

CANCER

Single-cell dissection of cellular and molecular features underlying human cervical squamous cell carcinoma initiation and progression

Chao Liu^{1,2†}, Min Zhang^{3†}, Xinlong Yan^{4†}, Yanli Ni^{5†}, Yandong Gong⁵, Cong Wang⁶, Xiaoling Zhang⁶, Lingfei Wan⁴, Hui Yang⁴, Chen Ge⁴, Yunqiao Li⁵, Wenxue Zou¹, Rui Huang¹, Xiaohui Li¹, Bing Sun⁷, Bing Liu^{5*}, Jinbo Yue^{1*}, Jinming Yu^{1,2*}

Cervical squamous cell carcinoma (CESC) is a prototypical human cancer with well-characterized pathological stages of initiation and progression. However, high-resolution knowledge of the transcriptional programs underlying each stage of CESC is lacking, and important questions remain. We performed single-cell RNA sequencing of 76,911 individual cells from 13 samples of human cervical tissues at various stages of malignancy, illuminating the transcriptional tumorigenic trajectory of cervical epithelial cells and revealing key factors involved in CESC initiation and progression. In addition, we found significant correlations between the abundance of specific myeloid, lymphoid, and endothelial cell populations and the progression of CESC, which were also associated with patients' prognosis. Last, we demonstrated the tumor-promoting function of matrix cancer-associated fibroblasts via the NRG1-ERBB3 pathway in CESC. This study provides a valuable resource and deeper insights into CESC initiation and progression, which is helpful in refining CESC diagnosis and for the design of optimal treatment strategies.

INTRODUCTION

Cervical cancer is the fourth most common malignancy in females worldwide (1), responsible for more than 500,000 new cases and more than 300,000 deaths each year (2). Cervical squamous cell carcinoma (CESC) accounts for approximately 75% of cases (3), with patient survival inversely correlated with International Federation of Gynecology and Obstetrics (FIGO) stage at diagnosis (4, 5). CESC initiation and progression from cervical intraepithelial neoplasia (CIN) to early and then advanced CESC are closely—but not exclusively—linked with chronic human papilloma virus (HPV) infection (6, 7). Studies using bulk RNA and DNA sequencing of tumor tissues have identified several molecular characteristics and gene mutations associated with progression in HPV-positive CESC (8–10); however, these approaches generate data that reflect the average characteristics of millions of tumor tissue cells of diverse types and, thus, are limited in their capacity to link transcriptional or genomic data with specific cellular identities and biologically relevant mechanisms (11, 12).

Single-cell RNA sequencing (scRNA-seq) is a powerful technology capable of revealing the heterogeneous cellular and molecular

characteristics of various cells within tumors (13). Recent advances in data analysis mean that scRNA-seq can now also be used to predict likely cell-to-cell interactions, which can be applied to the identification of key ligand-receptor pairs with the potential to promote tumor progression (14). In CESC, scRNA-seq has so far been used to investigate the heterogeneity of endothelial cells (ECs) within a single human CESC (15) and to describe the epithelial characteristics of cervical metaplasia in a mouse model (16). What remains unknown is how the molecular profile of epithelial cells evolves during CESC initiation and progression, how the transcriptomic profile of cells within CESC differs between stages, and how this might relate to the pro- or anti-tumorigenic functional features of specific cell types.

Here, we report findings from scRNA-seq analysis of 76,911 cells from 13 clinical samples ranging from normal cervix (NC) to advanced CESC. We uncovered evidence of a transcriptomic evolution occurring within the epithelial cell population during CESC initiation and progression, accompanied by marked changes in the frequency of specific subclusters of epithelial cells, T cells, natural killer (NK) cells, macrophages, and ECs. These changes correlated significantly with CESC prognosis and highlighted an important role of matrix cancer-associated fibroblasts (mCAFs) in promoting CESC progression. Together, our study comprehensively dissects tumorigenesis and progression of CESC at single-cell resolution.

RESULTS

Single-cell transcriptome landscape in CESC

To better understand CESC initiation and progression, we performed scRNA-seq on samples of NC from ectocervix ($n = 3$ samples), CIN ($n = 2$ samples), early CESC (FIGO stage IB2, $n = 3$ samples), and advanced CESC (FIGO stage IIB to IIIC, $n = 5$ samples) (Fig. 1A and table S1). After generating single-cell

¹Department of Radiation Oncology and Shandong Provincial Key Laboratory of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, China. ²Research Unit of Radiation Oncology, Chinese Academy of Medical Sciences, Jinan 250117, China. ³Department of Nephrology, First Medical Center of Chinese PLA General Hospital, Beijing 100853, China. ⁴Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China. ⁵State Key Laboratory of Experimental Hematology, Senior Department of Hematology, Fifth Medical Center of Chinese PLA General Hospital, Beijing 100071, China. ⁶Department of Gynecologic Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan 250117, China. ⁷Department of Radiation Oncology, Fifth Medical Center of Chinese PLA General Hospital, Beijing 100071, China.

[†]These authors contributed equally to this work.

*Corresponding author. Email: bingliu17@yahoo.com (B.L.); jbyue@sdfmu.edu.cn (J.B.Y.); sdyujinming@163.com (J.M.Y.)

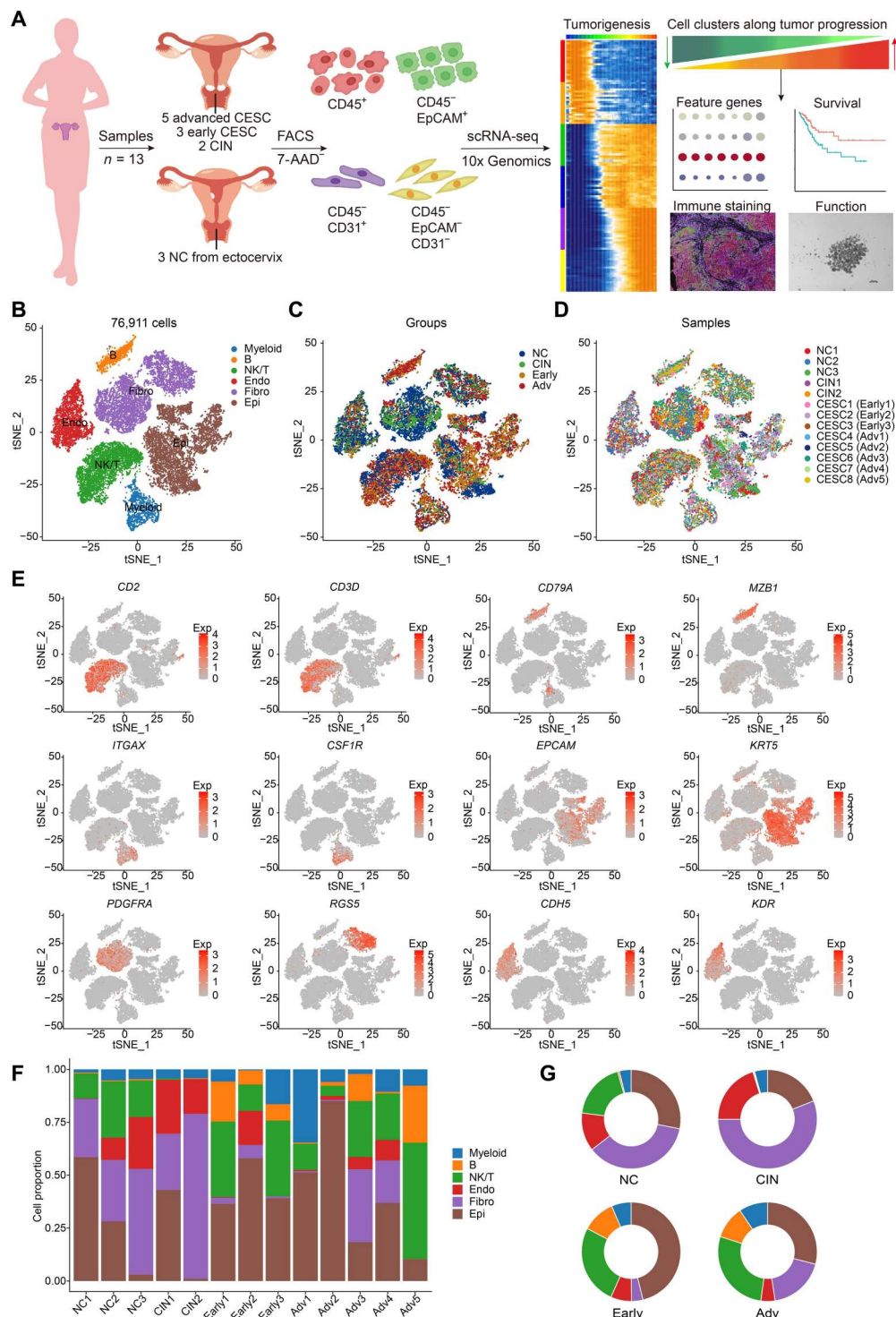


Fig. 1. scRNA-seq profile of 76,911 cells from human samples of NC, CIN, early CESC, and advanced CESC. (A) Study design. **(B to D)** t-distributed stochastic neighbor embedding (tSNE) plots of the 76,911 cells sequenced here, colored by cell type (B), disease stage groups (C), and individual samples (D). **(E)** tSNE plots showing marker gene expression for cell type identification. Legend shows a color gradient of normalized expression. **(F and G)** Proportions of each cell type in each sample (F) and in disease stage groups (G), colored by cell types.

suspensions, we subjected the samples to fluorescence-activated cell sorting (FACS) to exclude dead cells and to enrich immune (CD45⁺), endothelial (CD45⁺CD31⁺), epithelial (CD45⁺EPCAM⁺), or fibroblastic (CD45⁺EPCAM⁺CD31⁺) cells before scRNA-seq; this ensured that sufficient numbers of these cells were available for heterogeneity analyses. Following data quality control and filtering, we obtained scRNA-seq profiles for 76,911 cells with median gene and unique molecular identifier (UMI) counts of 2411 and 7555, respectively (fig. S1, A to C). Six major cell populations were identified by their expression of known lineage markers with 15,273 T/NK cells (CD2 and CD3D), 4463 B cells (CD79A and MZB1), 4324 myeloid cells (ITGAX and CSF1R), 24,343 epithelial cells (EPCAM and KRT5), 20,567 fibroblasts [platelet-derived growth factor receptor A (PDGFRA) and RGS5], and 7941 ECs (CDH5 and KDR) (Fig. 1, B to E, and fig. S1D). The relative abundance of the six major cell populations in the 13 samples and four disease stage groups is shown in Fig. 1 (F and G).

Transcriptional evolution of epithelial cells during CESC initiation and progression

To assess whether our samples from four continuous disease stage groups (NC, CIN, early CESC, and advanced CESC) represented the progress of tumorigenesis and progression of CESC, we examined the expression level of genes that were previously reported in cervical squamous epithelial cells at different stages of disease. Known squamous lineage markers, such as *KRT5*, *TP63*, *KRT17*, and *SOX2* (16, 17), were highly expressed in epithelial cells from all groups (fig. S2A), while the expression of cervical columnar lineage markers, such as *SOX17*, *MUC5B*, *MUC5AC*, and *TFF3* (17–20), was low among these cells (fig. S2B). This confirmed that cells from the four continuous disease stage groups were of squamous lineage. As a proven marker highly expressed in normal cervical squamous cells, *KRT14* (16, 18) showed high expression in NC samples but low expression in early and advanced CESC; in contrast, expression of malignant squamous cell marker (*KRT17*) and the canonical proliferation markers (*CDKN2A*, *MKI67*, and *TOP2A*) (18, 21–23) progressively increased from CIN to early and advanced CESC (Fig. 2A). This is consistent with the tumorigenic process and progression of CESC.

Gene expression profile analysis further revealed multiple significantly differentially expressed genes (DEGs) and differentially expressed transcription factors (TFs) between the squamous epithelial cells from the four disease stage groups (Fig. 2, B and C, and table S2). We identified numerous previously unknown feature genes uniquely expressed in each of the four disease stages: *SNHG29*, *ATF3*, *SNHG5*, *NR4A1*, *GADD45B*, and *PRRX1* in NC; *FGFBP1*, *LY6D*, *HMGCS1*, *BARX2*, *GRHL1*, *ZNF770*, *GRHL3*, *KLF8*, *ESRP2*, *ZNF219*, *FOXN1*, *LRRFIP1*, and *CYCS* in CIN; *TNIP2*, *GLUL*, *ANXA11*, and *ETV7* in early CESC; and *KCNK6*, *NDUFA4L2*, *ZNF519*, *ZNF146*, and *CEBPG* in advanced CESC.

We also used our scRNA-seq data to investigate the process of tumorigenic trajectory at single-cell resolution using Monocle analysis for trajectory inference (Fig. 2D). We found two different trajectories: one tumorigenic trajectory of squamous cells and another nonmalignant squamous cell development trajectory. The squamous cell subcluster at the beginning of the trajectory, epi1, was made up of *TP63*⁺*KRT5*⁺ basal stem cells (Fig. 2, D and E), which a previous study found to be the original *KRT5*⁺ squamous lineage of the ectocervix that gave rise to CESC (16). Thus, the epi1 cluster

might give rise to the malignant squamous cell subclusters, epi7 and epi8, with their increased expression of the malignant markers *MKI67* and *TOP2A* (Fig. 2, D and E) (22). On the other hand, epi1 could also differentiate into differentiated squamous cell subcluster, epi5, exhibiting more differentiated features including high expression of *SPRR3* and *SPRR2A* (Fig. 2, D and E). Furthermore, cell cycle analysis of epithelial cells showed gradual activation along the tumorigenic trajectory, which further confirmed CESC tumorigenesis (Fig. 2F).

We then went on to conduct pseudotime trajectory analysis of the changes in expression profile of epithelial cell genes and TFs along tumorigenic trajectory. This revealed five patterns of changing gene expression with pseudotime (Fig. 2, G and H, and table S3): Pattern 1 was characterized by high expression in normal squamous epithelial cells, decreasing with CESC initiation and progression, while gene expression in patterns 2 to 5 gradually increased but at different stages of tumorigenesis. Numerous genes that have not previously been linked with CESC progression were identified among the five pattern groups: *SNHG29*, *TLE5*, *AOPEP*, *SEPTIN7*, *RBIS*, and *ZFP36L2* (pattern 1); *FKBP2*, *GTF3A*, *RBCK1*, *TOX4*, *DNAJC2*, and *DNAJC1* (pattern 2); *PARP14*, *SMIM22*, *TSHZ2*, *ETV7*, *ZNFX1*, and *BBX* (pattern 3); *CENPW*, *MIS18BP1*, and *NFATC3* (pattern 4); and *KIF20B* and *MZT1* (pattern 5) (Fig. 2, G and H).

