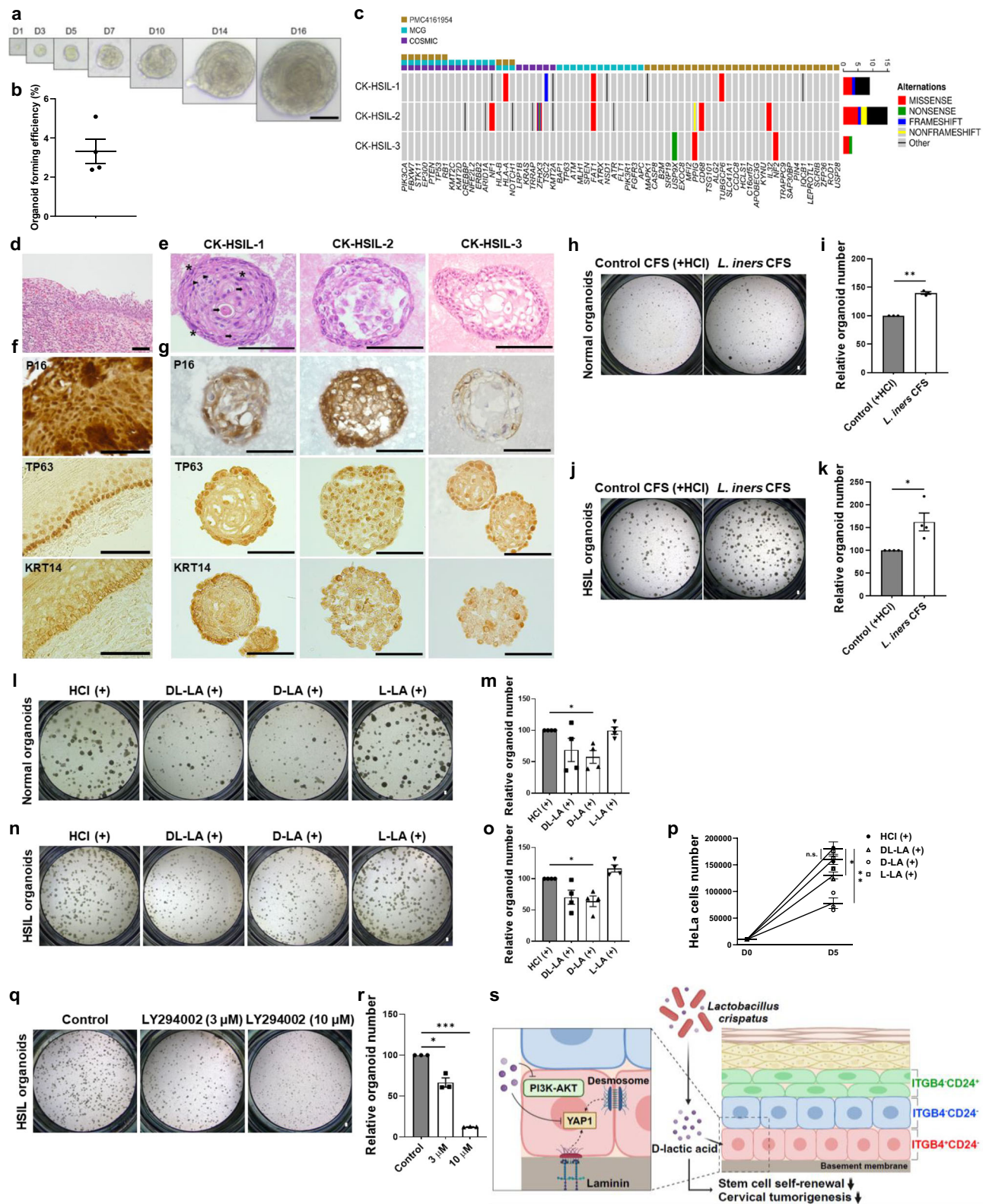
Lactobacillus colonization of the cervicovaginal tract can be protective against cervical cancer.

Establishment of human cervical precancerous organoids

Given our finding that LA treatment counteracts the effect of HPV E6 and E7 overexpression, we next asked whether LA has also an inhibitory effect in precancerous lesions and cervical cancers. To this end, we developed cervical precancerous (HSIL) organoid system. From punch biopsies of HSIL samples from nine patients infected by HPV16 or HPV33, we established HSIL organoids, which grew over two weeks, reaching up to 300 μ m in diameter (Fig. 8a). The efficiency of primary HSIL organoid formation was 3.3% (Fig. 8b). HSIL organoids were maintained and passaged over at least 6 passages and 3 months (Supplementary Fig. 10a). Since long-term culture and passaging can induce genetic and molecular changes in organoids compared to their original biopsy sample, we used organoids within the third passage for the subsequent experiments.

To analyze the mutational landscape of HSIL samples, we performed whole-exome sequencing (WES) for three precancerous cervical lesions. Since the amount of tissue is too limited to perform both WES and organoid culture, we cultured for organoids for the first two HSIL samples, then performed WES on two HSIL primary or the first passaged organoids (CK-HSIL-1 and CK-HSIL-2) and one HSIL tissue sample (CK-HSIL-3) (Fig. 8c). We identified 7156 somatic variants, including 3517 missense, 655 non-frameshift, 221 frameshift, 74 nonsense, and 2689 others (start loss, stop loss, stop gain, synonymous, and unknown). Since we could not obtain matched normal tissues or blood from the first two samples, we used a reference set that include the 65 most highly mutated SCC genes defined in COSMIC, MCG (My Cancer Genome), and a previous study based on 115 cervical carcinoma-normal paired samples⁵¹. 23 overlapping variants were identified and covered 20 genes out of 65 referred genes. Among these genes, we note *FAT1* and *KMT2A*, which are frequently mutated in many types of cancers, including cervical cancers^{52–55}.

We further investigated mutational signatures in terms of single base substitution (SBS) and doublet base substitution (DBS) patterns using SigProfilerExtractor, which provides ratios of mutation signatures defined in COSMIC. In total, 9 SBS and 2 DBS types were identified from 67 SBS and 19 DBS signature types, respectively. The SBS signature shared by all three samples was SBS5, which is related to aging (Supplementary Fig. 10b). CK-HSIL-1 and CK-HSIL-2 shared the SBS1 signature, which is strongly associated with spontaneous deamination of 5-methylcytosine and aging⁵⁶. Although these two signatures are detected in most types of cancers, SBS1 and SBS5 signatures show higher linkage with cervical cancers than other types of cancers, according to the indications on COSMIC. CK-HSIL-2 presents the SBS30 signature, which occurs due to deficiency in base excision repair (BER). BER deficiency is associated with high mutation rates and high risk of cervical cancers⁵⁷. In DBS signatures, all three samples displayed DBS17, which is closely connected with



Apolipoprotein B mRNA Editing Catalytic Polypeptide (APOBEC)-like mutagenesis (Supplementary Fig. 10c). Polymorphism in APOBEC genes frequently leads to oncogenic pathogenesis in cervical cancer⁵⁸. Taken together, these data indicate that HSILs display some of the mutational signatures frequently found in cervical cancers, which could serve as potential biomarkers for cervical carcinogenesis.

