Key Genetic Determinants Driving Esophageal Squamous Cell Carcinoma Initiation and Immune Evasion

Kyung-Pil Ko,[1] Yuanjian Huang,[1] Shengzhe Zhang,[1] Gengyi Zou,[1] Bongjun Kim,[1] Jie Zhang,[1] Sohee Jun,[1] Cecilia Martin,[2] Karen J. Dunbar,[2] Gizem Efe,[2] Anil K. Rustgi,[2] Hiroshi Nakagawa,[2] Jae-Il Park[1,3,4,*]

[1]Department of Experimental Radiation Oncology, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[2]Division of Digestive and Liver Diseases, Department of Medicine, Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032, USA

[3]The University of Texas MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences, Houston, TX 77030, USA

[4]Program in Genetics and Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

*Correspondence: Jae-Il Park
E-mail: jaeil@mdanderson.org; Tel: 713-792-3659; Fax: 713-794-5369
Keywords: Esophageal squamous cell cancer, CCL2, CCR2, organoids, esophageal neoplasia, immune evasion, single-cell RNA-sequencing

Abstract

Despite recent progress in identifying aberrant genetic and epigenetic alterations in esophageal squamous cell carcinoma (ESCC), the mechanism of ESCC initiation remain unknown. Using genetically engineered esophageal organoids (EOs), we identified the key genetic determinants that drive ESCC tumorigenesis. A single-cell transcriptomic analysis uncovered that *Trp53*, *Cdkn2a*, and *Notch1* (PCN) triple knockout (KO) induces neoplastic features of ESCC by generating distinct cell lineage trajectories with multiple root cells and high cell plasticity. Although *Trp53* and *Notch1* (PN) double KO was sufficient to induce esophageal neoplasia and cellular heterogeneity, additional inactivation of *Cdkn2a* was indispensable for immune landscape remodeling for *in vivo* tumorigenesis. *PCN* KO generated immunosuppressive niche enriched with exhausted T cells and M2 macrophages via the CCL2-CCR2 axis in an autochthonous ESCC mouse model. Moreover, genetic or pharmacological blockade of the CCL2-CCR2 axis suppressed ESCC tumorigenesis. Comparative single-cell transcriptomic analyses classified ESCC patient tumors into three subgroups and identified a specific subset recapitulating PCN-type ESCC signatures, including the high expression of CCL2 and CD274/PD-L1. Our study unveils that loss of *TP53*, *CDKN2A*, and *NOTCH1* induces esophageal neoplasia and immune evasion for ESCC initiation and proposes the CCL2 blockade as a viable approach to target a subset of ESCC.

Introduction

Esophageal squamous cell carcinoma (ESCC) accounts for over 80% of all cases of esophageal cancer and has a poor prognosis because of a lack of symptoms in the early stages. The 5-year overall survival rate of patients with esophageal cancer ranges from 10% to 25%[1]. ESCC develops from squamous dysplasia as a typical invasive histologic precursor lesion[2], which makes it difficult to detect early and leads to late detection. Therefore, understanding the mechanisms of esophageal neoplasia and ESCC initiation is imperative to improve the detection, diagnosis, treatment, and prevention of ESCC. However, despite a recent genome-wide analysis of ESCC patients, the key genetic factors that drive ESCC neoplasia and initiation remain elusive.

In addition to the genetic and epigenetic alteration of tumor cells, the tumor microenvironment (TME) plays a pivotal role in tumorigenesis. The TME comprises immune and stromal cells surrounding tumor cells and is known to design the immune landscape to elicit immunosuppressive effects on tumors and various immunotherapeutic responses[3,4]. Although the TME's effects on tumor cells are biphasic, as immune cells generally recognize tumor cells or neoantigens to generate anti-tumor immune responses[5], the tumor-favorable TME induces tumor growth, invasion, and metastasis[6]. The TME is also regarded as an indispensable factor in ESCC development, as it provides drug resistance and immunosuppressive immune cell infiltration[7,8].