Characteristics of epithelial cell subclusters during CESC progression

We then asked how the epithelial cell subclusters changed during malignant progression. First, we found that the copy number variance scores of epithelial cells across four disease stages were significantly elevated during CESC progression (fig. S3A); then, using unsupervised dimensionality reduction and clustering, we identified eight epithelial cell subclusters exhibiting distinct gene expression profiles (Fig. 3, A to C, and table S4). We saw that epi1 and epi3 were less abundant in CESC compared to NC, while epi7 and epi8 subclusters were enriched in early and advanced CESC groups compared to NC or CIN groups. This was confirmed by signature gene set scores of these clusters compared between groups ($P < 0.0001$, two-sided unpaired Wilcoxon rank sum test, Fig. 3D and fig. S3B). Gene set variation analyses (GSVAs) were used to evaluate the transcriptomic characteristics of these subclusters. This revealed that, for example, epi1 was enriched for expression of genes involved in stemness, angiogenesis, and differentiation pathways, while epi7 was enriched for cell cycle, DNA repair, and tumor invasion pathways (Fig. 3E).

Focusing on the transcriptional profiles of epithelial cell subclusters that were enriched or less abundant in CESC tissues, we saw that epi8 exhibited interesting dual-immune features including expression of an immune suppressive gene *IDO1*; antigen-presenting genes such as *HLA-DRA*, *HLA-DPA1*, *HLA-A*, and *HLA-B*; interferon genes (*IFIT1*, *IFIT2*, and *IFIT3*); and chemokine genes (*CXCL9*, *CXCL10*, and *CXCL11*) (Fig. 3F). Therefore, we speculated that these cells might be effective presenters of tumor antigens on major histocompatibility complex I (MHC I) and MHC II and could attract T cells into the tumor microenvironment via *CXCL9-11* (24); this latter point was further supported by the high expression of the corresponding genes encoding the receptors (*CXCR3*, *ACKR3*, *SDC4*, and *CCR3*) for these chemokines in T cells (Fig. 3, G and H). We further confirmed the presence of epi8 in

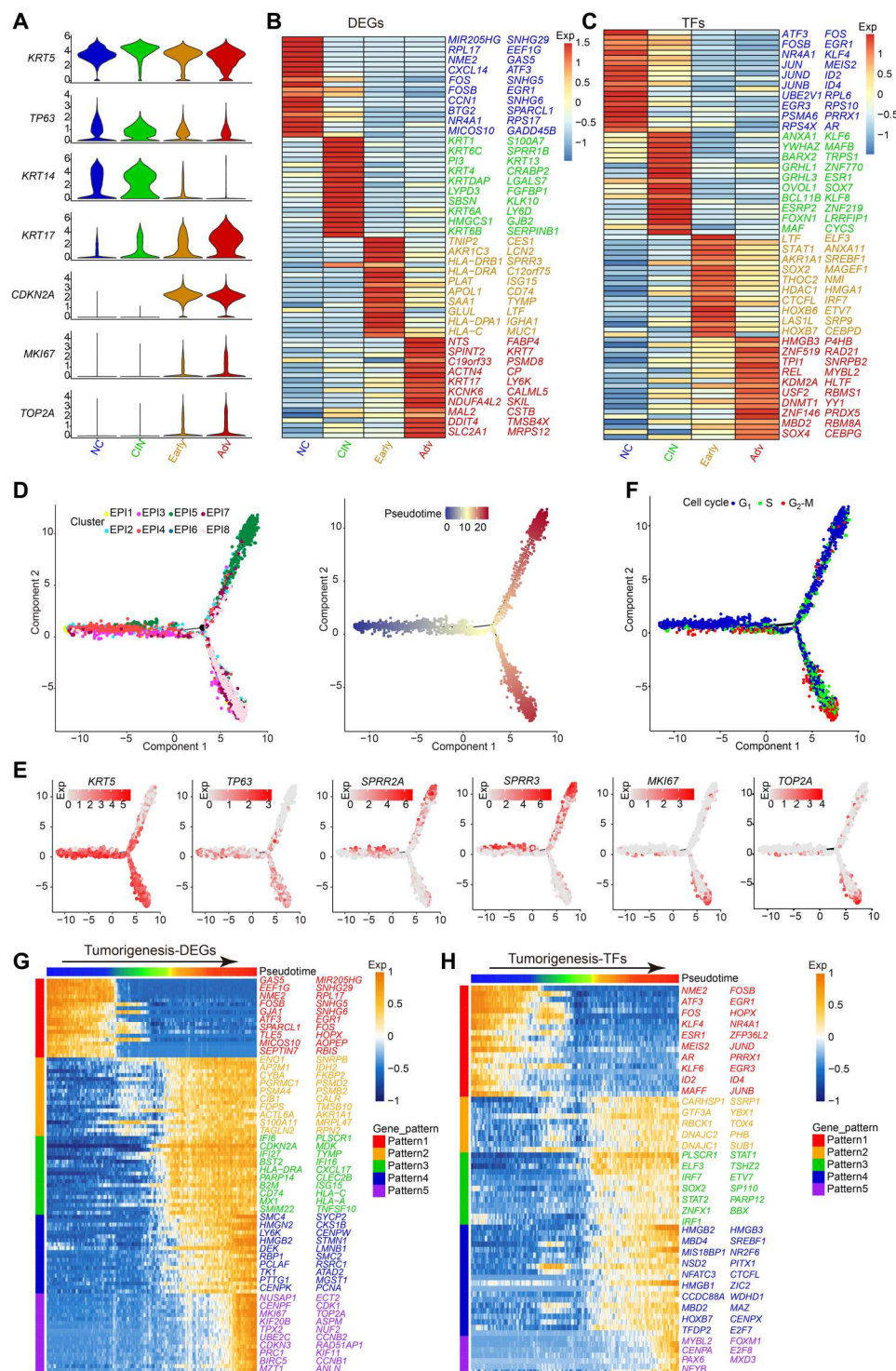


Fig. 2. Molecular profile of squamous epithelial cells during CESC initiation and progression by scRNA-seq. (A) Violin plots showing the expression of previously reported marker genes in the four disease stage groups. (B and C) Heatmap showing relative expression level of top 20 DEGs (B) and differentially expressed TFs (C) between squamous epithelial cells from the four disease stage groups. Intensity of color indicates average expression of genes. (D) Monocle analysis for trajectory inference of the epithelial cell subclusters in all samples, colored by cell cluster. (E) Representative gene expression in squamous epithelial cells during CESC initiation and progression. Intensity of color indicates normalized gene expression. (F) Cell cycle stage of squamous epithelial cells during CESC progression, colored by cell cycle types. (G and H) Heatmap showing the gene expression patterns and top 20 DEGs (G) and differentially expressed TFs (H) in each pattern along the pseudotime of CESC initiation and progression, colored by gene patterns. Intensity of color as in (B).

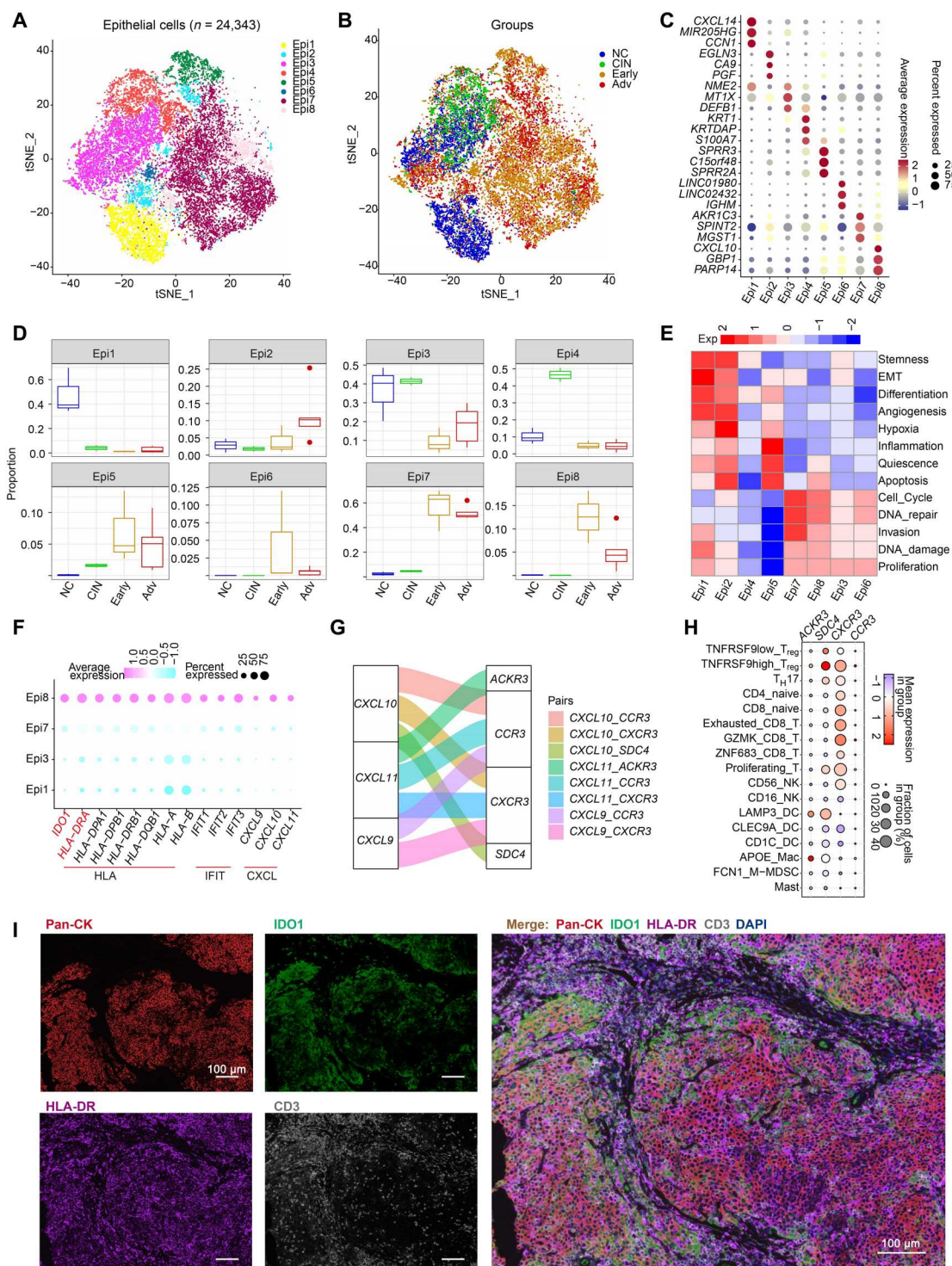


Fig. 3. Characteristics of squamous epithelial cell clusters during CESC progression. (A and B) tSNE plots showing all squamous epithelial cells colored by cell clusters (A) and four disease stage groups (B). (C) Dot plot showing top three DEGs from each of the eight epithelial cell subclusters. Intensity of color indicates average expression of genes. Dot size represents the percentage of cells expressing the gene. (D) Box plots showing proportions of epithelial cell subclusters in the four disease stage groups. (E) Heatmap showing gene set variation analysis scores of tumor characteristics in the eight squamous epithelial cell subclusters. Intensity of color indicates scores. (F) Feature gene expression of Epi8 subcluster. Dot size and colored scale as in (C). (G) Receptor-ligand pair analysis of CXCL9 to 11 for Epi8, colored by receptor-ligand pairs. (H) Dot plot showing the expression of receptors in immune cells. Dot size and colored scale as in (C). (I) Representative multiplex immunofluorescent labeling of Pan-CK (red), HLA-DR (blue), IDO1 (green) for Epi8, and CD3 (gray) for T cells in sections from CESC samples.

CESC by assessing the expression of the protein products of the key marker genes above using multiplex immunofluorescence. We labeled CESC sections for Pan-CK, IDO1, and HLA-DR, as well as for CD3, to detect T cells (Fig. 3I). While *IDO1* expression has been linked with immune tolerance in cancer (25–27), recent data from a study of human melanoma cells showed that IDO1 could in fact increase the recognition of tumor cells in the presence of proinflammatory cytokines (28). Thus, we propose that high relative abundance of epi8 in CESC could promote the high levels of T cell infiltration we observed and thereby enhance cross-talk with tumor cells.

Anti-inflammatory macrophages during CESC progression

Alongside epithelial cells, myeloid cells play an important role in shaping the tumor microenvironment (29–31). We analyzed these myeloid cells in our samples and identified eight subclusters exhibiting distinct patterns of gene and TF expression: two macrophage subclusters [(*CCL20*⁺ Mac (*CCL20*⁺) and *APOE*⁺ Mac (*APOE*⁺)], one subcluster of monocytic myeloid-derived suppressor cells (M-MDSCs) (*FCN1*⁺ M-MDSC (*FCN1*⁺), one subcluster of mast cells (*CPA3*⁺), and four dendritic cell (DC) subclusters [(*CD1C*⁺ DC (*CD1C*⁺), *LAMP3*⁺ DC (*LAMP3*⁺), *CLEC9A*⁺ DC (*CLEC9A*⁺), and pDC (*IL3RA*⁺)] (Fig. 4, A to D, and table S5).

Analyzing the relative proportions of each subcluster across the different patient groups revealed that *CCL20*⁺ and *APOE*⁺ macrophages were specifically enriched in the advanced CESC group, while *CLEC9A*⁺ DC became less abundant with progression, as confirmed by signature gene set scores of these clusters compared between groups ($P < 0.0001$, two-sided unpaired Wilcoxon rank sum test; Fig. 4E and fig. S4A). We then evaluated whether the abundance of these subclusters had any association with patient survival using bulk RNA-seq data from CESC. This showed that the frequency of both *CCL20*⁺ and *APOE*⁺ macrophages was negatively correlated with survival ($P < 0.0001$ and $= 0.023$, respectively, log-rank test; Fig. 4F), while the abundance of *CLEC9A*⁺ DC was positively correlated with survival ($P = 0.005$, log-rank test; Fig. 4F). We next examined anti-inflammatory and pro-inflammatory features (32) of the myeloid cell subclusters and found that both *APOE*⁺ and *CCL20*⁺ macrophages exhibited the anti-inflammatory features of M2-like macrophages (Fig. 4G). To further investigate the features of *CCL20*⁺ and *APOE*⁺ macrophages in CESC, we compared their transcriptional profiles and pathway enrichment across the four continuous disease stage groups, showing their potential functional heterogeneity in connection with epithelial cells, by increased enrichment of pathways linked to epithelial-mesenchymal transition (EMT) and hypoxia-inducible factors across the four continuous disease stage groups (fig. S4, B and C).

As the frequencies of *CCL20*⁺ and *APOE*⁺ macrophages in CESC both showed negative correlations with patient survival, we further investigated genes whose expression was up-regulated in both of these cell types in cervical cancer compared to NC. This analysis highlighted the significantly higher expression of several common up-regulated genes, including *SPP1* and *APOE* (fig. S4D). In addition, we investigated the expression of molecules that are known to coregulate the genes expressed by *APOE*⁺ macrophages and *CCL20*⁺ macrophages and found that expression of *PRDM1*, *PPARG*, and *RREB1* was co-up-regulated and that their overall expression level correlated with poor survival of patients with CESC (Fig. 4, H and I). Together, these data show important correlations between the

abundance of specific myeloid cell types and patient survival from CESC; moreover, we have identified potential key trends in expression of subtype-specific and shared DEGs that warrant further investigation.