We further characterized our HSIL organoids by H&E staining and immunohistochemistry (IHC) staining. Unlike the normal ectocervical

organoids, HSIL organoids displayed the dysplastic architecture of human cervical HSIL and higher cellularity in the center than normal cervical organoids (Figs. 1g and 8d, e), although both normal and HSIL organoids are composed of squamous cells. HSIL organoids have several features of SCCs, including anisocytosis, nuclear irregularity, intercellular bridge, and keratin pearl. However, unlike SCCs, their stratified features are conserved. IHC staining revealed that HSIL organoids express P16, a surrogate marker of HPV infection. HSIL

Fig. 8 | The establishment of long-term, three-dimensional cultures of human cervical precancerous organoids and their utility. **a** Time course images of human cervical precancerous HSIL organoids over 16 days. **b** Organoid-forming efficiency of HSIL organoids in our cervical organoid culture medium ($n = 4$). **c** Somatic mutations called from the cervical HSIL lesions. Mutation profiles are derived from organoids from primary HSIL cells or organoids within passage number 2. **d** H&E staining of human HSIL lesion. **e** H&E staining of human HSIL organoids. Asterisk: anisocytosis. Arrowhead: intercellular bridge. Arrow: keratin pearl. **f** Immunohistochemistry staining of human HSIL lesion. **g** Immunohistochemistry staining of human HSIL organoids. **h, i** The effect of *L. iners* CFS on human normal ectocervical organoid formation. Representative bright-field images (**h**) and bar graph for relative numbers (**i**) of organoids ($n = 3$). **j, k**, The effect of *L. iners* CFS on human HSIL organoid formation. Representative bright-field images (**j**) and bar graph for relative numbers (**k**) ($n = 4$). The effect of lactate isomers on normal cervical

organoid formation. Representative bright-field images (**l**) and bar graph for relative numbers (**m**) ($n = 4$). The effect of lactate isomers on HSIL organoid formation. Representative bright-field images (**n**) and bar graph for relative numbers (**o**) ($n = 4$). **p** The effect of lactate isomers on cell proliferation of HeLa cervical cancer cell line ($n = 3$). **q, r**, The effect of LY294002 on HSIL organoid formation. Representative bright-field images (**q**) and bar graph for relative numbers (**r**) ($n = 3$). **s** Model of *Lactobacilli* metabolite in human cervical stem cells. *Lactobacilli*-derived lactate regulates normal and precancerous stem cells in the human cervix and interferes with HPV-induced tumorigenesis. Created in BioRender. Myeong, J. (2025) <https://BioRender.com/a50d432>. All data are collected from indicated biological replicates and presented as mean \pm SEM ($*p < 0.05$; $**p < 0.01$; $***p < 0.001$). *P*-values were calculated by two-tailed unpaired Student's *t*-test (**i, k**) or one-way ANOVA with Tukey's multiple comparison test (**m, o, p, r**). Source data and exact *p*-values are provided as a Source Data file. All scale bars: 100 μ m.

organoids also express the basal cell markers KRT14 and TP63, which were sporadic and not restricted to the outer margin as SCCs are (Fig. 8f, g and Supplementary Fig. 10d). These data indicate the successful establishment of a human cervical HSIL organoid system that recapitulates precancerous lesions of the human cervix.

Differential effect of lactate isomers in normal and precancerous cervical organoids

We showed above (Fig. 7a,b and Supplementary Fig. 9a,b) that CFS of *Lactobacilli* species interferes with cervical organoid formation, implicating a potential protective role against cervical tumorigenesis. However, a recent study reported that *L. iners* is associated with worse prognosis of cervical cancer patients and that lactate isomers display differential effects in cervical SCC chemoresistance⁵⁹. Another study reported that as cervical cancers develop from the precancerous stage, the ratio of L-LA in vaginal secretions is increased⁶⁰. Of note, LA exists in two isoforms, D(-)-lactic acid (D-LA) and L(+)-lactic acid (L-LA). *L. iners* produces L-LA, while *L. crispatus* and *L. gasseri* produce D-LA.

Therefore, we hypothesized that *L. iners* has a cancer promoting effect distinct from *L. crispatus* and *L. gasseri* and that LA isomers elicit differential effects on cervical integrity and tumorigenesis. To test this hypothesis, we examined the effect of *L. iners* CFS, employing both normal cervical and HSIL organoids. Indeed, *L. iners* CFS enhanced the formation of both normal cervical and HSIL organoids at the same pH (Fig. 8h–k). We further examined the effect of each of D-LA, L-LA, and DL-LA, the mixed form of D-LA and L-LA, in both normal cervical and HSIL organoids. Remarkably, D-LA interfered with the formation of normal cervical organoids, while DL-LA and L-LA did not (Fig. 8l, m). Similarly, D-LA interfered with the formation of HSIL organoids, while DL-LA and L-LA exerts no inhibitory effect (Fig. 8n, o). We further examined the differential effects of LA isomers in cervical cancers, employing HeLa and SiHa cervical cancer cell lines. Consistent with findings from normal and HSIL organoids, D-LA inhibited proliferation of both cancer cell lines, while L-LA did not (Fig. 8p and Supplementary Fig. 10e).

Taken together, these data demonstrate the utility of our HSIL organoid system in cervical stem cell and cancer research. Furthermore, we find that *Lactobacilli* species and LA isomers have differential roles in cervical stem cells and carcinogenesis, underscoring the need for caution when considering *Lactobacilli* species and LA isomers for regenerative medicine or cancer prevention.

Lastly, we explored the regulatory effect of the PI3K-AKT pathway and YAP1 in HSIL organoids. LY294002 treatment dose-dependently decreased HSIL organoid formation (Fig. 8q, r). VP treatment also reduced HSIL organoid formation (Supplementary Fig. 10f, g). These data indicate that the PI3K-AKT pathway and YAP1 regulate the stem cell self-renewal even in HSIL lesions.

Discussion

In this study, we report the identification and isolation of a population enriched for human cervical stem cells (Fig. 4o), by the establishment

of highly efficient systems for human ectocervical organoids and an intraligular transplantation. Through single-cell transcriptomic analysis, we define the cell types of the human cervix and mapped their differentiation trajectories and intercellular interactions. Moreover, we show that commensal *Lactobacillus*-derived LA regulates the self-renewal and differentiation of human cervical stem cells through the PI3K-AKT signaling pathway and YAP1 (Fig. 8s). Finally, we established cervical precancerous organoids and demonstrate differential effects of *Lactobacilli*-derived lactate isomers in normal and precancerous organoids. Our work provides foundational new systems and knowledge that will enable a better understanding of human cervical biology and pathophysiology.