Immune cell profiling using single-cell transcriptomics has been used to elucidate the immune landscape of ESCC patients[9-11]. An immunosuppressive tumor niche is frequently observed in these patients, and immunotherapy is expected to be effective based on immune cell profiling results. Indeed, immune checkpoint inhibitors (ICIs) have been clinically tested in ESCC patients and have shown some promising results[12-16]. However, not all patients benefit from ICIs alone or combined with first-line therapy[7], possibly because of the lack of a clear classification of ESCC patients. The underlying genetic background or biomarkers in ICIs responders and non-responders have yet to be identified. Therefore, in this study, we sought to determine key genetic events driving ESCC initiation and generating tumor-favorable immune environments.

Results

Genetic ablation of *Trp53* and *Notch1* is sufficient to induce esophageal neoplasia

To elucidate the vital genetic events initiating ESCC, we analyzed genetic alterations in ESCC patients (n = 86). Recurrent loss-of-function mutations in the *TP53*, *CDKN2A*, *KMT2D*, *KMT2C*, *FAT1*, *FAT4*, *AJUBA*, *NOTCH1*, and *NOTCH3* genes were observed in ESCC patients, consistent with previous studies[17,18]. Genetic alterations of the nine genes frequently co-occurred, and the *TP53* and *CDKN2A* genes were the most commonly altered (> 70%) (Extended Data 1a). More than 50% of mutations in the *CDKN2A*, *KMT2D*, *NOTCH1*, and *FAT1* genes were frameshift or truncation mutations (Extended Data 1b). Also, ESCC patients displayed transcriptional downregulation of *KMT2C* and *FAT4* genes (Extended Data 1c). These results imply that genetic inactivation of the *TP53*, *CDKN2A*, *KMT2D*, *KMT2C*, *NOTCH1*, *NOTCH3*, *FAT1*, *FAT4*, and *AJUBA* genes might be associated with ESCC tumorigenesis.

To understand significance of the nine genes in ESCC initiation, we tested functional impact of their inactivation on transformation of murine esophageal organoids (EOs), as we recently established[19] (Fig. 1a and Supplementary Fig. 1a). Employing the CRISPR gene editing system, we genetically ablated the nine genes in 32 different combinations. Given that the *TP53* and *CDKN2A* genes were the most frequently mutated in patients, *Trp53* KO and *Cdkn2a* KO were chosen as the foundation for additional gene ablations, and double KO (dKO) or triple KO (tKO) EOs were established (Extended Data 1d). Gene KO was confirmed by immunoblot and genomic DNA sequencing (Supplementary Fig. 1b, c). Next, we assessed the size, morphology, and differentiation of 32 genetically engineered EOs (O1-O32) (Extended Data 1e, f).

Intriguingly, in combination with *Trp53* and *Cdkn2a* dKO, KO of *Notch1*, *Fat4*, or *Ajuba* led to a markedly increased EO size (Fig. 1b). However, only *Notch1* KO with *Trp53* and *Cdkn2a* dKO EOs displayed a loss of cell differentiation and hyperplastic growth (Fig. 1c). The increased number of proliferating cells was confirmed by Mki67 staining in dKO of *Trp53* and *Notch1* (PN) and tKO of *Trp53*, *Cdkn2a*, and *Notch1* (PCN) EOs (Fig. 1d, e and Extended Data 2a). Sox2 and Trp63, markers for ESCC stemness, were also highly expressed in PN and PCN compared to other EOs (Fig. 1d, f and Extended Data 2a). A bromodeoxyuridine (BrdU) incorporation assay confirmed the increased cell proliferation in PN and PCN EOs (Figs. 1g, h). *Notch1* KO suppressed organoid cell differentiation, as represented by decreased Krt13, a cell differentiation marker (Fig. 1f and Extended Data 1b), and this effect was validated using DAPT, a Notch signaling inhibitor (Extended Data Figs. 2c, d).