Immune suppressive NK and T cells during CESC progression

In addition to myeloid cells, tumor-infiltrating lymphocytes are key elements in the tumor microenvironment and can profoundly affect patient outcomes (33–35). On the basis of their transcriptomic profiles, we identified two NK cell subclusters consisting of *CD16*⁺ NK cells (*FCGR3A*⁺) and *CD56*⁺ NK cells (*NCAM1*⁺). We also uncovered nine subclusters of T cells (*CD3D*⁺, *CD4*⁺, *CD8A*⁺): four *CD4*⁺ T subclusters consisting of naïve *CD4*⁺ T cells (*CCR7*⁺), T helper 17 (*T_H17*) cells (*IL17A*⁺), *TNFRSF9*^{high} regulatory T (*T_{reg}*) cells (*TNFRSF9*^{high}), and *TNFRSF9*^{low} *T_{reg}* cells (*TNFRSF9*^{low}); four subclusters of *CD8*⁺ T cells consisting of naïve *CD8*⁺ T cells (*CCR7*⁺), *ZNF683*⁺ *CD8*⁺ T cells (*ZNF683*⁺), *GZMK*⁺ *CD8*⁺ T cells (*GZMK*⁺), and exhausted *CD8*⁺ T cells (*LAG3*⁺); and one proliferating T cell cluster (*MKI67*⁺) (Fig. 5, A to C, and table S6).

Comparing the proportions of NK subclusters between patient groups revealed that there were relatively fewer *CD16*⁺ NK cells in the early and advanced CESC groups than in the NC group, which was confirmed by signature gene set scores of these clusters compared between groups ($P < 0.0001$, two-sided unpaired Wilcoxon rank sum test; Fig. 5D and fig. S5A). Then, we evaluated whether the abundance of NK subclusters was linked with patient survival using bulk RNA-seq data of CESC: We found that that high *CD16*⁺ NK and low *CD56*⁺ NK cell frequencies were correlated with longer survival of patients with CESC ($P < 0.05$, log-rank test; Fig. 5E), indicative of a potentially antagonistic role of each subset. To investigate this possibility further, we performed DEG and Gene Ontology (GO) term analyses of the two subsets, revealing that *FGFBP2*, *NKG7*, and *GZMH* were more highly expressed in *CD16*⁺ NK cells than in *CD56*⁺ NK cells, as were GO pathways mediating cellular cytotoxicity (Fig. 5, F and G, and table S7). These results suggest a possible deficit of antitumor immune responses mediated by *CD16*⁺ NK cells in advancing CESC.

Comparing the proportions of T cell subclusters between patient groups revealed that *T_H17* and *TNFRSF9*^{high} *T_{reg}* cells were relatively more abundant in early and advanced CESC groups compared to NC and CIN, which was confirmed by signature gene set scores of these clusters compared between groups ($P < 0.05$, two-sided unpaired Wilcoxon rank sum test; Fig. 5D and fig. S5A). Previous studies demonstrated that the abundance of *T_H17* and *TNFRSF9*^{high} *T_{reg}* was associated with CESC progression (36, 37). We therefore asked about the possible relatedness of these cells based on their transcriptional similarities and differences during the progression of CESC. Using trajectory analysis, we found that naïve *CD4*⁺ T cells underwent divergent differentiation into *T_H17* and *TNFRSF9*^{high} *T_{reg}* cells (Fig. 5H); however, both cell populations exhibited high and specific expression of *CCL20* (Fig. 5I), which has been shown in fibroblasts to promote CESC progression (37). Here, we found that *CCL20* expression level in bulk RNA-seq data from CESC patient samples correlated with poor prognosis ($P = 0.002$, log-rank test; Fig. 5J). These findings were supported by a broader evaluation of the expression of immune-feature gene sets in the T cell subclusters, which revealed that *TNFRSF9*^{high} *T_{reg}* cells exhibited high immune inhibition scores (fig. S5, B and C). Thus, we show

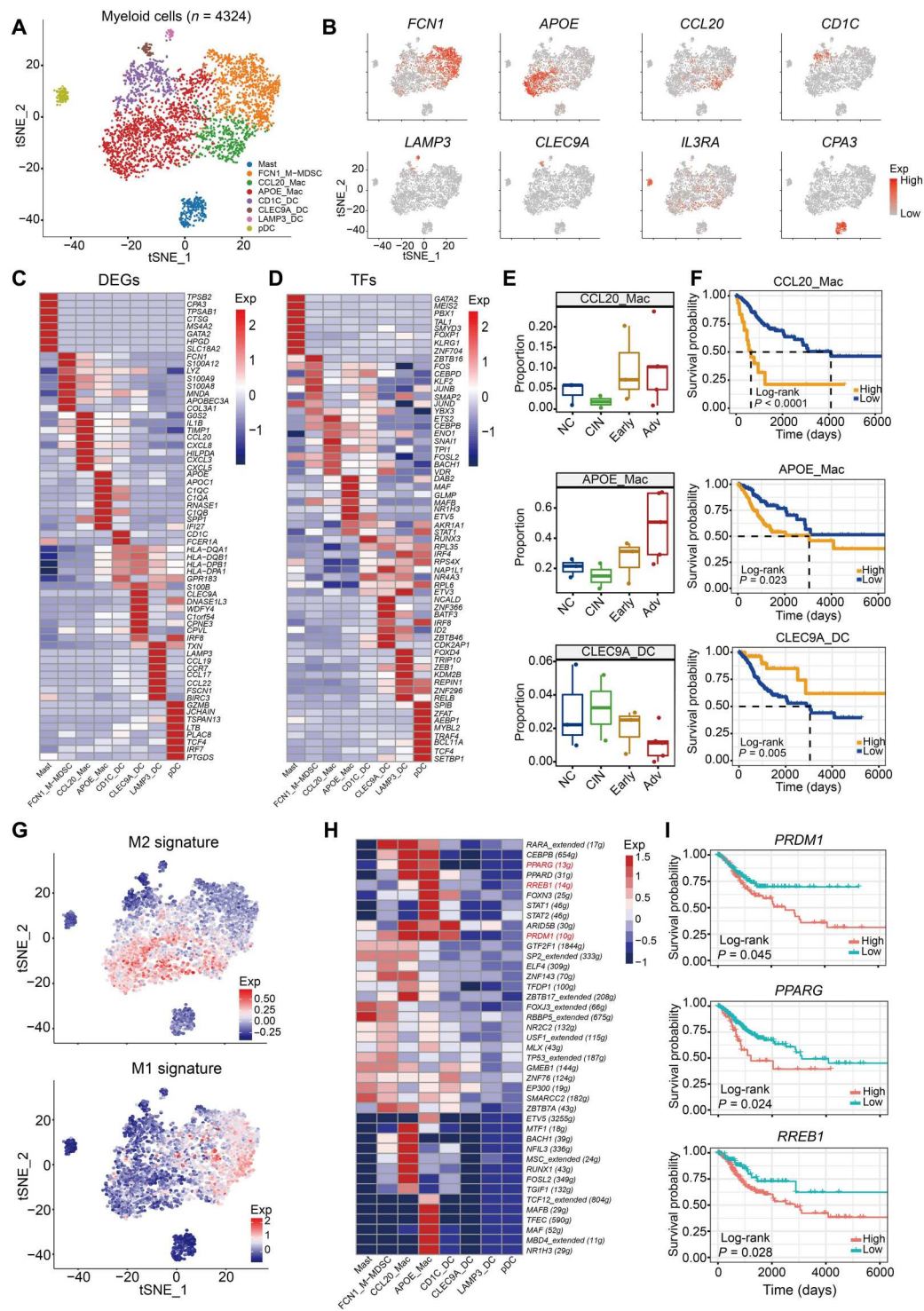
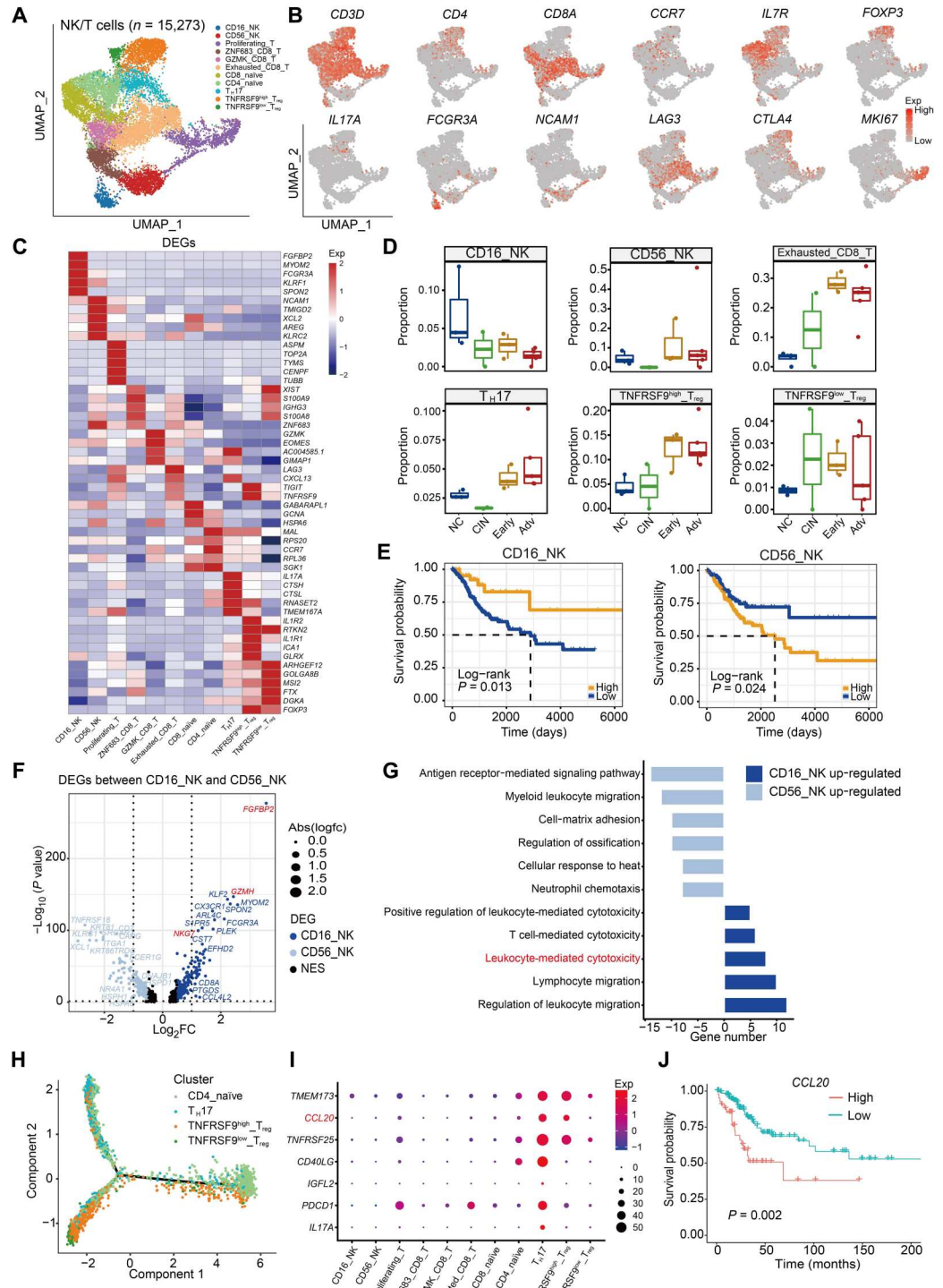


Fig. 4. Characteristics of myeloid cells during CESC progression. (A) tSNE plot showing all myeloid cells, colored by cell clusters. (B) tSNE plots showing the expression of selected marker genes for identification of myeloid cell clusters. Intensity of color indicates normalized gene expression. (C and D) Heatmap showing top eight DEGs (C) and differentially expressed TFs (D) of cell subclusters. Intensity of color indicates average expression of genes. (E) Box plots showing proportions of *CCL20*⁺ Mac, *APOE*⁺ Mac, and *CLEC9A*⁺ DC in the four disease stage groups. (F) Kaplan-Meier survival analysis of TCGA CESC patients stratified by high and low *CCL20*⁺ Mac, *APOE*⁺ Mac, and *CLEC9A*⁺ DC (with optimal cutoff value assigned using CutoffFinder). The *P* value of two-sided log-rank test is shown. (G) tSNE plots showing the expression scores of M1- and M2-like macrophage signatures in myeloid cells. Intensity of color indicates the score. (H) Heatmap showing expression of genes involved in regulating gene expression among distinct myeloid cell clusters, colored by cell clusters. Intensity of color as in (D). (I) Kaplan-Meier survival analysis of TCGA CESC patients stratified by high and low *PRDM1*, *PPARG*, and *RREB1* overall expression, which is defined as in (F). The *P* value as in (F).

Fig. 5. Characteristics of NK and T cells during CESC progression.

(A) Uniform Manifold Approximation and Projection (UMAP) plots showing all NK and T cells, colored by cell clusters. (B) UMAP plots showing the expression of selected marker genes for defining NK and T cell clusters. Intensity of color indicates normalized gene expression. (C) Heatmap showing top five DEGs of each cell sub-cluster. Intensity of color indicates average expression of genes. (D) Box plots showing proportions of several T and NK cell subclusters in the four disease stage groups. (E) Kaplan-Meier survival analysis of TCGA CESC patients stratified by high and low *CD16⁺* NK/*CD56⁺* NK, which is defined as in Fig. 4F. The *P* value of two-sided log-rank test is shown. (F) Volcano plot showing expression of DEGs between *CD16⁺* NK and *CD56⁺* NK, colored by cell cluster. (G) Gene ontology (GO) term analyses of *CD16⁺* NK and *CD56⁺* NK gene expression, colored by cell clusters. (H) Trajectory analysis of *CD4⁺* T cells inferred by Monocle v.2, colored by cell cluster. (I) Feature gene expression of *T_H17* and *TNFRSF9^{high}* *T_{reg}* cells. Dot size represents the fraction of cell expressing genes. Intensity of color indicates average expression of genes. (J) Kaplan-Meier survival analysis of TCGA CESC patients stratified by high and low *CCL20* expression, which is defined as in Fig. 4F. The *P* value of two-sided log-rank test is shown.



that the frequency of *T_H17* and *TNFRSF9^{high}* *T_{reg}* are elevated during CESC progression and that the total *CCL20* level is associated with shorter survival.