Tissue stem cells maintain tissue integrity and often serve as the cell of origin for cancers, but it can be difficult to map the stem cell subpopulation in a given tissue. In the human cervix, the oncogenic molecular mechanisms by which HPV transforms cervical cells are relatively well understood, but the early-stage cellular and tissue-level consequences of premalignant development remain unclear—partly because of the lack of knowledge about cervical stem cells and their differentiation trajectory. Here, we identified stem cells and mapped their differentiation, gene expression, and regulation in the human ectocervix, where most cervical cancers occur. Although a very recent paper reported a single-cell transcriptomic analysis of the human ectocervix, it focused on comparing the cellular and molecular differences between normal and malignant tissues^{26,61}. In contrast, we asked a different set of fundamental questions: which cells are the tissue stem cells of the human cervix?, how do they differentiate?, how are they regulated?, and how does HPV infection affect their self-renewal and differentiation?

We characterized cluster-specific molecular programs and differential cell surface markers, functionally validating our ability to enrich the stem cell subpopulation using the ITGB4 and CD24 cell surface markers. The ITGB4⁺CD24⁺ subpopulation constitutes ~30% of the cervical epithelial cell population (Fig. 4c), and the organoid-forming efficiency of bulk cervical epithelial cells is up to ~40% (Fig. 1e). Considering that some of the transit amplifying cells are able to generate organoids (Fig. 4h,i)^{24,62,63}, our data suggest that ITGB4⁺CD24⁺ cells are human cervical stem cells, which have in vivo self-renewal capability. By identifying cell-surface markers that can be used to isolate a population enriched for human cervical stem cells, we substantially advance previous work on human and murine cervical organoids and stem cells. Also, our approach yields organoids at much higher efficiency than in previous work^{18,20}. Our scRNA-seq data delineate cell types and their developmental trajectory in detail, while previous studies elucidated the differences among normal, precancerous, and cancer stages^{18,19,26,61}. Moreover, by analyzing using cell surface markers for cervical stem cells, we show that HPV infection expands ITGB4⁺CD24⁺ transit amplifying cells, in contrast to previous reports that cervical stem cells are the major target of HPV infection¹⁷. We speculate that HPV exerts its effect most effectively at the stage of

transit amplifying cells by regulating cell cycle, likely through RB destabilization, although it infects human cervical stem cells.

To evaluate *in vivo* regeneration capability of cervical stem cells and organoids, we developed an intralingual transplantation method. Although orthotopic injection of stem cells into their originating organs is ideal for evaluating their capacity to regenerate tissue, it is not always feasible in organs such as the esophagus and intestine due to the high risk of puncture. Although subcutaneous transplantation is sometimes used, it is difficult to identify regenerated normal tissues afterward since loose subcutaneous space allows movement. Currently, subrenal capsule transplantation is widely used in these cases to examine tissue reconstitution by stem cells^{24,64–66}. The kidney's affluent blood flow supports the survival and growth of transplanted stem cells, and the renal capsule limits their movement²². However, the renal capsule is very fragile, and it is technically challenging to surgically implant stem cells or organoids underneath it. Our new method, intralingual transplantation, provides the advantages of the renal capsule system (blood flow and containment) with a much more accessible and less technically challenging procedure. Indeed, we demonstrate that human cervical stem cells and organoids transplanted into the tongues of immunocompromised mice regenerate stratified squamous epithelium-like structures. We believe that the intralingual transplantation technique will lower the technical hurdle of subrenal capsule transplantation and ease the demonstration of stem cells regenerative capacity in many contexts.

In addition, we established a precancerous cervical organoid system. In our characterization, HSIL organoids display dysplastic but still stratified histology and also express squamous genes, recapitulating the HSIL lesion *in vivo*. Importantly, exome sequencing of HSIL organoids detected somatic mutations of *FAT1*, *KMT2A*, and *NFI*, but not *TP53*; *TP53* mutations are rarely detected in cervical cancers since viral E6 protein mimics the effect of p53 mutations⁵². We further show experimental utility of HSIL organoids by testing the effects of LA isomers, LY294002, and VP in HSIL organoid formation. Our precancerous 3D organoid system provides new experimental avenues for cervical biology and can help to elucidate aspects of cervical carcinogenesis in precancerous stages.

Previous studies have implicated the local microbiome and pathogens in regulating cervical stem cells and cervical cancer development. *Lactobacilli* inversely correlate with HPV infection and malignant progression^{5,8}, while overexpression of HPV effector proteins E6 and E7 can increase cervical organoid formation and pathogenic *Chlamydia* infection enhances cell proliferation in cervical organoids⁵⁰. Importantly, our study extends these findings to show that commensal *Lactobacilli*, the most common taxon in the human cervicovaginal tract⁴⁸, can regulate human cervical stem cells through the production of lactic acid. Indeed, LA treatment counteracts the increase in cervical organoid formation induced by HPV E6 and E7 overexpression (Fig. 7s,t). Consistent with our data, previous studies reported that *L. crispatus* culture supernatant induces cell cycle arrest of diverse cancer cells^{9,10,67}. Considering that cervical stem cells are thought to be one of the major targets of HPV infection and subsequently transform to malignant cells, *Lactobacilli* could play an important preventive role against cervical carcinogenesis by interfering with aberrant self-renewal of HPV-infected cervical stem cells.

Importantly, our experiments with normal and precancerous organoids show a differential effect of LA isomers. A recent study reported that cancer-derived *L. iners* CFS and its major metabolite, L-LA, but not D-LA, promote therapeutic resistance of cervical cancer cells against irradiation or anticancer drugs and that cervical cancer patients with *L. iners* have worse overall survival than patients without it³⁹. We found that D-LA reduces normal cervical and HSIL organoids, while L-LA does not. Although our and previous

studies measured different outcomes of organoid formation or therapeutic resistance, they are in accord on some important conclusions: (1) lactate isomers exert different effects; (2) only D-LA is likely to protect the human cervix from cervical cancer; and (3) L-LA is likely to aggravate cervical cancer. Further studies are required to reveal the mechanisms by which lactate isomers exert these distinct influences.

We characterized the regulatory networks in the human ectocervix at multiple levels. By comparing the basal and suprabasal clusters, we show that the PI3K-AKT pathway regulates cervical stem cells' self-renewal and differentiation. We further delineated inter-cell type and intra-epithelial signaling networks in the human ectocervix and found that YAP1 regulates ectocervical stem cell self-renewal as a downstream mediator of desmosomes and laminins. It appears that YAP1 regulates stem cells of most surface epithelia, since YAP1 also regulates epidermal, intestinal, esophageal, and tracheal stem cells^{42–45,68}. Notably, LA induces the gene sets of YAP1 down and PTEN up, in addition to suppressing the expression of YAP target genes and PI3K-AKT pathway genes (Fig. 7o–r and Supplementary Fig. 9g). Therefore, we suggest a model in which commensal *Lactobacilli* regulate human cervical stem cells via the metabolite, LA, through YAP1 and the PI3K-AKT pathway (Fig. 8s). Furthermore, given that YAP1 and *PIK3CA* function as oncogenes while *PTEN* serves as a tumor suppressor gene in many cancers^{69–72}, we speculate that *Lactobacilli* can prevent cervical carcinogenesis by interfering with YAP1 and the PI3K-AKT signaling pathway, which could be explored in future work.

In summary, we report the development of highly efficient human ectocervical normal and precancerous organoid culture systems and an intralingual transplantation method. Employing these systems and characterizing cellular heterogeneity from single-cell transcriptomic analysis, we identified human cervical stem cells and explored their regulatory mechanisms via the PI3K pathway, YAP1, and *Lactobacillus*-produced LA isomers.