Since ESCC invades the stromal cell layers at an early stage[20], we analyzed the cell invasiveness of EOs by co-culturing them on the stromal cell layer (Extended Data 2e). While WT EO-derived epithelial cells formed distinct epithelial cell colonies, PN and PCN EO-derived cells were mixed with the stromal cell layer (Fig. 1i, j), recapitulating the invasive feature of ESCC. We further examined the tumorigenicity of each group of EOs in the 2D-culture system to assess cell-autonomous growth without the supporting materials and growth factors included in the organoid culture system. In 2D-culture, only PN and PCN EOs showed colony formation and exponential cell growth, while other EOs failed to grow. Of note, PN EOs showed relatively faster growth than PCN EOs (Figs. 1k, l and Extended Data 2f). PN and PCN cells also displayed slightly increased

cell migration compared to 2D-cultured PC cells (Extended Data 2g). Of note, PN and PCN EOs were not related to EAC (Extended Data 2h). These results suggest that *Trp53* and *Notch1* dKO (PN) and *Trp53*, *Cdkn2a, Notch1* tKO (PCN) are sufficient to induce neoplastic transformation exhibiting the characteristics of early-stage ESCC.

Dysregulated cell plasticity and lineage trajectories of neoplastic EOs

We next sought to understand the mechanism of ESCC initiation by single-cell transcriptomic analysis. We performed single-cell RNA-sequencing (scRNA-seq) of WT, PC, PN, and PCN EOs. Each library was sequenced with the multiplexing workflow, and the results were integrated using sctransform (R) or Harmony (Python) (Fig. 2a-c and Extended Data 3a). Thirty-five cell clusters were annotated with multiple markers that were specifically expressed in proliferative (Mki67, Top2a, and Stmn1 markers), early suprabasal (Ptma, Itgb4, and Krt5 markers), intermediate suprabasal (Itga6 marker), and stratified (Krt13, Krt4, and Sprr3 markers) cell types[19,21] (Extended Data 3b-d). The intermediate suprabasal cell cluster showed a moderate expression level of early suprabasal markers (Itgb4 and Krt5) and stratified markers. A cell proportion analysis showed that the proliferative cells were significantly enriched in the PC, PN, and PCN organoids compared to the WT EOs enriched with differentiated (stratified and suprabasal) cells (Fig. 2d and Extended Data 3e). These results were consistent with the hyperproliferative phenotypes of PN or PCN EOs (Fig. 1).

Having observed distinct cell hyperproliferation and de-differentiation by PN and PCN, we examined the PN or PCN-specific cell lineages through RNA velocity–based cell trajectory inferences with re-clustered UMAP projections (Fig. 2e). The cell lineages, indicated by a line of arrows, show that PN and PCN have multiple root cell clusters, while WT and PC have a single root cell cluster (Fig. 2f). Cell lineage trajectories were also assessed by the latent time and PAGA analysis (Fig. 2g-i). The PAGA analysis visualized that PN and PCN harbored three root cell clusters (PN3, PN4, and PN8 and PCN3, PCN4, and PCN7, respectively), while WT showed one root cell cluster (WT4) (Fig. 2h, i), consistent with the RNA velocity results (Fig. 2f). Notably, PC displayed two root cell clusters (PC4 and PC5) with relatively more straightforward lineage trajectories than those of PN and PCN (Fig. 2h, i). These results suggest that the combined ablation of *Trp53* and *Notch1* generates distinct cell plasticity and higher cellular heterogeneity, as shown in malignant tumors[22].