Characteristics of EC subclusters during CESC progression

ECs play a pivotal role in tumor angiogenesis, supporting growth and metastasis of cancer (38). Using unsupervised dimensionality

reduction and clustering, we identified three EC subclusters with distinct patterns of marker gene expression: *PODXL⁺* EC (*PODXL⁺* and *INSR⁺*), *MHC^{high}* EC (*HLA-DPA1⁺* and *SELP⁺*), and *GJA4⁺* EC (*GJA4⁺* and *CXCL12⁺*) (Fig. 6, A to C). These EC subclusters also showed more broadly distinct gene and TF expression profiles (Fig. 6, D and E, and table S8). Analysis of the relative proportions of each subcluster indicated that the *PODXL⁺* EC

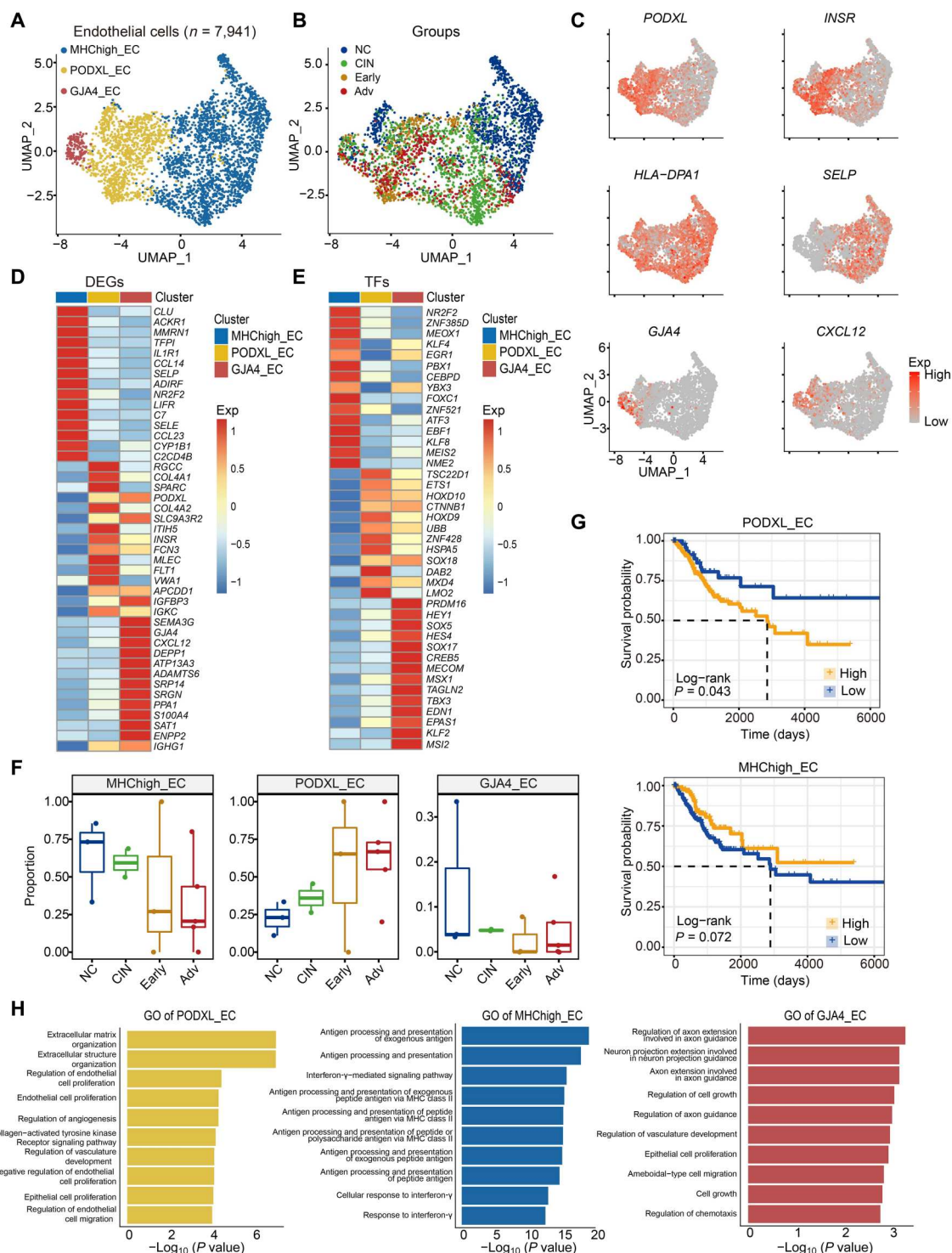


Fig. 6. Characteristics of ECs during CESC progression. (A and B) tSNE plots of all ECs, color-coded by cell type (A) and in the four disease stage groups (B). (C) tSNE plots showing the expression of selected marker genes for EC subclusters. Intensity of color indicates normalized gene expression. (D and E) Heatmap showing expression of top 15 DEGs (D) and differentially expressed TFs (E) in three EC subclusters, colored by cell cluster. Intensity of color indicates average expression of genes. (F) Box plots showing proportions of EC subclusters in the four disease stage groups. (G) Kaplan-Meier survival analysis of TCGA CESC patients stratified by *PODXL*⁺ EC/*MHC*^{high} EC, which is defined as in Fig. 4F. The *P* value of two-sided log-rank test is shown. (H) GO terms enriched in the three EC subclusters.

subcluster was enriched in early and advanced CESC groups compared to NC and CIN, while the abundance of the *MHC^{high}* EC subcluster was lower, which was confirmed by signature gene set scores of these clusters compared between groups ($P < 0.0001$, two-sided unpaired Wilcoxon rank sum test; Fig. 6F and fig. S6A). Survival analysis showed that the abundance of *PODXL⁺* EC was negatively correlated with CESC patients' survival ($P = 0.043$, log-rank test; Fig. 6G), while there was some evidence of a possible positive correlation between *MHC^{high}* EC frequency and survival time ($P = 0.072$, log-rank test; Fig. 6G). We further revealed that the overall expression levels of the marker genes for these two EC subsets, *PODXL* and *SELP*, were both significantly correlated with patient survival with antagonistic role ($P < 0.05$, log-rank test; fig. S6B). To investigate possible reasons why these subsets showed opposing prognostic indications in CESC, we performed pathway enrichment analysis. This revealed that *PODXL⁺* ECs showed enrichment for transcripts associated with EC and epithelial cell proliferation, while *MHC^{high}* ECs exhibited enrichment for transcripts associated with interferon- γ , antigen-processing, and presentation pathways (Fig. 6H). Therefore, it is possible that the potential antitumor effect mediated by *MHC^{high}* ECs is deficient, while the protumor effect mediated by *PODXL⁺* ECs is enhanced, in cases of advancing CESC.

Characteristics of fibroblast subclusters during CESC progression

Further enriching the emerging picture of CESC progression, we identified six fibroblast subclusters across four disease stage groups: two subsets of normal matrix fibroblasts (Fib1 and Fib2), mainly derived from NC samples; one subset of mCAFs; and three subsets of vascular cancer-associated fibroblasts (vCAFs; comprising vCAF1, vCAF2, and vCAF3), mainly derived from CESC and CIN samples (Fig. 7, A and B). mCAFs were characterized by high expression of extracellular matrix signature genes including *PDGFRA*, *LUM*, and *DCN*, while vCAFs showed high expression of *MCAM* (CD146), *RGS5*, and *MYH11* (Fig. 7C and fig. S7, A to C). Among these, we confirmed the expression of *PDGFRA* and melanoma cell adhesion molecule (MCAM) by multiplex immunofluorescence labeling of CESC sections, which also revealed that mCAFs were predominantly located within the collagen-rich stromal streaks (Fig. 7D). The fibroblast subclusters also exhibited more broadly distinct transcriptomic features (Fig. 7E and table S9): For example, mCAFs expressed higher levels of some genes, including *VCAN*, *LUM*, *TWIST2*, *NRG1*, *SFRP2*, *MMP11*, and *WNT5A*, than Fib1 and Fib2. Notably, mCAFs expressed high levels of genes involved in EMT and vital pathways for cancer development and progression, such as regulation of WNT signaling and hypoxia (Fig. 7F). In addition, expression of genes associated with cancer stemness and/or progression or metastasis-associated TFs, such as *ZEB1*, *RUNX1*, *HIF1A*, *SIX2*, *TCF12*, and *TWIST2*, was specifically enriched in mCAFs (Fig. 7G). Notably, mCAF abundance was significantly correlated with poor survival of patients with CESC (Fig. 7H), indicating that mCAFs may promote CESC progression.

Cell clusters in the CESC ecosystem and their correlation with patient survival

While characterization of individual cell subtypes is important for advancing our in-depth knowledge of disease, the immune system

does not operate in a compartmentalized way, and so we must also consider the bigger picture, which we term here the "CESC ecosystem." Therefore, we evaluated the correlation between all cell clusters' abundance with patient survival using bulk RNA-seq data of CESC from The Cancer Genome Atlas (TCGA) (Fig. 7I). The relative abundance of several cell clusters showed significant correlation with shorter survival: *CCL20⁺* Mac, *APOE⁺* Mac, epi7, *CD56⁺* NK, *T_H17*, exhausted *CD8⁺* T, *PODXL⁺* EC, *TNFRSF9^{high}* T_{reg}, and mCAF (Fig. 7I). In contrast, the abundance of other cell clusters was significantly associated with longer survival: pDC, *CD16⁺* NK, *GZMK⁺* *CD8⁺* T, *ZNF683⁺* *CD8⁺* T, *CLEC9A⁺* DC, epi8, and mast cells (Fig. 7I). Together, these findings represent a comprehensive immune signature with potential prognostic significance that warrants further investigation.

mCAFs promote CESC progression primarily through the NRG1/ERBB3 pathway

CAFs promote tumor progression by secreting multiple cytokines, chemokines, and growth factors, which are essential for cancer cell survival and metastasis (39); however, in many cases, the precise cellular targets for these CAF-produced factors are unknown. Here, we found that mCAFs were enriched in CESC and CIN samples and had high expression of cancer progression-associated genes, regulators, and pathways, which was correlated with poor survival in CESC. Therefore, we further asked about the possible cellular interaction between mCAFs/normal matrix fibroblasts (Fib1 and Fib2) and tumor cells across the four disease stage groups (Fig. 8A and fig. S7D) using intercellular communication analysis (40). We uncovered distinct interaction levels across the four disease stage groups: the expression level of ligands including *NRG1* and *WNT5A* was higher in mCAFs derived from CIN and early and advanced CESC samples than in normal matrix fibroblasts (Fib1 and Fib2) from NC samples (Fig. 8, A and B, and fig. S7E), while tumor cells of CESC samples expressed the corresponding receptors, *ERBB3* and *FZD6*, respectively (Fig. 8, A and B). These transcriptional data were mirrored by immunolabeling of CESC sections, confirming that mCAFs mainly expressed neuregulin 1 (NRG1) and wnt family member 5A (WNT5A), while tumor cells correspondingly expressed high levels of the receptors including *ERBB3* and *FZD6* (Fig. 8C and fig. S8, A and B).

The results from the intercellular communication analysis led us to further explore the interactions between mCAFs and tumor cells mediated by NRG1-ERBB3 and WNT5A-FZD6. We first isolated mCAF subpopulations from CESC samples by FACS based on *PDGFRA* expression (CD140a). Thereafter, quantitative reverse transcription polymerase chain reaction (qRT-PCR), immunofluorescence labeling, and flow cytometry analysis indicated that the in vitro-cultured mCAFs also expressed canonical mesenchymal markers, including abundant expression of NRG1 and WNT5A (fig. S8, C to E), which corresponded well with the transcriptomic features of mCAFs via scRNA-seq analysis. We then asked about the functional profile of mCAFs using in vitro coculture with CESC cells. The presence of mCAFs significantly enhanced tumor sphere formation and the migration of CESC cells (Fig. 8D). Moreover, the presence of mitomycin C-treated mCAFs or inclusion of mCAF-conditioned medium significantly reduced inhibition of tumor sphere formation by the cytotoxic antitumor drugs carboplatin (CBP) and cisplatin (DDP) (Fig. 8E and fig. S8F). Thus, mCAFs

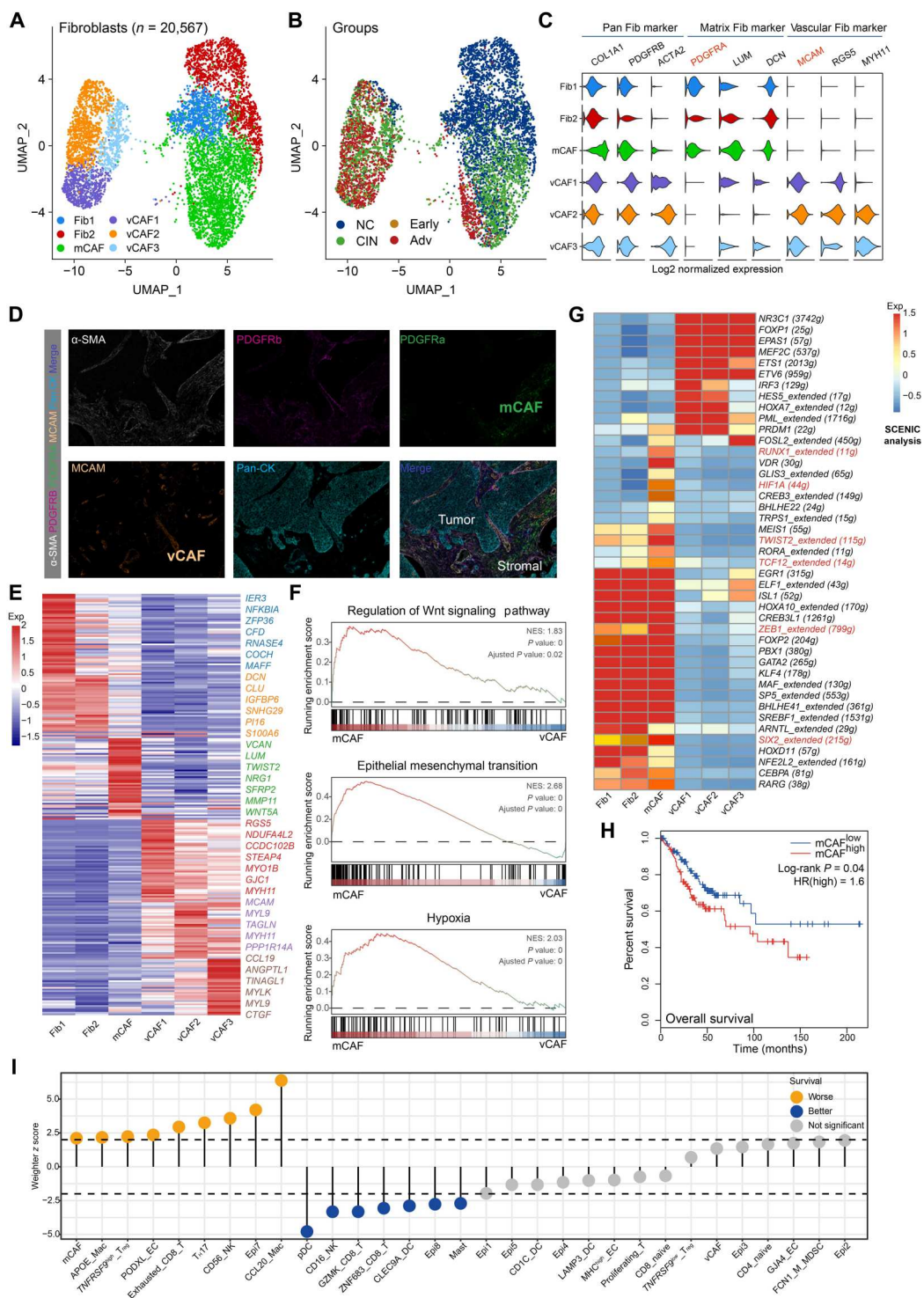


Fig. 7. Characteristics of fibroblasts during CESC progression. (A and B) UMAP plots showing total fibroblasts, color-coded by cell type (A) and the four disease stage groups (B). (C) Violin plots showing the expression of pan-fibroblast, matrix-fibroblast, and vascular-fibroblast markers across six fibroblast subclusters. (D) Multiplex immunofluorescent labeling of specific markers on mCAFs and vCAFs in sections from CESC samples. α -SMA, α -smooth muscle actin. (E) Heatmap showing relative expression of distinct genes in six fibroblast subclusters. (F) Differences in pathway activity (scored per cell by Gene Set Enrichment Analysis) in mCAFs and vCAFs. NES, normalized enrichment score. (G) Heatmap showing the expression of key regulatory TFs (estimated using SCENIC) between different fibroblast subclusters. (H) Kaplan-Meier survival analysis of TCGA CESC patients stratified by high and low mCAFs (with optimal cutoff value assigned using CutoffFinder). The P value of two-sided log-rank test is shown. (I) Correlation between abundance of the annotated cell clusters with patient survival in CESC. The Cox proportional hazards model was conducted, which included age and tumor stage. The results of the Cox regression model among cell subtypes were visualized using weighted Z scores.