Methods

Human materials for organoid culture

All experiments using human materials were approved by the internal review boards of DGIST and Chilgok Kyungpook National University Hospital, Gyeongsangbuk-do, Korea, in accordance with all relevant ethical regulations (KNUCH-2020-12-020-001 and DGIST-20210401-BR-112-01). Human ecto- and endo-cervical tissues were provided by the Department of Gynecology at the Chilgok Kyungpook National University Hospital. Healthy cervical tissues were obtained from patients who underwent total laparoscopic hysterectomy (TLH) for their benign uterine diseases. Informed consent was obtained from all patients. Ectocervical and endocervical tissues from a total 14 donors were used in this study (Supplementary Table 1).

Materials availability

All organoids established in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Mice

All experimental work complied with the Republic of Korea Animal Scientific Procedures Act. The animal study was approved by the Ethical Committee and the Institutional Animal Care and Use Committee at DGIST. C57BL/6J mice and *NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ* (NSG) mice were obtained from Jackson laboratories, bred, and maintained in Daegu Gyeongbuk Institute of Science and Technology Animal Facility. All mice were accommodated under specific pathogen-free (SPF) conditions and maintained on a 12-h light/dark circadian cycle with free access to food and water. Ambient temperature was maintained at 20–24 °C. 1–6 months old female C57BL/6J mice and 1–12 months old male and female NSG mice were used.

Isolation of epithelial cells from human and mouse cervical tissues

Human and mouse ectocervical tissues were washed once in sterile PBS and incubated in dispase solution (2.4 U/ml of dispase, Gibco) at 37 °C for 1 h. Then, the epithelial layer was peeled off from connective tissue, incubated in TrypLE Express (Gibco) at 37 °C for 5 min and neutralized with blocking buffer (DPBS containing 2% bovine calf serum). After centrifugation, the pellets were incubated with 1X RBC lysis buffer and 1X DNase I (Enzymomics) at room-temperature (RT) for 5 min. After neutralization with the blocking buffer, the cell suspension was filtered through a 40 µm cell strainer. The dissociated cells were pelleted by centrifugation at $300 \times g$ for 5 min and resuspended in the optimized medium for further experiments.

Cervical organoid culture

Media containing cervical cells were mixed with growth factor-reduced Matrigel (Corning) at a 1:1 ratio, and a final volume of 100 µl was seeded into the inside of the cell culture insert. After solidification, 400 µl of the growth medium was added to the external side of the cell culture inserts. The optimized medium for human is composed of Advanced DMEM/F12 (Gibco) supplemented with 1× GlutaMAX supplement (Gibco), 1× penicillin–streptomycin (Biowest), 10 mM HEPES (Gibco), 1X B27 supplement (Gibco), 100 ng/ml Noggin (PEPROTECH), 1.25 mM *N*-acetyl-L-cysteine (Sigma), 10 µM Y-27632 (STEMCELL Technologies), 50 ng/ml FGF2 (PEPROTECH), 25 ng/ml FGF7 (PEPROTECH), 500 nM A83-01 (Selleckchem), 10 µM forskolin (Sigma), 100 ng/ml neuregulin 1 (PEPROTECH), 20 ng/ml NGF (PEPROTECH). For human endocervical organoid culture, 50 ng/ml EGF (PEPROTECH), 50 ng/ml Wnt-3a (R&D System) and 300 nM CHIR 99021 (Selleckchem) were additionally added to the optimized ectocervical organoid medium. For murine ectocervical organoid culture, medium was modified by adding 100 ng/ml R-spondin 1 and withdrawing FGF2, neuregulin 1, and NGF.

For the evaluation of effects of metabolites produced by *Lactobacillus crispatus* (*L. crispatus*), *L. gasseri*, and *L. iners* on human ectocervical stem cells, we plated cervical cells in either our optimized medium supplemented with *L. crispatus*, *L. gasseri*, *L. iners* cell free supernatant (CFS), or the control bacterial culture media at 5% v/v with hydrochloric acid solution added to adjust the pH to the same level as the other treatment (pH 7.07 for *L. crispatus* CFS, pH 6.65 for *L. gasseri* CFS, and pH 7.14 for *L. iners* CFS). For the evaluation of effects of lactic acid, human ectocervical cells were cultured for organoid formation in our optimized medium supplemented with 1/10 diluted lactic acid solution (Sigma). For the control group, hydrochloric acid solution was added to adjust pH to 6.5 (equivalent to the pH of LA-supplemented media). After 8–10 days of culture, the number of organoids of each condition was counted. Data were analyzed using ImageJ software and experiments were performed in at least three biological replicates.

Organoid passaging

Organoids were incubated with dispase solution at 37 °C for 40 min, followed by treatment with TrypLE Express at 37 °C for 15 min. Dissociated cells were then passed through a 40 µm cell strainer and seeded for secondary organoid culture.

In vivo transplantation of ectocervical organoids

When organoid size reached 50–70 µm, ectocervical organoids were resuspended in a mixture of Matrigel and medium, and then 50 µl of organoid-containing mixture of Matrigel and medium was injected underneath the tongue epithelium of NSG mice using a 30 G needle. Alternatively, 300–10,000 ectocervical cells were transplanted. After 3–4 weeks, the grafted tongues were surgically removed for a series of analyses.

Cell culture

HeLa cells, a human cervical adenocarcinoma epithelial cell line (KCLB 10002), and SiHa cells, a human squamous cell carcinoma cell line (KCLB 30035), were obtained from Korean cell line bank. For the cell proliferation assay, cells were seeded at 1×10^4 cells per well in 24-well plates (SPL) and cultured in 5% CO₂ at 37 °C in Dulbecco's modified Eagle medium (D-MEM) (Biowest), supplemented with 10% FBS (Gibco) and 1% Penicillin–Streptomycin solution (Biowest).

To evaluate the effect of lactic acid isoforms on cell proliferation, cells were cultured in optimized media supplemented with DL-lactic acid solution (Sigma), D(-)-lactic acid (TCI), and L(+)-lactic acid (Sigma). In the control group, hydrochloric acid solution was added to adjust the pH to 6.5 (same as the pH of the lactic acid-added medium). 5 days after seeding, the number of cells in each well was counted.

Untargeted metabolomics

Control media or *L. crispatus* cell-free supernatant (CFS) were freeze-dried overnight using PVTFD 20 R (Ilisin, Republic of Korea), and stored at –80 °C for later LC-QTOF/MS analysis. Lyophilized samples were dissolved in 1 ml of MeOH and then vortexed at 1600 rpm for 5 min. After centrifugation ($15,000 \times g$ for 30 min at 4 °C), the supernatants were filtered through a 0.2 µm nylon syringe filter into a glass LC vial. This yielded a total of 18 samples (2 treatment groups, each with 3 biological replicates and 3 extraction replicates). A quality control sample was prepared by pooling equal aliquots from each sample, bringing the total to 19.