Human ESCC features in *Trp53* / *Notch1* dKO and *Trp53* / *Cdkn2a* / *Notch1* tKO EOs

We assessed the pathological relevance of genetically engineered EOs to human ESCC by comparing the transcriptional signatures of EOs with those of ESCC patients. Compared to WT and PC EOs, ESCC cancer stem cell markers, including Trp63, Krt5, and Tfrc, were intensively enriched in the proliferating cells of PN and PCN EOs (Fig. 3a). Sox2 and Gnl3 were highly expressed in the proliferating cells of PN and PCN EOs, respectively, while the other genes did not show the specificity to the PN or PCN cell types. We also compared the transcriptional features of our EOs with those of ESCC patients from the TCGA database using the Scissor package. The gene expression feature comparison analysis showed the similarity of each EO to ESCC patients: WT (35.4%), PC (52.2%), PN (50.5%), and PCN (52.2%) (Fig. 3b, c). Additionally, PCN (21.9%)

showed the highest similarity with the poor survival–associated ESCC patients compared to the other organoids (Fig. 3d, e). A gene set enrichment analysis further confirmed that the PCN highly expressed gene set was enriched in ESCC-specific gene features, while the PN highly expressed gene set enrichment was not statistically significant (Fig. 3f). To increase the depth of the sequencing read, we also performed bulk RNA-seq of WT, PN, and PCN EOs and reanalyzed the gene expression of each dataset (Fig. 3g). Using the differentially expressed genes (DEGs) of PN to WT or PCN to WT, we ran the Enrichr analysis with PN or PCN highly expressed genes. Both PN and PCN highly expressed genes were shown to be enriched with features of the cell cycle, DNA replication, and mitotic cells in the Bioplanet and REACTOME databases (Fig. 3h). The "pathways in cancer" were also associated with the PCN highly expressed genes in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis (Fig. 3i). These results suggest that the transcriptomic patterns of human ESCC are most similar to the transcriptomic features of *Trp53*, *Cdkn2a*, *Notch1* tKO EOs.

Trp53 / *Cdkn2a* / *Notch1* tKO cells are tumorigenic *in vivo* via Ccl2-mediated immune evasion

Having determined that either PN dKO or PCN tKO is sufficient to induce the neoplastic phenotype *in vitro* (Fig. 1, 2), we determined the impact of PN dKO and PCN tKO on ESCC tumorigenesis *in vivo* by congenic transplantation. Intriguingly, only PCN cells developed tumors, with a 60% success rate (3 of 5), while PN did not grow (0 of 5) in the immunocompetent (C57BL/6) mice (Fig. 4a, b), which was unexpected considering the superior cell growth of PN compared to that of PCN *in vitro* (Fig. 1k, l). PCN cell-derived tumors had substantially invaded the muscular layer with poorly differentiated squamous cell carcinoma (Fig. 4b). PCN tumors were hyperproliferative (Mki67), with ESCC stemness marker (Sox2 and Trp63) expression, recapitulating the features of ESCC (Fig. 4c).

These results led us to hypothesize that PCN cells evade immune surveillance, while PN cells are easily targeted by immune cells. Indeed, $Cd4^+$ T cells, $Cd8^+$ T cells, neutrophils, and dendritic cells infiltrated into the PCN-derived tumors, implying an active interaction between tumor and immune cells (Extended Data 4a). To test this, we comparatively examined PN and PCN transcriptomes. An Enrichr analysis showed that the genes associated with T-cell receptor signaling and antigen processing and presentation were highly enriched in the PCN compared to the PN (Extended Data 4b). A DEG analysis of bulk RNA-seq showed that the genes related to immune response and cytokine-related signaling pathways were highly expressed in PCN compared to in PN (Fig. 4d and Extended Data 4c), suggesting that cytokines or chemokines specifically expressed in PCN cells contribute to immune evasion. A comparative DEG analysis of scRNA-seq data sets identified three cytokines (Ccl2, Cxcl1, and Cxcl2) that were highly upregulated in PCN (Fig. 4e). Notably, only Ccl2 expression was specifically enriched in the PCN compared to WT, PC, and PN (Fig. 4f and Extended Data 4d-h). Consistently, immunostaining validated the high expression of Ccl2 in PCN allograft tumors and Ccr2, a receptor for Ccl2, in the immune cells surrounding the tumor (Extended Data 4i).