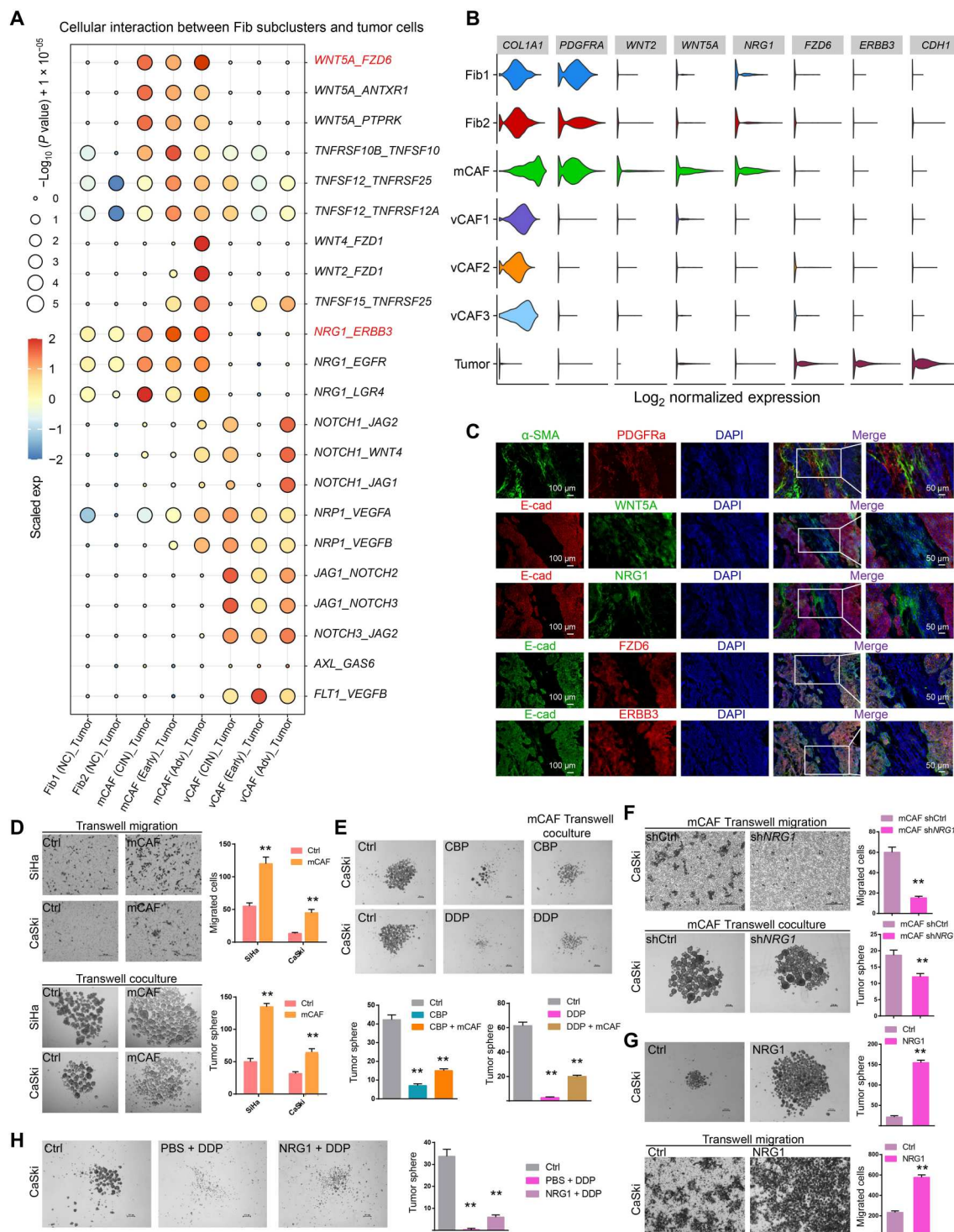


Fig. 8. ScRNA-seq profiling and functional experiments reveal intercellular communication between mCAFs and tumor cells in CESC. (A) ScRNA-seq profiling identifies potential ligand-receptor pairs between fibroblast subclusters and tumor cells of CESC across four disease stage groups. (B) Violin plots confirmed the enriched expression of ligand-receptor pairs in fibroblast subclusters and tumor cells of CESC. (C) Multiplex immunofluorescent labeling of mCAF tumor-specific ligand-receptor pairs, including NRG1-ERBB3 and WNT5A-FZD6, respectively. (D) Patient-derived mCAFs significantly enhanced the capacity of CESC cells to migrate and form tumor spheres in vitro. (E) mCAFs induced drug resistance of CESC cells in the sphere formation assays. (F) Lentivirus-mediated RNA interference of NRG1 expression in mCAFs significantly rescued their migration- and sphere formation-promoting effects on CESC cells. (G and H) Exogenous addition of NRG1 enhanced tumor sphere proliferation, migration, and drug resistance of CESC cells. $**P < 0.01$ (two-sided unpaired *t* test).

appear to both directly and indirectly promote tumor sphere formation.

NRG1 was expressed by CAFs at high levels and has been shown to promote antiandrogen resistance in prostate cancer (41). Silencing NRG1 in mCAFs significantly repressed their promotion of tumor sphere formation and tumor cell migration (Fig. 8F and fig. S9A). Moreover, the addition of exogenous NRG1 significantly promoted tumor sphere formation, migration, and drug resistance (Fig. 8, G and H, and fig. S9B). In addition, silencing ERBB3 in CESC cells or adding an ERBB3 inhibitor, TX1-85-1, significantly repressed their sphere formation ability (fig. S9, C to E). Moreover, lentivirus-mediated silencing of WNT5A in mCAFs (fig. S10A) resulted in a significant reduction in CESC cell migration in coculture but not of tumor sphere formation (fig. S10, B and C). These results were further verified by the exogenous addition of WNT5A or silencing of the receptor FZD6 in CESC cells (fig. S10, D to F). Thereafter, to explore whether NRG1 and WNT5A exhibit a synergistic effect on tumor cell migration, we simultaneously silenced WNT5A and NRG1 in the mCAFs and found that tumor migration ability was notably repressed (fig. S11, A and B). This synergistic migration-promoting effect was verified by simultaneous addition of exogenous WNT5A and NRG1 (fig. S11C). However, this approach did not significantly promote drug resistance compared with NRG1-alone group (fig. S11D). Collectively, these results indicate that both NRG1 and WNT5A produced by mCAFs act in a paracrine manner to enhance tumor cell migration, while NRG1 from mCAFs promotes proliferation and drug resistance in CESC cells.

DISCUSSION

Although numerous bulk transcriptome analyses have been conducted on CESC, these approaches have been unable to provide high-resolution insights into the main cellular players, their interaction partners, and key molecular pathways driving disease initiation and progression. Here, we comprehensively investigated the dynamic cellular and molecular alterations that occur in specific cell subpopulations during the various stages of human CESC at single-cell resolution. The data we generated represent a valuable resource for improving CESC diagnosis, refining patient stratification, and informing future studies aimed at the discovery of potential previously unidentified therapeutic targets.

Previous scRNA-seq studies that investigated the heterogeneity of tumor cells and stromal cells in human cancers (24, 29, 31, 38, 42–49) have only captured snapshots of the overall picture. In this study, we investigated tumor initiation and progression using samples of CESC from human patients enriched for a broad range of specific cell types and subjected to scRNA-seq. Human CESC begins with normal cervical columnar cells being replaced by metaplastic squamous cells, which, in turn, become malignant squamous cells (7). Here, we describe the molecular characteristics and reveal new markers that are expressed during CESC initiation and progression. After validating our approach by confirming expression of known marker genes (17, 50), we defined five patterns of changing gene expression during CESC progression. Among the genes exhibiting these clear patterns, we identified numerous candidates warranting further investigation for their roles in initiation and progression of CESC, including *FKBP2*, *GTF3A*, *RBCK1*, *TOX4*, *DNAJC2*, *DNAJC1*, *PARP14*, *SMIM22*, *TSHZ2*, *ETV7*, *ZNFX1*, *BBX*, *CENPW*, *MIS18BP1*, *NFATC3*, *KIF20B*, and *MZT1*.

Our scRNA-seq data thus provide a comprehensive resource showing the changing molecular characteristics of epithelial cells during CESC initiation and progression.

We also exploited the resolution of scRNA-seq to ask directly about epithelial cell heterogeneity, which is not possible to study using bulk transcriptomic approaches. We identified eight epithelial cell subclusters, among which epi7 to 8 were specifically enriched in early and advanced CESC and exhibited the expression of genes involved in active tumor-associated pathways such as cell cycle, DNA repair, and tumor invasion, consistent with previous studies (51). Epi8 expressed both potentially immune-suppressive (*IDO1*) and proinflammatory (*HLA-DRA*, *HLA-DPA1*, *IFITs*, and *CXCL9-11*) genes and was predicted to communicate directly with T cells by ligand-receptor interactions. A similar immune-featured epithelial cell cluster has been reported in nasopharyngeal carcinoma (52), but our highly detailed characterization indicates a more complicated profile whose functional outcome requires further investigation.

Among myeloid cell subclusters, two subclusters of M2-like macrophages were enriched in advanced CESC, which was also seen in a recent renal cell carcinoma study (30). We further found that the abundance of *CCL20*⁺ and *APOE*⁺ macrophages and the expression level of several coregulatory molecules were associated with shorter survival of patients with CESC. *CD16*⁺ and *CD56*⁺ NK cell clusters showed different functional enrichment and opposing prognostic roles: Abundance of the *CD16*⁺ NK cell cluster was associated with improved survival, and this cell population was typically deficient in advanced CESC. Among T cell subclusters, *T_H17* and *TNFRSF9*^{high} *T_{reg}* were enriched in advanced CESC, which was consistent with the previously reported role of *T_H17* in promoting tumor development (37); we further found that their coexpression of the gene *CCL20* was associated with poor survival. Among EC subclusters, *PODXL*⁺ and *MHC*^{high} ECs exerted opposite effects during CESC progression: *PODXL*⁺ ECs were enriched in advanced CESC, and their abundance was negatively associated with CESC patient survival, while *MHC*^{high} ECs were decreased in advanced CESC and positively correlated with CESC patient prognosis. Functional enrichment analyses of the two EC subclusters demonstrated their opposing roles in the CESC environment: *PODXL*⁺ ECs were enriched for expression of genes involved in EC and epithelial cell proliferation, while *MHC*^{high} ECs were enriched for the function of antigen processing and presentation.

As a key component in the tumor microenvironment, CAFs promote malignant growth and invasion by diverse mechanisms (39). However, the heterogeneity of CAFs has so far challenged the investigation of their cross-talk with tumor cells. We have previously reported the role of vCAFs in promoting intrahepatic cholangiocarcinoma proliferation by interleukin-6 using scRNA-seq analysis and functional assays (53). Here, we revealed two main subclusters of mCAFs and vCAFs and validated their presence in human CESC by multiplex immunofluorescent labeling. We found mCAFs expressing high levels of genes associated with pro-tumor pathways, including hypoxia and EMT. Intercell interaction analysis revealed that mCAFs could promote CESC progression primarily through the NRG1/ERBB3 pathway, which is involved in antiandrogen resistance to prostate cancer (41) and RAF inhibitor resistance to mutant BRAF melanoma (54) but has not been reported in previous studies on CAFs in CESC (36, 55, 56). Further functional experiments demonstrated that mCAFs enhanced CESC

proliferation, migration, and chemotherapy resistance via the NRG1/ERBB3 pathway.

While these data represent an important advance in knowledge, future studies should take into account the limitations of the current work. First, the sample size was small, and larger-scale studies would be beneficial to confirm and extend our findings. Second, spatial CESC information, which is important for understanding the location of cell clusters and their interactions, was lost during the dissociation process for scRNA-seq analysis. Using the data that we have generated here, more detailed microscopy studies on human tissues can be directed toward answering specific questions that could not be addressed by scRNA-seq. Third, further validation of our findings on CESC samples is needed, including the molecular profile of cells during CESC initiation and progression.

In conclusion, our study provides a comprehensive map of the cellular and molecular characteristics of CESC during tumorigenesis and progression. By showing the dynamic changes of the molecular profiles and relative abundance of epithelial cells and various cell clusters involved in CESC initiation and progression, we provide potential therapeutic targets, as well as molecular and cellular diagnostic and prognostic biomarkers for CESC.

MATERIALS AND METHODS

Patients and sample collection

This study was approved by the Ethical Committee of Shandong Cancer Hospital and Institute (SDTHEC201906009). Ten patients with pathologically confirmed cervical abnormality (eight patients with CESC and two with CIN) and three patients with endometrial cancer, who were treated at Shandong Cancer Hospital, were enrolled into the study. All patients provided written informed consent. Eight CESC samples were collected by surgery or biopsy and divided into three early and five advanced CESC according to the 2018 FIGO stage (5). Two samples of CIN cervical tissue were collected by cervical loop electrosurgical excision. Three samples of normal cervical squamous tissues were collected from the ectocervix by surgery from three patients with endometrial cancer.

Sample processing and cell sorting

Fresh samples were stored in MACS Tissue Storage Solution (Miltenyi Biotec) on ice and immediately transferred to the laboratory. After washing with phosphate-buffered saline (PBS), the samples were enzymatically digested to obtain single-cell suspensions. Following similar procedures previously described (57, 58), the samples were minced into $<1\text{ mm}^3$ pieces and digested with 5 ml of digestion buffer containing collagenase IV (2 mg/ml; Sigma-Aldrich) and deoxyribonuclease I (1 mg/ml; Sigma-Aldrich) for 30 min at 37°C. Then, the resulting suspension was mixed with 5 ml of 2% fetal bovine serum (FBS)/PBS and centrifuged at 300g for 5 min at 4°C to get the cell pellet. Next, the cell pellet was mixed with red blood cell lysis buffer (BD) for 3 min at room temperature and centrifuged at 300g for 5 min at 4°C. After that, the cells were incubated with antibodies recognizing EpCAM (BD Biosciences), CD45 (BD Biosciences), and CD31 (BD Biosciences) for 30 min at 4°C, followed by staining with 7-aminoactinomycin D (7-AAD) (eBioscience) before cell sorting. Then, we sorted and collected 7-AAD-negative cells for the following scRNA-seq, including immune cells (CD45⁺), ECs (CD45⁺CD31⁺), epithelial cells

(CD45⁺EpCAM⁺), and fibroblasts (CD45⁺EpCAM⁺CD31⁺), using a BD FACSARIA II (BD).