Untargeted metabolomics analysis was performed using a 1290 Infinity II LC system (Agilent Technologies, USA) coupled with a 6545XT AdvanceBio Q-TOF (Agilent Technologies, USA) equipped with a Dual AJS ESI source under negative mode. The separation of compounds was achieved with an InfinityLab Poroshell 120 EC-C18 column (2.7 µm, 2.1×150 mm). The volume of sample injected into the column was 2 µl. In ESI- mode, binary eluent system of water (A) and acetonitrile (B), both containing 0.1% formic acid, was used at a flow rate of 0.15 ml/min using the following 36 min-cycle: increased to 30% B within 18 min, 30–70% B in 7 min, 70–95% B within 7 min, 95–100% B within 3 min, hold at 100% B for 1 min. The ion source conditions for the ESI- were as follows: gas temperature, 325 °C; gas flow, 7 L/min; nebulizer gas, 25 psi; sheath gas temperature, 275 °C; sheath gas flow, 12 L/min; capillary voltage, 4000 V; nozzle voltage, 2000 V; fragmentor, 180 V; skimmer, 45 V; octopole RF, 750 V; mass range, 30–1700 m/z; collision energy, 20 eV.

The LC-QTOF data was exported to MassHunter Qualitative Analysis version B.07.00 software (Agilent Technologies, USA) for feature extraction and peak picking. To remove information considered as background noise, peaks with ion abundances <50,000 were excluded from further analysis. The extracted feature files as compound exchange format (.cef files) were imported into Agilent Mass Profiler Professional version 15.1 software (Agilent Technologies, USA). The identification browser was used to identify the metabolites using MassHunter METLIN Metabolite Accurate Mass-Personal Compound Database and Library (AM-PCDL) version B.08.00 database. Normalized data were exported as a comma-separated value file (.CSV file) with annotated metabolites, and further analysis was performed using R 4.0.2 software.

Lactate concentration measurement

Lactate concentration was measured using an EZ-Lactate Assay Kit (DogenBio, Korea) according to the manufacturer's instructions. The fluorescent signals were detected at Ex/Em = 535/590 nm using a fluorescence microplate reader (Bio-Rad, USA).

Histology and immunostaining

Tissue paraffin embedding and sectioning. Cervical tissues and organoids were fixed in 4% paraformaldehyde (PFA) at RT for 2 hours and washed once with 100 mM glycine and then once with PBS. Fixed cervical tissues and organoids were dehydrated sequentially in 50%, 70%, 90%, 95%, and 100% ethanol, and then with xylene. Next, they were embedded in paraffin, sliced with a microtome at a thickness of 10 μ m, and directly mounted onto a slide. The sliced cervical tissues and organoids were deparaffinized and rehydrated sequentially in xylene, 100%, 95%, 90%, 70%, and 50% ethanol, and water.

H&E staining. H&E staining was performed using a hematoxylin and eosin stain kit (Vector laboratories, Cat# H-3502) according to the manufacturer's instructions. In brief, slides were stained with hematoxylin at RT for 5 min and washed twice with distilled water. After incubation with a bluing reagent at RT for 10–15 s, slides were washed twice with distilled water and then, immersed in 100% ethanol for 10 s. Slides were stained with eosin at RT for 2–3 min, washed with 100% ethanol for 10 s and dehydrated three times with 100% ethanol for 1–2 min. Thereafter, stained slides were mounted and imaged with a microscope with a digital camera.

Immunostaining. Cervical tissue or organoid slides were subjected to a heat-induced antigen-retrieval procedure using IHC Antigen Retrieval Solution (Invitrogen) at 90 °C for 20 min. After cooling and washing the slides with 1× PBS, blocking buffer containing 0.1–0.3% Triton X-100, 3% BSA (MP Biomedicals), 3% normal donkey serum (Abcam) and 3% normal goat serum (Cell Signaling Technology) was applied at RT for 1 h. Then, sample slides were incubated with primary antibodies in Dako REAL Antibody Diluent (Dako) overnight at 4 °C. Primary antibodies used in this study are as follows: rabbit anti-keratin 14 (1:1000, Abcam), mouse anti-keratin 14 (1:1000, Invitrogen), rabbit anti-keratin 1 (1:1000, Biolegend), rabbit anti-loricrin (1:100, Biolegend), mouse anti-keratin 7 (1:100, Invitrogen), rabbit anti-keratin 5 (1:200, Biolegend), rabbit anti-YAP (1:100, Cell Signaling Technology), mouse anti-ITGB4 (1:20, Santacruz), rabbit anti-Ki67 (1:100, Abcam), mouse APC anti-human CD24 (1:50, Biolegend), mouse FITC anti-human CD271 (NGFR) (1:50, Biolegend), mouse FITC anti-human HLA-A,B,C (1:25, Biolegend), and rabbit HRP-conjugated anti-p63 (1:100, Abcam), rat FITC anti-mouse ITGB4 (1:50, Biolegend), and rat APC anti-mouse CD24 (1:50, Biolegend). After washing the slides three times with 0.1% Triton X-100 solution in 1X PBS, sample slides were incubated with secondary antibodies at RT for 1 hour. After washing again three times with Triton X-100 solution, slides were incubated with DAPI solution (1:1000, Roche) at RT for 3 min and mounted using an anti-fade fluorescence mounting medium (Abcam), and imaged using an epifluorescence microscope (ZEISS, Axio Vert.A1) or a confocal microscope (ZEISS, LSM800).

Image acquisition. The images were taken in a confocal microscope (ZEISS, LSM800).

Flow cytometry analysis

Cells were suspended in PBS containing 2% FBS. Antibodies were added at the following ratios; FITC anti-human ITGB4 (1:50, Biolegend), PerCP-Cy5.5 anti-human CD24 (1:50, Biolegend), PE anti-human EpCAM (1:50, Biolegend), FITC anti-mouse ITGB4 (1:50, Biolegend), PE-Cy7 anti-mouse CD24 (1:50, Biolegend), and APC anti-mouse EpCAM (1:50, Biolegend). Dead cells were discriminated using Zombie dye (1:100, Biolegend) or propidium iodide (20 μ g/ml, Sigma).

Cell cycle analysis

Single cells were suspended in 150 μ l of PBS, and 350 μ l of absolute ethanol was added in a dropwise manner while vortexing. After incubating cells at 4 °C for at least 2 h, fixed cells were centrifuged at

1500 \times g for 3 min, and supernatant was removed. After washing with PBS, pellets were resuspended in staining solution containing 100 μ g/ml RNase A and 50 μ g/ml propidium iodide and incubated at RT for 30 min. After filtration through a cell strainer, cells were analyzed by the flow cytometer.

RNA isolation and RT-qPCR

Total RNA was isolated using the Monarch Total RNA Miniprep Kit (New England Biolabs) according to the manufacturer's instructions. Reverse transcription was performed using High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. cDNA was pre-amplified by 16–18 cycles depending on the initial amount of RNA quantities using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific). Quantitative PCR was performed using StepOnePlus Real-Time PCR System (Applied Biosystems). Expression of genes was analyzed by delta-delta-Ct method with β -Actin for normalization. PCR primers used in this paper are listed in Supplementary Table 2.