Since CCL2 expressed by tumor cells is known to recruit myeloid cells (monocytes, tumor-associated macrophages [TAMs], and myeloid-derived suppressor cells [MDSCs]) to the TME, inhibiting the Ccl2-Ccr2 axis was shown to enhance the immune response to tumors[23-26]. To

determine the role of Ccl2 in immune evasion of PCN cells, we genetically ablated the *Ccl2* gene in PCN cells (PCN-*Ccl2* KO) and compared tumorigenicity of PCN and PCN-*Ccl2* KO cells using allograft transplantation. PCN and PCN-*Ccl2* KO showed no difference in cell growth *in vitro* (Extended Data 5a, b). However, unlike PCN cells, PCN-*Ccl2* KO cells barely developed tumors in immunocompetent mice (Fig. 4g-i). Although immune cells were abundant in both PCN parental and *Ccl2* KO tumors, the keratin pearls, one of the signatures of ESCC, were mainly observed in PCN tumors compared to *Ccl2* KO tumors, indicating that PCN tumors develop with less interruption by immune cells (Extended Data 5c). In line with this, the number of exhausted T (T_{ex}) cells (expressing Pdcd1/Pd-1) was increased among the immune cells surrounding PCN tumors compared to PCN-*Ccl2* KO tumors (Fig. 4j, k and Extended Data 5d). Moreover, PCN-*Ccl2* KO tumors showed a higher number of cells with perforin, an effector $Cd8^+$ T cell marker, $Cd8^+$ T cells, and cleaved-Caspase-3^+ (an apoptotic cell marker) cells compared to PCN tumors, suggesting that more cell death mediated by T cells occurs in PCN-*Ccl2* KO tumors than in PCN tumors (Fig. 4j, k and Extended Data 5d). Simultaneously, a more significant number of $Cd206^+$ (an M2 macrophage marker) cells was observed in PCN tumors than in PCN-*Ccl2* KO tumors (Fig. 4j, k and Extended Data 5d).

Although we found fewer M2 macrophage marker–expressing cells in PCN-*Ccl2* KO tumors than in PCN tumors, we did not find significant differences in the number of M1 macrophage ($CD68^+$/$Cd80^+$ cells) or MDSC ($Cd11b^+$/$Ly6g^+$ cells) infiltration (Extended Data 5d-f). Nonetheless, the proliferation of PCN and PCN-*Ccl2* KO tumors was comparable based on the $Mki67^+$ cell quantification (Fig. 4j, k and Extended Data 5d). These results suggest that Ccl2 secreted from PCN tKO tumor cells contributes to immune evasion in ESCC tumorigenesis.

Emerging Ccl2-Ccr2 interaction between tumor and immune cells during ESCC tumorigenesis

To complement allograft models, we analyzed single-cell transcriptomics of an autochthonous mouse model[27]. In this model, treatment of 4-NQO (4-nitroquinoline 1-oxide) resulted in development of inflammation, hyperplasia, dysplasia, cancer in situ, and invasive cancer in the esophageal epithelium in a time-dependent manner. From the scRNA-seq datasets of the 4-NQO model, the epithelial cells (proliferating, basal, suprabasal, stratified, and other cells) and the immune cells (T, B, and myeloid cells) were sub-clustered based on marker gene expression (Extended Data 6a-f). This ESCC model showed decreased expression of Trp53, Cdkn2a, and Notch signaling pathways-related genes through ESCC development, similar to PCN organoids (Extended Data 7a). A comparative analysis of epithelial cell transcriptomes at each stage with ESCC patients showed increasing similarity as murine ESCC develops under 4-NQO exposure (Extended Data 7b, c), suggesting that the 4-NQO mouse model recapitulates human ESCC tumorigenesis. Notably, in hyperplasia, *Ccl2* and *Ccr2* expression was the highest in the esophageal epithelial cells and immune cells, respectively (Fig. 5a, b and Extended Data 7d, e). A further subtype analysis showed that *Ccr2* expression was enriched in the T_{ex} cells and regulatory T (T_{reg}) cells (Extended Data 7f). Moreover, we found that T_{ex} cell marker (Cd160, Havcr2/Tim-3, Pdcd1/Pd-1, and Lag3) expression was increased in the T cells at the early stage of ESCC tumorigenesis (Extended Data 7g), consistent with our PCN-transplanted tumor cells (Fig. 4 and Extended Data 5). To investigate the intercellular CCL-mediated ligand-receptor interaction, epithelial and immune cell communications were analyzed using the Squidpy and CellChat