Single-cell library preparation and sequencing

Following similar procedures previously described (57, 58), the Single-Cell 3' Library Kit v3 (10x Genomics) was used for single-cell transcriptome amplification and library preparation according to the manufacturer's instructions. The sorted single-cell suspension was loaded onto a microfluidic chip from 10x Genomics to generate the cDNA library. Next, cDNA libraries were prepared and sequenced across six lanes on an Illumina NovaSeq 6000 system (Illumina Inc., San Diego, CA, USA).

Preprocessing of scRNA-seq data

The data were processed as previously described (57, 58). Using the cellranger count function of CellRanger (10x Genomics, v4), raw sequencing FASTQ files were aligned to the GRCh38 reference genome to produce a gene expression library via the STAR algorithm. Next, the raw gene expression matrices were processed by the Seurat R package (version 3.1.4) (59). Included genes were expressed in at least 10 cells in a sample. Low-quality cells were removed according to the following criteria: Cells had fewer than 2001 UMIs, more than 6000 or less than 501 expressed genes, or over 20% of UMIs derived from the mitochondrial genome. Cell doublets were removed using the DoubletFinder R package. Then, we integrated the single-cell transcriptome expression matrices of the remaining high-quality cells using the harmony R package, normalized them to the total cellular UMI count, and scaled (scale factor = 1×10^4) them by regressing out the total cellular UMI counts and percentage of mitochondrial genes. After that, highly variable genes (HVGs) were selected for principal components analysis (PCA), and the top 30 significant principal components (PCs) were selected for Uniform Manifold Approximation and Projection (UMAP) and t-distributed stochastic neighbor embedding (tSNE) dimension reduction, and visualization of gene expression.

Determination of cell type

The DEGs of each cell subcluster were identified by the "FindAllMarker" function with default parameters provided by Seurat. The cell types and subtypes were annotated according to their expression of the known canonical marker genes of the respective cell types. Cell subclusters with similar gene expression patterns were annotated as the same cell type.

Trajectory analysis by Monocle 2 analysis

To further investigate the tumorigenesis of CESC, trajectory analysis was performed using Monocle 2. First, the DEGs of eight epithelial cell subclusters were identified and subjected to Monocle analysis. Then, dimensionality reduction and visualization were performed using "DDRTree" and "plot_cell_trajectory" functions in the Monocle 2 package. Monocle 2 analysis was also used to illustrate the process of CD4⁺ T cell differentiation. First, PCA analysis was performed using HVGs, and the top 50 correlating and anticorrelating genes of the top three PCs on the basis of gene loading were used to sort the cells in pseudotime order. Then, the DDRTree function was used for dimensionality reduction, and the plot_cell_trajectory function was used for visualization.

Pathway analysis

DEGs with $|\log FC| > 0.5$ and adjusted P value < 0.05 were used for GO enrichment analysis. The compareCluster function in the clusterProfiler R package was used to find different enriched GO terms between distinct fibroblast subclusters. To assess the different pathways between distinct fibroblast subsets, GSVA analysis was performed and calculated with a linear model offered by the limma package.

Survival analysis

Bulk RNA-seq data along with the curated clinical data from patients with CESC were obtained from the UCSC Xena database website (<https://xenabrowser.net/datapages/>). To evaluate the signatures of cell subtypes identified in our work with respect to CESC patients' survival, GSVA analysis was applied to calculate the combined expression value of the cell type-specific signatures (top 30 DEGs). The optimal low/high cutoffs for expression level of these marker genes were assessed by the CutoffFinder (<http://molpath.charite.de/cutoff/>), and patients were classified into gene expression -high or -low groups. Kaplan-Meier survival curves were plotted by the R package "survminer." Then, the Cox proportional hazards model was conducted that included age and tumor stage. The results of the Cox regression model among cell subtypes were visualized using weighted Z scores.

Single-cell regulatory network analysis

As previously described (58), we analyzed the differentially expressed TFs with a standard pipeline implemented in R using the SCENIC R package (<https://github.com/aertslab/SCENIC>) to explore the single-cell gene regulatory network (GRN) among different cell subclusters. Briefly, we estimated the gene expression matrices of the cell subclusters using GENIE3 to build the initial coexpression GRNs. Next, we analyzed the regulon data using the RcisTarget package to find TF motifs. Then, we calculated the regulon activity score of each cell by the AUCell package. Last, the regulons with a correlation coefficient of > 0.3 with at least one other regulon were filtered, and the regulons that were activated in at least 30% of the cell subclusters were selected for the subsequent visualization.

Differentiation of tumor cells and nonmalignant epithelial cells based on inferred CNV

To distinguish tumor cells and nonmalignant epithelial cells, the initial copy number variation (CNV) signal of each region (60) was estimated with default parameters using the inferCNV R package (<https://github.com/broadinstitute/inferCNV/wiki>). The quadratic sum of the CNV region was calculated as the CNV signal, and cells with a CNV signal of > 0.04 were defined as putative tumor cells in CESC.

Cell-cell interaction analysis using CellPhoneDB

To investigate the cell-cell interaction network among fibroblast subsets and CESC malignant cells, ligand-receptor pairs between niche cell subtypes and malignant cells were explored using the CellPhoneDB package (www.cellphonedb.org), as previously reported (40). We chose the receptors and ligands expressed in more than 10% of the cells in the specific cluster for subsequent analysis. The interaction pairs whose ligands belonged to the vascular endothelial growth factor, NOTCH, WNT, NRG1, TNFSF, or

GAS6 families and had P values < 0.05 returned by CellPhoneDB were selected for the evaluation of the relationships between the distinct fibroblast subsets and CESC malignant cells, as visualized using the ggplot2 package.

Multiplex immunofluorescent labeling

To identify epithelial cell subclusters [expressing pan-ck (pan-CK), indole 2,3-dioxygenase (IDO), HLA-DRA, and 4',6-diamidino-2-phenylindole (DAPI)] and fibroblast subclusters [expressing α -smooth muscle actin (α -SMA), MCAM, PDGFR α , PDGFR β , and DAPI] in the CESC tissues, we used a PANO 7-plex immunohistochemistry kit (0004100100, Panovue) for multiplex immunofluorescent labeling. We applied the above primary antibodies sequentially to CESC tissue sections, followed by incubation with horseradish peroxidase-conjugated secondary antibodies and tyramide signal amplification (TSA). After each TSA operation, the slides were treated with microwave heat. After labeling of all human antigens, the nuclei were stained with DAPI. To obtain multispectral images, the stained slides were scanned using the Mantra System (PerkinElmer), which captured fluorescence spectra at 20-nm wavelength intervals in the wavelength range of 420 to 720 nm with identical exposure times. On the basis of the extracted images, a spectral library required for multispectral unmixing was established using the InForm image analysis software (PerkinElmer) and reconstructed images of sections with autofluorescence removed were obtained. The antibodies are listed in table S10.

Immunofluorescent labeling

Tumor tissues were dissected, embedded in OCT compound (4583, Sakura), and frozen immediately at -80°C . Serial 8- μm sections were collected from the cryostat. After fixation in 4% paraformaldehyde for 20 min and blocking in 10% donkey serum in 1 \times PBS at room temperature for 1 hour, the sections were permeabilized in PBS containing 0.1% Triton X-100. The sections were then incubated overnight at 4°C with the following primary antibodies: anti-E-cad (1:200; 610181, BD Biosciences), anti-WNT5A (1:50; ab235966, Abcam), anti-NRG1 (1:100; ab191139, Abcam), anti- α -SMA (1:200; A5228, Sigma-Aldrich), anti-ERBB3 (1:100; 12708, Cell Signaling Technology), anti-PDGFR α (1:200; 3174, Cell Signaling Technology), and anti-FZD6 (1:100; NBP1-89702, Novus Biologicals). Tissue sections were washed three times with PBS and incubated with the appropriate Alexa Fluor-conjugated secondary antibodies [donkey anti-rabbit Alexa Fluor 488 (1:500; A-21206, Invitrogen), donkey anti-mouse Alexa Fluor 568 (1:500; A10037, Invitrogen), donkey anti-mouse Alexa Fluor 568 (1:500; A10042, Invitrogen), or donkey anti-mouse Alexa Fluor 488 (1:500; A-21202, Invitrogen)] for 1 hour at room temperature, after which they were incubated with DAPI for 15 min. Images were obtained on a TissueGnostics imaging system (TissueFAXS Plus, Austria).

Immunohistochemistry

Tumor tissues were embedded in paraffin, sectioned, deparaffinized in xylene, and rehydrated in graded alcohol to distilled water. Slides were heated for antigen retrieval as follows: sodium citrate buffer (pH 6) for WNT5A, NRG1, PDGFR α , α -SMA, and FZD6 and tris-EDTA buffer (pH 9) for ERBB3. The endogenous peroxidase activity was blocked by incubation with 3% hydrogen peroxide for 10 min. The sections were incubated overnight at 4°C with the following primary antibodies: anti-WNT5A (1:100; ab235966,

Abcam), anti-NRG1 (1:100; ab191139, Abcam), anti- α -SMA (1:5000; A5228, Sigma-Aldrich), anti-ERBB3 (1:50; 12708, Cell Signaling Technology), anti-PDGFR α (1:100; 3174, Cell Signaling Technology), and anti-FZD6 (1:100; NBP1-89702, Novus Biologicals). Tissue sections were washed three times with PBS, incubated with the appropriate horseradish peroxidase-conjugated secondary antibodies for 1 hour at room temperature, and detected using the 3,3'-diaminobenzidine (DAB) Substrate Kit (SK-4100, Vector). The sections were dehydrated and mounted with mounting medium (10004160, Sinaharm).

Western blotting

Whole-cell protein extracts were acquired by placing cells in lysis buffer with protease cocktail inhibitor ampt;Iota; (Calbiochem, San Diego, CA, USA). The resulting extracts were then resolved by SDS-polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes. The Western blots were blocked in tris-buffered saline tween 20 (TBST) containing 5% skimmed milk and incubated with primary antibodies overnight at 4°C. The membranes were then washed with TBST and incubated with horseradish peroxidase-conjugated secondary antibodies (Beijing Zhongshan Biotechnology, China) for 1 hour at room temperature. The antigen-antibody reaction was visualized by the enhanced chemiluminescence assay (Thermo Fisher Scientific) after three additional washes with TBST. The antibodies used are listed in table S10.

Isolation and expansion of mCAFs from cervical cancer tissues

CESC samples from patients were enzymatically dissociated as in our previous study (53) and grown in culture dishes in a modification of minimum essential medium containing basic fibroblast growth factor (bFGF) (1 ng/ml; R&D Systems, Minneapolis, MN), 10% selected FBS, streptomycin (100 U/ml), and penicillin (100 μ g/ml; Invitrogen, NY, USA). After 48 hours, nonadherent cells and tissue debris were removed, and the cells were washed twice with PBS. Adherent fibroblasts were further incubated for 6 to 10 days until 80 to 90% confluence was achieved. The mixed fibroblasts then underwent FACS using PDGFR α -allophycocyanin (1:50; catalog no. 323512, BioLegend) to enrich for mCAF subpopulations. The mCAFs were further expanded in the above medium and used in this study between passages 5 and 10.

Cell culture

The human cervical cancer cell lines CaSki, SiHa, and HeLa were purchased from Beijing Beina Chuanglian Biotechnology Institute, and 293T cells were obtained from the American Type Culture Collection. CaSki cells were cultured in RPMI 1640 medium supplemented with 10% FBS (Invitrogen). SiHa and HeLa cells were grown in high-glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS. All cell lines were cultured in a humidified incubator containing 5% CO₂ at 37°C.

Migration assays

For the migration assays, approximately 5×10^5 cervical cancer cells were seeded into the upper inserts of six-well Transwell plates (Corning, USA). In some experiments, the cervical cancer cells were incubated with the Wnt5a inhibitor Box5 (681673, Sigma-Aldrich) at a concentration of 100 μ M for 48 hours, before

seeding into Transwells. The medium in the upper inserts was free of FBS, while the bottom medium was supplemented with 10% FBS. After 48 hours, the cell inserts were fixed and stained with hematoxylin and eosin (H&E). Representative fields were photographed, and the number of migrated cells per field was counted.

Migration assays in the mCAF/cervical cancer cell Transwell coculture system

To study the effect of mCAFs on cervical cancer cell migration assays via the Transwell coculture system, confluent mCAFs were treated with mitomycin C for 1.5 hours, and 5×10^5 mCAFs were placed into the bottom of six-well Transwell plates. After 24 hours, approximately 5×10^5 cervical cancer cells were seeded into the upper inserts of six-well Transwell plates. The medium in the inserts was free of FBS, while the bottom medium was supplemented with 10% FBS. After 48 hours, the cell inserts were fixed and stained with H&E.

Tumor sphere formation

Two thousand cervical cancer cells were seeded into six-well ultralow attachment plates (Corning, USA) and cultured in sphere formation medium: DMEM/F-12 supplemented with B27 (1:50; Life Technologies, USA), epidermal growth factor (20 ng/ml), bFGF (10 ng/ml), penicillin (100 U/ml), N2 (1:100; Life Technologies), and streptomycin (100 ng/ml). Cervical cancer cells were treated with 3 μ M ErbB3 inhibitor (TX1-85-1; HY-100848, MCE) or dimethyl sulfoxide as the control. Tumor spheres larger than 100 μ m were counted under a stereomicroscope after 7 to 9 days. Small molecular inhibitors and growth factors are shown in the table S11.

Tumor sphere formation in the mCAF/cervical cancer cell Transwell coculture system

To study the effect of mCAFs on sphere formation by cervical cancer cells in the Transwell coculture system, confluent mCAFs were treated with mitomycin C for 1.5 hours, and 5×10^5 mCAFs were placed in the upper inserts. After culturing for 24 hours, the inserts were washed with PBS, and the cells were transferred into the fresh wells of ultralow attachment plates supplemented with tumor sphere culture medium. Then, 2000 cervical cancer cells were seeded into the bottom chamber of a six-well ultralow attachment plate. The tumor spheres were counted under a stereomicroscope after 7 to 9 days. In the drug resistance assays, CBP (10 μ g/ml) and DDP (1 μ g/ml) were added into the bottom well of the mCAF/cervical cancer cell Transwell coculture system.

Short hairpin RNA and lentivirus package

To generate the lentiviral plasmid for stable RNA interference, we designed short hairpin RNA (shRNA) using online software (<http://rnaidesigner.lifetechnologies.com/rnaiexpress/design.do>). The effective shRNAs had the following sequences: shNRG1 (GGCTGATTCTGGAGAGTATAT), shERBB3 (TGCAGCTTGTCACTCAATATT), shWNT5A (CAAAGAATGCCAGTATCAATT), and shFZD6 (ACCCAGAGAGACCAATTATAT). A nontargeting scrambled silencing RNA was used as the control. Lentivirus packaging was performed in 293T cells after cotransfection of packaging plasmids (pRRE, pCMV-VSVG, and pRSV-REV, Addgene) using Lipofectamine 2000 (Thermo Fisher Scientific).