Nucleofection

Nucleofection of cervical epithelial cells was performed using Amaxa 4D-Nucleofector (EA-125 pulse code) and P4 Primary Cell 4D-Nucleofector X kit (Lonza). After nucleofection, cells were recovered at RT for 10 min and resuspended in growth media. Cells were plated into cell culture inserts with Matrigel for the organoid formation and seeded on ultralow attachment plates for 48 hours for real time PCR to confirm the silencing of target genes. siRNA oligo sequences used in this paper are listed in Supplementary Table 3.

Single-cell RNA sequencing analysis

scRNA sequencing reads were processed by Cell Ranger version 6.1.1 with hg38 reference transcriptome. Output files were imported and analyzed by Seurat v3.0 in R. Datasets were integrated, and cells with low gene numbers (<200) or with high mitochondrial gene counts (>25) were filtered out. After the filtering process, doublets were eliminated by Scrublet. 19,172 cells with 21,925 genes were included in single cell objects. Whole UMI counts were normalized and then scaled by regressing out cell-cycle score and mitochondrial gene counts. We performed PCA and selected 15 principal components that were based on the top 2000 variable features. The overall UMAP projection was organized by graph-based clustering with resolution 0.3, and 8 clusters were generated. Performing “FindAllMarkers” function in Seurat, DEGs were generated for each cluster, and significance was determined with adjusted p -value < 0.05, Log₂FC (Fold Change) > 0.25. With reference to the representative cell type marker genes, we annotated 3 cell types in each cluster. In order to examine epithelial cell conditions in detail, we re-clustered epithelial cells alone, with an overall processing procedure identical to above. We constructed a UMAP projection with 18,083 epithelial cells and annotated 5 cell type clusters by referring to epithelial cell type marker genes of human esophagus⁷³ lined by the stratified squamous epithelium like human ectocervix.

To identify DEGs between basal cells and suprabasal cells, “FindMarker” function in Seurat with the parameter of Wilcoxon test was used. DEGs for each cell type with p -value < 0.05 were depicted in volcano plots and dot plots using ggplot2 and ggdotchart R packages, respectively. Then, pre-ranked GSEA analyses were processed by GSEA version 4.1.0. GSEA pathways used in overall plots and graphs were selected under the condition p -value < 0.05. Whole represented heatmaps of DEGs were illustrated by the “DoHeatmap” function in Seurat. Pie charts were created using the ggplot2 package in R.

For the epithelial cell trajectory inference analysis, Slingshot package version 2.4.0 and Monocle package version 2.24.1 in R was used to predict the transition lineage of cells according to pseudo-time.

For the identification of murine ectocervical stem cell surface markers, we downloaded and analyzed the mouse ectocervix epithelium scRNA-seq dataset (NCBI GEO GSE128987). Raw data was pre-processed in the same way as above and 1856 cells with 17,240 genes were in the resultant object. We selected 15 principal components to construct UMAP of mouse data. We checked 3 representative basal cell type markers to specify cell types of the clusters.

For cell-cell interaction analysis, CellChat package version 1.6.1 in R was used. To identify DEGs specifically between basal cells and suprabasal cells, we used the “FindMarker” function in Seurat with the parameters of the Wilcoxon test. We identified DEGs for each cell type at p -value < 0.05 . These were depicted in volcano plots and dot plots by using ggplot2, ggdotchart R package, respectively. Whole represented heatmaps of DEGs were illustrated by “DoHeatmap” function in Seurat. A pie chart was generated by ggplot2 package in R. Enrichment scores for 2 gene sets (‘HIPPO pathway’ and ‘YAP downstream’) were assigned to basal or suprabasal cells by using UCell package version 2.0.1 in R.

For the gene set enrichment analysis, pre-ranked GSEA analyses were processed by GSEA version 4.1.0. GSEA pathways used in overall plots and graphs were selected under the condition $P < 0.05$.

Bulk RNA sequencing analysis

Raw files were processed by the STAR aligner package version 2.7.10 using the GRCh38 reference genome. Whole count matrices from each dataset were merged into one. Batch effect was corrected by “ComBat-seq()” function in sva R package version 3.44.0. DEG analyses were processed using edgeR R package version 3.38.4. GSEA analyses were processed by GSEA version 4.1.0. GSEA pathways used in overall plots and graphs were selected using the condition FDR q -value < 0.05 .

Whole exome sequencing

Overall fastq files from 3 HSIL samples and one blood sample were pre-processed with Trimmomatic version 0.39⁷⁴ by trimming adapter sequences and discarding sequences having low base quality. Using bwa2 (Burrow-Wheeler Aligner) version 0.7.17⁷⁵, whole sequence reads were mapped on GRCh38. Using GATK version 4.5.0⁷⁶, generated bam files were sorted by several properties, assigned to new read groups, and purged of duplicated reads caused by PCR. Based on various covariates, base quality in bam files could be specified and recalibrated. Somatic SNVs were called based on assembly of haplotypes using Mutect2 version 4.1.0. Tumor-only mode was applied to CK-HSIL-1 and CK-HSIL-2 samples, while Tumor-normal mode was applied to CK-HSIL-3 and matched WBC. Throughout the somatic variants calling procedures, germline / normal artifacts were excluded employing gnomAD, 1000 Genomes filter resource. Functional annotations on variants were organized by computing ANNOVAR version 2020-06-08⁷⁷. Only variants with depth > 10 and allele frequency > 0.2 were used in subsequent analyses. Exonic / intronic alternations profiling on the most significant 65 cervical SCC mutations was computed by oncoPrint function in ComplexHeatmap version 2.18.0⁷⁸. Single-base substitution (SBS) and doublet base substitution (DBS) pattern compositions on overall sample variants were calculated by SigProfilerExtractor⁷⁹ for pre-defined mutation signatures of COSMIC mutational signatures version 3.4.

Statistics and reproducibility

Statistical analyses were performed using two-sided Student t -test between two groups, and statistical significance was defined based on $P < 0.05$. Data are presented as mean \pm SEM of at least 3 or more independent biological replicates. All immunofluorescence and immunohistochemistry staining were performed at least twice, and the representative images were included.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Bulk and single cell RNA sequencing data generated in this study have been deposited at NCBI and are publicly available under the following accession codes: PRJNA963403 (hCervix scRNA-seq data), PRJNA963409 (Bulk RNA-seq data with or without VP treatment) and PRJNA964006 (Bulk RNA-seq data with HCl or LA treatment). Whole exome sequencing data generated in this study have been deposited at NCBI and are publicly available under the following accession codes: PRJNA1213881. Previously published murine cervical scRNA-seq data that are re-analyzed here are available at (NCBI GEO GSE128987). All other data, including organoid, cell cycle, and gene expression data, are provided in the Supplementary Information/Source Data file. Source data are provided with this paper.

Code availability

All codes used for analysis are available at <https://github.com/MinhoLee-DGU/2025.Myeong.et.al.NatureComms>.

References

- Singh, D. et al. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. *Lancet Glob. Health* **11**, e197–e206 (2023).
- International, World Cancer Research Fund. <https://www.wcrf.org/preventing-cancer/cancer-statistics/global-cancer-data-by-country/> (2022)
- Small, W. Jr. et al. Cervical cancer: a global health crisis. *Cancer* **123**, 2404–2412 (2017).
- Cohen, P. A., Jhingran, A., Oaknin, A. & Denny, L. Cervical cancer. *Lancet* **393**, 169–182 (2019).
- Lee, J. E. et al. Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. *PLoS ONE* **8**, e63514 (2013).
- Gao, W., Weng, J., Gao, Y. & Chen, X. Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. *BMC Infect. Dis.* **13**, 271 (2013).
- Brusselaers, N., Shrestha, S., van de Wijert, J. & Verstraelen, H. Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: systematic review and meta-analysis. *Am. J. Obstet. Gynecol.* **221**, 9–18.e18 (2019).
- Kang, G. U. et al. Potential Association between vaginal microbiota and cervical carcinogenesis in Korean women: a cohort study. *Microorganisms* **9**, 294 (2021).
- Sungur, T., Aslim, B., Karaaslan, C. & Aktas, B. Impact of exopolysaccharides (EPSs) of *Lactobacillus gasseri* strains isolated from human vagina on cervical tumor cells (HeLa). *Anaerobe* **47**, 137–144 (2017).
- Sharma, P., Kaur, S., Kaur, R., Kaur, M. & Kaur, S. Proteinaceous secretory metabolites of probiotic human commensal *Enterococcus hirae* 20c, *E. faecium* 12a and L12b as antiproliferative agents against cancer cell lines. *Front. Microbiol.* **9**, 948 (2018).
- Jeong, Y. et al. Role of KEAP1/NRF2 and TP53 mutations in lung squamous cell carcinoma development and radiation resistance. *Cancer Discov.* **7**, 86–101 (2017).
- Hayakawa, Y., Nakagawa, H., Rustgi, A. K., Que, J. & Wang, T. C. Stem cells and origins of cancer in the upper gastrointestinal tract. *Cell Stem Cell* **28**, 1343–1361 (2021).
- Bamodu, O. A., Chung, C. C., Pisanic, T. R. 2nd & Wu, A. T. H. The intricate interplay between cancer stem cells and cell-of-origin of

- cancer: implications for therapeutic strategies. *Front. Oncol.* **14**, 1404628 (2024).
14. Hou, Y. et al. A diet-microbial metabolism feedforward loop modulates intestinal stem cell renewal in the stressed gut. *Nat. Commun.* **12**, 271 (2021).
 15. Liu, X. et al. The impact of gut microbial signals on hematopoietic stem cells and the bone marrow microenvironment. *Front. Immunol.* **15**, 1338178 (2024).
 16. Regauer, S. & Reich, O. The origin of Human Papillomavirus (HPV) - induced cervical squamous cancer. *Curr. Opin. Virol.* **51**, 111–118 (2021).
 17. Frazer, I. H. Prevention of cervical cancer through papillomavirus vaccination. *Nat. Rev. Immunol.* **4**, 46–54 (2004).
 18. Chumduri, C. et al. Opposing Wnt signals regulate cervical squamocolumnar homeostasis and emergence of metaplasia. *Nat. Cell Biol.* **23**, 184–197 (2021).
 19. Zhao, Z. et al. Single-cell analysis defines the lineage plasticity of stem cells in cervix epithelium. *Cell Regen.* **10**, 36 (2021).
 20. Lohmussaar, K. et al. Patient-derived organoids model cervical tissue dynamics and viral oncogenesis in cervical cancer. *Cell Stem Cell* **28**, 1380–1396.e1386 (2021).
 21. Doorbar, J. Latent papillomavirus infections and their regulation. *Curr. Opin. Virol.* **3**, 416–421 (2013).
 22. Jeong, Y. Methods to identify epithelial stem cells. *Organoid* **2**, e24 (2022).
 23. Dutta, D., Heo, I. & Clevers, H. Disease modeling in stem cell-derived 3D organoid systems. *Trends Mol. Med.* **23**, 393–410 (2017).
 24. Jeong, Y. et al. Identification and genetic manipulation of human and mouse oesophageal stem cells. *Gut* **65**, 1077–1086 (2016).
 25. Rock, J. R. et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc. Natl. Acad. Sci. USA* **106**, 12771–12775 (2009).
 26. Liu, C. et al. Single-cell dissection of cellular and molecular features underlying human cervical squamous cell carcinoma initiation and progression. *Sci. Adv.* **9**, eadd8977 (2023).
 27. Wang, S. et al. Single cell transcriptomics of human epidermis identifies basal stem cell transition states. *Nat. Commun.* **11**, 4239 (2020).
 28. Zou, Z. et al. A single-cell transcriptomic atlas of human skin aging. *Dev. Cell* **56**, 383–397.e388 (2021).
 29. Polito, M. P., Marini, G., Palamenghi, M. & Enzo, E. Decoding the human epidermal complexity at single-cell resolution. *Int. J. Mol. Sci.* **24**, 8544 (2023).
 30. Piedrafito, G. et al. A single-progenitor model as the unifying paradigm of epidermal and esophageal epithelial maintenance in mice. *Nat. Commun.* **11**, 1429 (2020).
 31. Tanaka, T. et al. Identification of stem cells that maintain and regenerate lingual keratinized epithelial cells. *Nat. Cell Biol.* **15**, 511–518 (2013).
 32. Yin, H., Hu, M. & Li, D. Regulation of epidermal stratification and development by basal keratinocytes. *J. Cell Physiol.* **238**, 742–748 (2023).
 33. Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
 34. He, X. C. et al. PTEN-deficient intestinal stem cells initiate intestinal polyposis. *Nat. Genet.* **39**, 189–198 (2007).
 35. Wang, S. et al. Pten deletion leads to the expansion of a prostatic stem/progenitor cell subpopulation and tumor initiation. *Proc. Natl. Acad. Sci. USA* **103**, 1480–1485 (2006).
 36. Yanagi, S. et al. Pten controls lung morphogenesis, bronchioalveolar stem cells, and onset of lung adenocarcinomas in mice. *J. Clin. Investig.* **117**, 2929–2940 (2007).
 37. Muller, L., Hatzfeld, M. & Keil, R. Desmosomes as signaling hubs in the regulation of cell behavior. *Front. Cell Dev. Biol.* **9**, 745670 (2021).
 38. Garrod, D. & Chidgey, M. Desmosome structure, composition and function. *Biochim. Biophys. Acta* **1778**, 572–587 (2008).
 39. Matlin, K. S., Myllymaki, S. M. & Manninen, A. Laminins in epithelial cell polarization: old questions in search of new answers. *Cold Spring Harb. Perspect. Biol.* **9**, a027920 (2017).
 40. Aumailley, M. The laminin family. *Cell Adh. Migr.* **7**, 48–55 (2013).
 41. Kechagia, J. Z., Ivaska, J. & Roca-Cusachs, P. Integrins as biomechanical sensors of the microenvironment. *Nat. Rev. Mol. Cell Biol.* **20**, 457–473 (2019).
 42. Barry, E. R. et al. Restriction of intestinal stem cell expansion and the regenerative response by YAP. *Nature* **493**, 106–110 (2013).
 43. Zhao, R. et al. Yap tunes airway epithelial size and architecture by regulating the identity, maintenance, and self-renewal of stem cells. *Dev. Cell* **30**, 151–165 (2014).
 44. Zhang, H., Pasolli, H. A. & Fuchs, E. Yes-associated protein (YAP) transcriptional coactivator functions in balancing growth and differentiation in skin. *Proc. Natl. Acad. Sci. USA* **108**, 2270–2275 (2011).
 45. Choi, Y. J. et al. YAP1 regulates esophageal stem cells' self-renewal and differentiation. *Biochem. Biophys. Res. Commun.* **726**, 150280 (2024).
 46. Wang, Y. et al. Comprehensive molecular characterization of the hippo signaling pathway in cancer. *Cell Rep.* **25**, 1304–1317.e1305 (2018).
 47. Ruff, W. E., Greiling, T. M. & Kriegel, M. A. Host-microbiota interactions in immune-mediated diseases. *Nat. Rev. Microbiol.* **18**, 521–538 (2020).
 48. Ravel, J. et al. Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. USA* **108**, 4680–4687 (2011).
 49. MSigDB. <https://www.gsea-msigdb.org/gsea/msigdb> (2022).
 50. Koster, S. et al. Modelling Chlamydia and HPV co-infection in patient-derived ectocervix organoids reveals distinct cellular reprogramming. *Nat. Commun.* **13**, 1030 (2022).
 51. Ojesina, A. I. et al. Landscape of genomic alterations in cervical carcinomas. *Nature* **506**, 371–375 (2014).
 52. Cancer Genome Atlas Research, N. et al. Integrated genomic and molecular characterization of cervical cancer. *Nature* **543**, 378–384 (2017).
 53. Sanchez-Vega, F. et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell* **173**, 321–337.e310 (2018).
 54. Bolouri, H. et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat. Med.* **24**, 103–112 (2018).
 55. Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
 56. Guo, Q. et al. The mutational signatures of formalin fixation on the human genome. *Nat. Commun.* **13**, 4487 (2022).
 57. Bajpai, D., Banerjee, A., Pathak, S., Jain, S. K. & Singh, N. Decreased expression of DNA repair genes (XRCC1, ERCC1, ERCC2, and ERCC4) in squamous intraepithelial lesion and invasive squamous cell carcinoma of the cervix. *Mol. Cell Biochem.* **377**, 45–53 (2013).
 58. Revathidevi, S., Murugan, A. K., Nakaoka, H., Inoue, I. & Munirajan, A. K. APOBEC: a molecular driver in cervical cancer pathogenesis. *Cancer Lett.* **496**, 104–116 (2021).
 59. Colbert, L. E. et al. Tumor-resident *Lactobacillus iners* confer chemoradiation resistance through lactate-induced metabolic rewiring. *Cancer Cell* **41**, 1945–1962.e1911 (2023).
 60. de Magalhaes, C. C. B. et al. Comparative measurement of D- and L-lactic acid isomers in vaginal secretions: association with high-grade cervical squamous intraepithelial lesions. *Arch. Gynecol. Obstet.* **305**, 373–377 (2022).
 61. Zhang, T. et al. Identification of cervical cancer stem cells using single-cell transcriptomes of normal cervix, cervical premalignant lesions, and cervical cancer. *EBioMedicine* **92**, 104612 (2023).