Quantitative polymerase chain reaction

Total RNAs were purified using TRIzol (Invitrogen) and reverse-transcribed into cDNAs with the miScript Reverse Transcription Kit (Qiagen Sciences Inc., USA). The miScript SYBR-Green PCR Master Mix on the ABI Prism 7900 system was then used to assay gene expression. The level of mRNA expression of each gene was normalized to 18S. qPCR was performed according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments guidelines. The mRNA primer sequences are listed in table S12.

Supplementary Materials

This PDF file includes:

Figs. S1 to S11

Other Supplementary Material for this

manuscript includes the following:

Tables S1 to S12

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- R. L. Siegel, K. D. Miller, H. E. Fuchs, A. Jemal, Cancer statistics, 2021. *CA Cancer J. Clin.* **71**, 7–33 (2021).
- M. Arbyn, E. Weiderpass, L. Bruni, S. de Sanjosé, M. Saraiya, J. Ferlay, F. Bray, Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis. *Lancet Glob. Health* **8**, e191–e203 (2020).
- W. Small Jr., M. A. Bacon, A. Bajaj, L. T. Chuang, B. J. Fisher, M. M. Harkenrider, A. Jhingran, H. C. Kitchener, L. R. Mileskhan, A. N. Viswanathan, D. K. Gaffney, Cervical cancer: A global health crisis. *Cancer* **123**, 2404–2412 (2017).
- P. A. Cohen, A. Jhingran, A. Oaknin, L. Denny, Cervical cancer. *Lancet* **393**, 169–182 (2019).
- K. Matsuo, H. Machida, R. S. Mandelbaum, I. Konishi, M. Mikami, Validation of the 2018 FIGO cervical cancer staging system. *Gynecol. Oncol.* **152**, 87–93 (2019).
- I. Kalliala, A. Athanasiou, A. A. Veroniki, G. Salanti, O. Efthimiou, N. Raftis, S. Bowden, M. Paraskeva, K. Aro, M. Arbyn, P. Bennett, P. Nieminen, E. Paraskevaidis, M. Kyrgiou, Incidence and mortality from cervical cancer and other malignancies after treatment of cervical intraepithelial neoplasia: A systematic review and meta-analysis of the literature. *Ann. Oncol.* **31**, 213–227 (2020).
- V. Giroux, A. K. Rustgi, Metaplasia: Tissue injury adaptation and a precursor to the dysplasia-cancer sequence. *Nat. Rev. Cancer* **17**, 594–604 (2017).
- A. C. Berger, A. Korkut, R. S. Kanchi, A. M. Hegde, W. Lenoir, W. Liu, Y. Liu, H. Fan, H. Shen, V. Ravikumar, A. Rao, A. Schultz, X. Li, P. Sumazin, C. Williams, P. Mestdagh, P. H. Gunaratne, C. Yau, R. Bowlby, A. G. Robertson, D. G. Tiezzi, C. Wang, A. D. Cherniack, A. K. Godwin, N. M. Kuderer, J. S. Rader, R. E. Zuna, A. K. Sood, A. J. Lazar, A. I. Ojesina, C. Adebamowo, S. N. Adebamowo, K. A. Baggerly, T. W. Chen, H. S. Chiu, S. Lefever, L. Liu, K. MacKenzie, S. Orsulic, J. Roszik, C. S. Shelley, Q. Song, C. P. Vellano, N. Wentzensen, J. N. Weinstein, G. B. Mills, D. A. Levine, R. Akbani, A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer Cell* **33**, 690–705.e9 (2018).
- The Cancer Genome Atlas Research Network, Integrated genomic and molecular characterization of cervical cancer. *Nature* **543**, 378–384 (2017).
- L. Zammataro, S. Lopez, S. Bellone, F. Pettinella, E. Bonazzoli, E. Perrone, S. Zhao, G. Menderes, G. Altwerger, C. Han, B. Zeybek, A. Bianchi, A. Manzano, P. Manara, E. Cocco, N. Buza, P. Hui, S. Wong, A. Ravaggi, E. Bignotti, C. Romani, P. Todeschini, L. Zanotti, F. Odicino, S. Pecorelli, C. Donzelli, L. Ardighieri, R. Angioli, F. Raspagliesi, G. Scambia, J. Choi, W. Dong, K. Bilguvar, L. B. Alexandrov, D. A. Silasi, G. S. Huang, E. Ratner, M. Azodi, P. E. Schwartz, V. Pirazzoli, A. L. Stiegler, T. J. Boggon, R. P. Lifton, J. Schlessinger, A. D. Santin, Whole-exome sequencing of cervical carcinomas identifies activating ERBB2 and PIK3CA mutations as targets for combination therapy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22730–22736 (2019).
- J. E. Visvader, Cells of origin in cancer. *Nature* **469**, 314–322 (2011).
- D. Hanahan, R. A. Weinberg, Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
- S. H. Gohil, J. B. Iorgulescu, D. A. Braun, D. B. Keskin, K. J. Livak, Applying high-dimensional single-cell technologies to the analysis of cancer immunotherapy. *Nat. Rev. Clin. Oncol.* **18**, 244–256 (2021).
- S. Z. Wu, A. Swarbrick, Single-cell advances in stromal-leukocyte interactions in cancer. *Immunol. Rev.* **302**, 286–298 (2021).
- C. Li, L. Guo, S. Li, K. Hua, Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and transcriptional activities of ECs in CC. *Mol. Ther. Nucleic Acids* **24**, 682–694 (2021).
- C. Chumduri, R. K. Gurumurthy, H. Berger, O. Dietrich, N. Kumar, S. Koster, V. Brinkmann, K. Hoffmann, M. Drabkina, P. Arampatzis, D. Son, U. Klemm, H. J. Mollenkopf, H. Herbst, M. Mangler, J. Vogel, A. E. Saliba, T. F. Meyer, Opposing Wnt signals regulate cervical squamocolumnar homeostasis and emergence of metaplasia. *Nat. Cell Biol.* **23**, 184–197 (2021).
- A. N. H. Hopman, J. M. Mosh, K. J. Hoogduin, M. Ummelen, M. E. R. Henfling, M. van Engeland, K. A. D. Wouters, H. Stoop, L. H. J. Looijenga, F. C. S. Ramaekers, SOX17 expression and its down-regulation by promoter methylation in cervical adenocarcinoma in situ and adenocarcinoma. *Histopathology* **76**, 383–393 (2020).
- K. Löhmussaar, R. Oka, J. Espejo Valle-Inclan, M. H. H. Smits, H. Wardak, J. Korving, H. Begthel, N. Proost, M. van de Ven, O. W. Kranenburg, T. G. N. Jonges, R. P. Zweemer, S. Veersema, R. van Bortel, H. Clevers, Patient-derived organoids model cervical tissue dynamics and viral oncogenesis in cervical cancer. *Cell Stem Cell* **28**, 1380–1396.e6 (2021).
- A. Wiede, M. Hinz, E. Canzler, K. Franke, C. Quednow, W. Hoffmann, Synthesis and localization of the mucin-associated TFF-peptides in the human uterus. *Cell Tissue Res.* **303**, 109–115 (2001).
- L. Riethdorf, J. T. O'Connell, S. Riethdorf, A. Cviko, C. P. Crum, Differential expression of MUC2 and MUC5AC in benign and malignant glandular lesions of the cervix uteri. *Virchows Arch.* **437**, 365–371 (2000).
- N. Wentzensen, M. A. Clarke, R. Bremer, N. Poitras, D. Tokugawa, P. E. Goldhoff, P. E. Castle, M. Schiffman, J. D. Kingery, K. K. Grewal, A. Locke, W. Kinney, T. S. Lorey, Clinical evaluation of human papillomavirus screening with p16/Ki-67 dual stain triage in a large organized cervical cancer screening program. *JAMA Intern. Med.* **179**, 881–888 (2019).
- A. L. Peres, E. S. K. M. Paz, R. F. de Araújo, J. L. de Lima Filho, M. R. de Melo Júnior, D. B. Martins, N. T. de Pontes Filho, Immunocytochemical study of TOP2A and Ki-67 in cervical smears from women under routine gynecological care. *J. Biomed. Sci.* **23**, 42 (2016).
- L. F. Escobar-Hoyos, J. Yang, J. Zhu, J. A. Cavallo, H. Zhai, S. Burke, A. Koller, E. I. Chen, K. R. Shroyer, Keratin 17 in premalignant and malignant squamous lesions of the cervix: Proteomic discovery and immunohistochemical validation as a diagnostic and prognostic biomarker. *Mod. Pathol.* **27**, 621–630 (2014).
- M. Hornburg, M. Desbois, S. Lu, Y. Guan, A. A. Lo, S. Kaufman, A. Elrod, A. Lotstein, T. M. DesRochers, J. L. Munoz-Rodriguez, X. Wang, J. Giltane, O. Mayba, S. J. Turley, R. Bourgon, A. Daemen, Y. Wang, Single-cell dissection of cellular components and interactions shaping the tumor immune phenotypes in ovarian cancer. *Cancer Cell* **39**, 928–944.e6 (2021).
- J. Huang, P. Chen, K. Liu, J. Liu, B. Zhou, R. Wu, Q. Peng, Z. X. Liu, C. Li, G. Kroemer, M. Lotze, H. Zeh, R. Kang, D. Tang, CDK1/2/5 inhibition overcomes IFN γ -mediated adaptive immune resistance in pancreatic cancer. *Gut* **70**, 890–899 (2021).
- M. D. Sharma, R. Pacholczyk, H. Shi, Z. J. Berrong, Y. Zakharia, A. Greco, C. S. Chang, S. Eathiraj, E. Kennedy, T. Cash, R. J. Bollag, R. Kolhe, R. Sadek, T. L. McGaha, P. Rodriguez, J. Mandula, B. R. Blazar, T. S. Johnson, D. H. Munn, Inhibition of the BTK-IDO-mTOR axis promotes differentiation of monocyte-lineage dendritic cells and enhances anti-tumor T cell immunity. *Immunity* **54**, 2354–2371.e8 (2021).
- E. Proietti, S. Rossini, U. Grohmann, G. Mondanelli, Polyamines and kynurenines at the intersection of immune modulation. *Trends Immunol.* **41**, 1037–1050 (2020).
- O. Bartok, A. Pataskar, R. Nagel, M. Laos, E. Goldfarb, D. Hayoun, R. Levy, P. R. Korner, I. Z. M. Kreuger, J. Champagne, E. A. Zaal, O. B. Bleijerveld, X. Huang, J. Kenski, J. Wargo, A. Brandis, Y. Levin, O. Mizrahi, M. Alon, S. Lebon, W. Yang, M. M. Nielsen, N. Stern-Ginossar, M. Altelaar, C. R. Berkers, T. Geiger, D. S. Peeper, J. Olweus, Y. Samuels, R. Agami, Anti-tumour immunity induces aberrant peptide presentation in melanoma. *Nature* **590**, 332–337 (2021).
- L. Zhang, Z. Li, K. M. Skrzypczynska, Q. Fang, W. Zhang, S. A. O'Brien, Y. He, L. Wang, Q. Zhang, A. Kim, R. Gao, J. Orf, T. Wang, D. Sawant, J. Kang, D. Bhatt, D. Lu, C. M. Li, A. S. Rapaport, K. Perez, Y. Ye, S. Wang, X. Hu, X. Ren, W. Ouyang, Z. Shen, J. G. Egen, Z. Zhang, X. Yu, Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* **181**, 442–459.e29 (2020).
- D. A. Braun, K. Street, K. P. Burke, D. L. Cookmeyer, T. Denize, C. B. Pedersen, S. H. Gohil, N. Schindler, L. Pomerance, L. Hirsch, Z. Bakouny, Y. Hou, J. Forman, T. Huang, S. Li, A. Cui, D. B. Keskin, J. Steinharter, G. Bouchard, M. Sun, E. M. Pimenta, W. Xu, K. M. Mahoney, B. A. McGregor, M. S. Hirsch, S. L. Chang, K. J. Livak, D. F. McDermott, S. A. Shukla, L. R. Olsen, S. Signoretti, A. H. Sharpe, R. A. Irizarry, T. K. Choueiri, C. J. Wu, Progressive immune

- dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell* **39**, 632–648.e8 (2021).
31. S. Cheng, Z. Li, R. Gao, B. Xing, Y. Gao, Y. Yang, S. Qin, L. Zhang, H. Ouyang, P. Du, L. Jiang, B. Zhang, Y. Yang, X. Wang, X. Ren, J. X. Bei, X. Hu, Z. Bu, J. Ji, Z. Zhang, A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792–809.e23 (2021).
 32. E. Azizi, A. J. Carr, G. Plitas, A. E. Cornish, C. Konopacki, S. Prabhakaran, J. Nainys, K. Wu, V. Kiseliovas, M. Setty, K. Choi, R. M. Fromme, P. Dao, P. T. McKenney, R. C. Wasti, K. Kadaveru, L. Mazutis, A. Y. Rudensky, D. Pe'er, Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308.e36 (2018).
 33. R. J. P. Bonnal, G. Rossetti, E. Lugli, M. De Simone, P. Gruarin, J. Brummelman, L. Drufuca, M. Passaro, B. Bason, F. Gervasoni, G. D. Chiara, C. D'Oria, M. Martinovic, S. Curti, V. Ranzani, C. Cordigliari, G. Alvisi, E. M. C. Mazza, S. Oliveto, Y. Silvestri, E. Carelli, S. Mazzara, R. Bosotti, M. L. Sarnicola, C. Godano, V. Bevilacqua, M. Lorenzo, S. Siena, E. Bonoldi, A. Sartore-Bianchi, A. Amatu, G. Veronesi, P. Novellis, M. Alloisio, A. Giani, N. Zucchini, E. Opocher, A. P. Ceretti, N. Mariani, S. Biffo, D. Prati, A. Bardelli, J. Geginat, A. Lanzavecchia, S. Abrignani, M. Pagani, Clonally expanded EOMES⁺ Tr1-like cells in primary and metastatic tumors are associated with disease progression. *Nat. Immunol.* **22**, 735–745 (2021).
 34. N. D. Mathewson, O. Ashenberg, I. Tirosh, S. Gritsch, E. M. Perez, S. Marx, L. Jerby-Arnon, R. Chanoch-Myers, T. Hara, A. R. Richman, Y. Ito, J. Pyrdol, M. Friedrich, K. Schumann, M. J. Poitras, P. C. Gokhale, L. N. Gonzalez Castro, M. E. Shore, C. M. Hebert, B. Shaw, H. L. Cahill, M. Drummond, W. Zhang, O. Olawoyin, H. Wakimoto, O. Rozenblatt-Rosen, P. K. Brastianos, X. S. Liu, P. S. Jones, D. P. Cahill, M. P. Frosch, D. N. Louis, G. J. Freeman, K. L. Ligon, A. Marson, E. A. Chiocci, D. A. Reardon, A. Regev, M. L. Suvà, K. W. Wucherpfennig, Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* **184**, 1281–1298.e26 (2021).
 35. P. Gueguen, C. Metoikidou, T. Dupic, M. Lawand, C. Goudot, S. Baulande, S. Lameiras, O. Lantz, N. Girard, A. Seguin-Givélet, M. Lefevre, T. Mora, A. M. Walczak, J. J. Waterfall, S. Amigorena, Contribution of resident and circulating precursors to tumor-infiltrating CD8⁺ T cell populations in lung cancer. *Sci. Immunol.* **6**, eabd5778 (2021).
 36. B. Walch-Rückheim, R. Stroder, L. Theobald, J. Pahne-Zeppenfeld, S. Hegde, Y. J. Kim, R. M. Bohle, I. Juhasz-Boss, E. F. Solomayer, S. Smola, Cervical cancer-instructed stromal fibroblasts enhance IL23 expression in dendritic cells to support expansion of Th17 cells. *Cancer Res.* **79**, 1573–1586 (2019).
 37. B. Walch-Rückheim, R. Mavrova, M. Henning, B. Vicinus, Y. J. Kim, R. M. Bohle, I. Juhasz-Böss, E. F. Solomayer, S. Smola, Stromal fibroblasts induce CCL20 through IL6/C/EBP β to support the recruitment of Th17 cells during cervical cancer progression. *Cancer Res.* **75**, 5248–5259 (2015).
 38. J. Goveia, K. Rohlenova, F. Taverna, L. Treps, L. C. Conradi, A. Pircher, V. Geldhof, L. de Rooij, J. Kalucka, L. Sokol, M. Garcia-Caballero, Y. Zheng, J. Qian, L. A. Teuwen, S. Khan, B. Boeckx, E. Wauters, H. Decaluwe, P. De Leyn, J. Vansteenkiste, B. Weynand, X. Sagaert, E. Verbeke, A. Walthuis, B. Topal, W. Everaert, H. Bohnenberger, A. Emmert, D. Panovska, F. De Smet, F. J. T. Staal, R. J. McLaughlin, F. Impens, V. Lagani, S. Vincier, M. Mazzone, L. Schoonjans, M. Dewerchin, G. Eelen, T. K. Karakach, H. Yang, J. Wang, L. Bolund, L. Lin, B. Thienpont, X. Li, D. Lambrechts, Y. Luo, P. Carmeliet, An integrated gene expression landscape profiling approach to identify lung tumor endothelial cell heterogeneity and angiogenic candidates. *Cancer Cell* **37**, 21–36.e13 (2020).
 39. E. Sahai, I. Astsaturov, E. Cukierman, D. G. DeNardo, M. Egeblad, R. M. Evans, D. Fearon, F. R. Greten, S. R. Hingorani, T. Hunter, R. O. Hynes, R. K. Jain, T. Janowitz, C. Jorgensen, A. C. Kimmelman, M. G. Kolonin, R. G. Maki, R. S. Powers, E. Puré, D. C. Ramirez, R. Scherz-Shouval, M. H. Sherman, S. Stewart, T. D. Tlsty, D. A. Tuveson, F. M. Watt, V. Weaver, A. T. Weeraratna, Z. Werb, A framework for advancing our understanding of cancer-associated fibroblasts. *Nat. Rev. Cancer* **20**, 174–186 (2020).
 40. R. Vento-Tormo, M. Efremova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer, J.-E. Park, E. Stephenson, K. Polański, A. Goncalves, L. Gardner, S. Holmqvist, J. Henriksson, A. Zou, A. M. Sharkey, B. Millar, B. Innes, L. Wood, A. Wilbrey-Clark, R. P. Payne, M. A. Ivarsson, S. Lisgo, A. Filby, D. H. Rowitch, J. N. Bulmer, G. J. Wright, M. J. T. Stubbington, M. Haniffa, A. Moffett, S. A. Teichmann, Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* **563**, 347–353 (2018).
 41. Z. Zhang, W. R. Karthaus, Y. S. Lee, V. R. Gao, C. Wu, J. W. Russo, M. Liu, J. M. Mota, W. Abida, E. Linton, E. Lee, S. D. Barnes, H. A. Chen, N. Mao, J. Wongvipat, D. Choi, X. Chen, H. Zhao, K. Manova-Todorova, E. de Stanchina, M. E. Taplin, S. P. Balk, D. E. Rathkopf, A. Gopalan, B. S. Carver, P. Mu, X. Jiang, P. A. Watson, C. L. Sawyers, Tumor microenvironment-derived NRG1 promotes antiandrogen resistance in prostate cancer. *Cancer Cell* **38**, 279–296.e9 (2020).
 42. Y. Kieffer, H. R. Hocine, G. Gentric, F. Pelon, C. Bernard, B. Bourachot, S. Lameiras, L. Albergante, C. Bonneau, A. Guyard, K. Tarte, A. Zinovyev, S. Baulande, G. Zalcman, A. Vincent-Salomon, F. Mechta-Grigoriou, Single-cell analysis reveals fibroblast clusters linked to immunotherapy resistance in cancer. *Cancer Discov.* **10**, 1330–1351 (2020).
 43. Y. Sun, L. Wu, Y. Zhong, K. Zhou, Y. Hou, Z. Wang, Z. Zhang, J. Xie, C. Wang, D. Chen, Y. Huang, X. Wei, Y. Shi, Z. Zhao, Y. Li, Z. Guo, Q. Yu, L. Xu, G. Volpe, S. Qiu, J. Zhou, C. Ward, H. Sun, Y. Yin, X. Xu, X. Wang, M. A. Esteban, H. Yang, J. Wang, M. Dean, Y. Zhang, S. Liu, X. Yang, J. Fan, Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* **184**, 404–421.e16 (2020).
 44. M. Zhang, S. Hu, M. Min, Y. Ni, Z. Lu, X. Sun, J. Wu, B. Liu, X. Ying, Y. Liu, Dissecting transcriptional heterogeneity in primary gastric adenocarcinoma by single cell RNA sequencing. *Gut* **70**, 464–475 (2021).
 45. S. Chen, G. Zhu, Y. Yang, F. Wang, Y. T. Xiao, N. Zhang, X. Bian, Y. Zhu, Y. Yu, F. Liu, K. Dong, J. Mariscal, Y. Liu, F. Soares, H. Loo Yau, B. Zhang, W. Chen, C. Wang, D. Chen, Q. Guo, Z. Yi, M. Liu, M. Fraser, D. D. De Carvalho, P. C. Boutros, D. Di Vizio, Z. Jiang, T. van der Kwast, A. Berlin, S. Wu, J. Wang, H. H. He, S. Ren, Single-cell analysis reveals transcriptomic remodelling in distinct cell types that contribute to human prostate cancer progression. *Nat. Cell Biol.* **23**, 87–98 (2021).
 46. R. Wang, M. Dang, K. Harada, G. Han, F. Wang, M. Pool Pizzi, M. Zhao, G. Tatlonghari, S. Zhang, D. Hao, Y. Lu, S. Zhao, B. D. Badgwell, M. Blum Murphy, N. Shanbhag, J. S. Estrella, S. Roy-Chowdhuri, A. A. F. Abdelhakeem, Y. Wang, G. Peng, S. Hanash, G. A. Calin, X. Song, Y. Chu, J. Zhang, M. Li, K. Chen, A. J. Lazar, A. Futreal, S. Song, J. A. Ajani, L. Wang, Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat. Med.* **27**, 141–151 (2021).
 47. T. D. Wu, S. Madireddi, P. E. de Almeida, R. Banchereau, Y. J. Chen, A. S. Chitre, E. Y. Chiang, H. Iftikhar, W. E. O'Gorman, A. Au-Yeung, C. Takahashi, L. D. Goldstein, C. Poon, S. Keerthivasan, D. E. de Almeida Nagata, X. Du, H. M. Lee, K. L. Banta, S. Mariathasan, M. Das Thakur, M. A. Huseini, M. Ballinger, I. Estay, P. Caplazi, Z. Modrusan, L. Delamarre, I. Mellman, R. Bourgon, J. L. Grogan, Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature* **579**, 274–278 (2020).
 48. C. Liu, H. Yu, R. Huang, T. Lei, X. Li, M. Liu, Q. Huang, Q. Du, L. Xing, J. Yu, Radioimmunotherapy-induced intratumoral changes in cervical squamous cell carcinoma at single-cell resolution. *Cancer Commun (Lond)* **42**, 1407–1411 (2022).
 49. X. Li, M. Zhang, T. Lei, W. Zou, R. Huang, F. Wang, Q. Huang, C. Liu, Single-cell RNA-sequencing dissects cellular heterogeneity and identifies two tumor-suppressing immune cell subclusters in HPV-related cervical adenosquamous carcinoma. *J. Med. Virol.* **94**, 6047–6059 (2022).
 50. S. A. Hong, S. H. Yoo, J. Choi, S. J. Robboy, K. R. Kim, A review and update on papillary immature metaplasia of the uterine cervix: A distinct subset of low-grade squamous intraepithelial lesion, proposing a possible cell of origin. *Arch. Pathol. Lab. Med.* **142**, 973–981 (2018).
 51. J. Peng, B. F. Sun, C. Y. Chen, J. Y. Zhou, Y. S. Chen, H. Chen, L. Liu, D. Huang, J. Jiang, G. S. Cui, Y. Yang, W. Wang, D. Guo, M. Dai, J. Guo, T. Zhang, Q. Liao, Y. Liu, Y. L. Zhao, D. L. Han, Y. Zhao, Y. G. Yang, W. Wu, Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.* **29**, 725–738 (2019).
 52. S. Jin, R. Li, M. Y. Chen, C. Yu, L. Q. Tang, Y. M. Liu, J. P. Li, Y. N. Liu, Y. L. Luo, Y. Zhao, Y. Zhang, T. L. Xia, S. X. Liu, Q. Liu, G. N. Wang, R. You, J. Y. Peng, J. Li, F. Han, J. Wang, Q. Y. Chen, L. Zhang, H. Q. Mai, B. E. Gewurz, B. Zhao, L. S. Young, Q. Zhong, F. Bai, M. S. Zeng, Single-cell transcriptomic analysis defines the interplay between tumor cells, viral infection, and the microenvironment in nasopharyngeal carcinoma. *Cell Res.* **30**, 950–965 (2020).
 53. M. Zhang, H. Yang, L. Wan, Z. Wang, H. Wang, C. Ge, Y. Liu, Y. Hao, D. Zhang, G. Shi, Y. Gong, Y. Ni, C. Wang, Y. Zhang, J. Xi, S. Wang, L. Shi, L. Zhang, W. Yue, X. Pei, B. Liu, X. Yan, Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J. Hepatol.* **73**, 1118–1130 (2020).
 54. C. Capparelli, S. Rosenbaum, A. C. Berger, A. E. Aplin, Fibroblast-derived neuregulin 1 promotes compensatory ErbB3 receptor signaling in mutant BRAF melanoma. *J. Biol. Chem.* **290**, 24267–24277 (2015).
 55. W. F. Wei, X. J. Chen, L. J. Liang, L. Yu, X. G. Wu, C. F. Zhou, Z. C. Wang, L. S. Fan, Z. Hu, L. Liang, W. Wang, Periostin⁺ cancer-associated fibroblasts promote lymph node metastasis by impairing the lymphatic endothelial barriers in cervical squamous cell carcinoma. *Mol. Oncol.* **15**, 210–227 (2021).
 56. M. Nagura, N. Matsumura, T. Baba, R. Murakami, B. Kharma, J. Hamanishi, K. Yamaguchi, K. Abiko, M. Koshiyama, M. Mandai, T. Murata, S. K. Murphy, I. Konishi, Invasion of uterine cervical squamous cell carcinoma cells is facilitated by locoregional interaction with cancer-associated fibroblasts via activating transforming growth factor- β . *Gynecol. Oncol.* **136**, 104–111 (2015).
 57. L. Ma, M. O. Hernandez, Y. Zhao, M. Mehta, B. Tran, M. Kelly, Z. Rae, J. M. Hernandez, J. L. Davis, S. P. Martin, D. E. Kleiner, S. M. Hewitt, K. Ylaya, B. J. Wood, T. F. Greten, X. W. Wang, Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. *Cancer Cell* **36**, 418–430.e6 (2019).
 58. M. Zhang, L. Wu, Y. Deng, F. Peng, T. Wang, Y. Zhao, P. Chen, J. Liu, G. Cai, L. Wang, J. Wu, X. Chen, Single cell dissection of epithelial-immune cellular interplay in acute kidney injury microenvironment. *Front. Immunol.* **13**, 857025 (2022).
 59. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).

60. I. Tirosh, A. S. Venteicher, C. Hebert, L. E. Escalante, A. P. Patel, K. Yizhak, J. M. Fisher, C. Rodman, C. Mount, M. G. Filbin, C. Neftel, N. Desai, J. Nyman, B. Izar, C. C. Luo, J. M. Francis, A. A. Patel, M. L. Onozato, N. Riggi, K. J. Livak, D. Gennert, R. Satija, B. V. Nahed, W. T. Curry, R. L. Martuza, R. Mylvaganam, A. J. Iafrate, M. P. Frosch, T. R. Golub, M. N. Rivera, G. Getz, O. Rozenblatt-Rosen, D. P. Cahill, M. Monje, B. E. Bernstein, D. N. Louis, A. Regev, M. L. Suvà, Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309–313 (2016).

Acknowledgments: We thank all patients enrolled in this study. We thank R. Li for the help with artwork. We thank L. Robinson for the editing support. **Funding:** This study was supported by funds provided to J. Yu from the National Natural Science Foundation of China (82030082), the Academic Promotion Program of Shandong First Medical University (2019ZL002), and Radiation Oncology Innovate Unit, Chinese Academy of Medical Sciences (2019RU071). This study was also supported by grants provided to J. Yue from the National Natural Science Foundation of China (81871895), Young Taishan Scholars and Academic Promotion Program of Shandong First Medical University (2019RC003). The work was also supported by grants provided to C.L. from the Shandong Provincial Natural Science Foundation (ZR2021QH006) and the Beijing Bethune Charitable Foundation (flzh202118). Last, the work was also supported by grants from the National Natural Science Foundation of China (81772617 and 82173183), the Great Wall

Scholar Project (CIT&TCD20190311), and the R&D program of Beijing Municipal Education Commission (KZ202210005010) to X.Y. **Author contributions:** B.L., J. Yue, J. Yu, C.L., M.Z., and X.Y., conceived the project, designed the study, and interpreted the results. C.W., X.Z., W.Z., R.H., X.L., and B.S. contributed to sample and clinical data collection. Y.L. and Y.N. provided flow cytometry expertise. C.L., M.Z., X.Y., Y.N., Y.G., L.W., H.Y., and C.G. processed the samples for scRNA-seq and in vitro experiments and performed data analysis. C.L., M.Z., and X.Y. prepared figures and tables and wrote the manuscript. B.L., J. Yue, and J. Yu jointly supervised this work. All authors reviewed and approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The scRNA-seq data in the study have been uploaded to the Genome Sequence Archive (GSA) of the National Genomics Data Center (access number: PRJCA008573), including raw data accessible at <https://ngdc.cncb.ac.cn/gsa-human> and processed data available at <https://ngdc.cncb.ac.cn/omix>. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 12 July 2022

Accepted 23 December 2022

Published 27 January 2023

10.1126/sciadv.add8977