62. McQualter, J. L., Yuen, K., Williams, B. & Bertoncello, I. Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. *Proc. Natl. Acad. Sci. USA* **107**, 1414–1419 (2010).
63. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
64. Chan, C. K. F. et al. Identification of the human skeletal stem cell. *Cell* **175**, 43–56.e21 (2018).
65. Maxwell, K. G., Kim, M. H., Gale, S. E. & Millman, J. R. Differential function and maturation of human stem cell-derived islets after transplantation. *Stem Cells Transl. Med.* **11**, 322–331 (2022).
66. Gurung, S., Deane, J. A., Darzi, S., Werkmeister, J. A. & Gargett, C. E. In vivo survival of human endometrial mesenchymal stem cells transplanted under the kidney capsule of immunocompromised mice. *Stem Cells Dev.* **27**, 35–43 (2018).
67. Wang, K. D. et al. Inhibitory effect of vaginal *Lactobacillus* supernatants on cervical cancer cells. *Probiotics Antimicrob. Proteins* **10**, 236–242 (2018).
68. Mahoney, J. E., Mori, M., Szymaniak, A. D., Varelas, X. & Cardoso, W. V. The hippo pathway effector Yap controls patterning and differentiation of airway epithelial progenitors. *Dev. Cell* **30**, 137–150 (2014).
69. Moroishi, T., Hansen, C. G. & Guan, K. L. The emerging roles of YAP and TAZ in cancer. *Nat. Rev. Cancer* **15**, 73–79 (2015).
70. Zancanato, F., Cordenonsi, M. & Piccolo, S. YAP/TAZ at the roots of cancer. *Cancer Cell* **29**, 783–803 (2016).
71. Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
72. Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582 (2015).
73. Rochman, M. et al. Single-cell RNA-Seq of human esophageal epithelium in homeostasis and allergic inflammation. *JCI Insight* **7**, e159093 (2022).
74. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
75. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
76. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
77. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
78. Gu, Z. Complex heatmap visualization. *Imeta* **1**, e43 (2022).
79. Islam, S. M. A. et al. Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. *Cell Genom.* **2**, None (2022).

Acknowledgements

This work was supported by grants from the National Research Foundation (NRF) of the Ministry of Science and ICT (2020R1A2C2009359 and RS-2023-00221112 for Y.J. and 2022R1A2C2093050 for M.L.), the DGIST R&D Program (24-CoE-BT-01 for Y. Jeong), a grant from the Ministry of Education (2022R111A1A01071095 for B.L.), a grant from the Institute for Basic Science (IBS) (IBS-R801-D1-2024-a03 for Y.J.), a grant from the Ministry of Health and Welfare (RS-2024-00439434 for Y.J.),

and a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (HI23C0515 for G.C.) of Korea. We additionally thank Drs. June Myoung Kwak and Younghoon Kee for their invaluable comments and Messrs. Young Jun Lee and Daeyoung Lee for their technical assistance of tissue preparation and H&E staining. We also appreciate the KNU NGS core facility for analyzing bacterial culture.

Author contributions

Conception and design: G. O. Chong, Y. Jeong, Performing experiments: J. Myeong, B. Lee, Y. Choi, J. Kim, S. Y. Jeon, D. Jung, M. Jeong, B. Oh, Y. Nam, C. S. Kim, N. J. Park. Analysis and interpretation of data: J. Myeong, B. Lee, J. H. Kim, Y. Shin, D. Jung, J. Jung, H. Shim, H. S. Han, J. Shin, N. J. Park, M. Lee, Y. Jeong. Clinical sampling: Y. H. Lee, G. O. Chong. Writing, review, and/or revision of the manuscript: B. Lee, M. Lee, Y. Jeong.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-57323-6>.

Correspondence and requests for materials should be addressed to Gun Oh Chong or Youngtae Jeong.

Peer review information *Nature Communications* thanks Nardhy Gomez-Lopez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025