packages. The CCL2-CCR2 interaction was inferred to be strong from the hyperplasia state and maintained until invasive cancer (Fig. 5c). Moreover, the total interactions related to the CCL pathway between epithelial cells and immune cells were increased during ESCC development (Fig. 5d and Extended Data 7h). The CCL pathway–specified cell-cell communications were observed starting from the inflammation status in epithelial cells, T_{ex} cells, MDSCs, and macrophages (Fig. 5e and Extended Data 7i).

Given the role of the CCL2-CCR2 axis in the immune evasion of tumor cells[28], we tested whether the CCL2-CCR2 axis is indispensable for ESCC initiation using two different Ccr2 inhibitors (Ccr2 22 and BMS-813160). To increase tumorigenic efficiency, we established PCN cells that stably expressed Sox2 (PCNS) as Sox2 promotes ESCC tumorigenicity[29]. Sox2 overexpression accelerated the tumor development of PCN cells while not affecting Ccl2 expression (Extended Data 8a-c). Although Ccr2 inhibitors did not suppress PCNS cell growth *in vitro*, allograft tumors grown with Ccr2 inhibitors were drastically reduced in size and incidence compared to the tumors treated with vehicle (Fig. 5f and Extended Data 8d). All tumors that were exposed to vehicle and Ccr2 inhibitors showed abundant immune cell infiltration into the tumors and a comparable number of proliferating ($Mki67^+$) tumor cells (Extended Data 8e, f). Consistent with the *Ccl2* KO tumor results, T_{ex} cell ($Pdcd1^+$) infiltration into Ccr2-inhibited tumors was markedly decreased, while active $Cd8^+$ T cells ($perforin^+$ cells) were enriched in the tumors from Ccr2 inhibitor-injected mice compared to in the tumors from vehicle-injected control mice (Fig. 5g, h and Extended Data 8f, g). On the other hand, fewer M2 macrophages were observed in the Ccr2-inhibited tumors than in the *Ccl2* KO tumors (Fig. 5g, h and Extended Data 8f). However, the number of other monocyte-derived cells ($Cd68^+$ or $Cd209^+$) was lower in the Ccr2 inhibitor–treated TME, indicating reduced monocyte recruitment by Ccr2 blockade (Extended Data 8f, g). These results suggest that the CCL2-CCR2 axis elicits immune evasion for ESCC initiation.

NF-κB pathway, the upstream regulator of Ccl2

Next, we determined the mechanism of CCL2 upregulation during ESCC tumorigenesis. Given that CCL2 is upregulated mainly by transcription factors (TFs)[30], we sought to identify the key TFs that transactivate *CCL2*. An scRNA-seq-based regulon analysis identified PCN-specific regulons (RELA, FOXA1, MYC, TAF1, and TCF12) that may bind to the *CCL2* promoter (Extended Data 9a, b). Among them, Rela was the most highly activated in PCN, and the *Ccl2* gene is expected to be directly transactivated by Rela from a refined analysis using iRegulon (Extended Data 9c, d). Compared to PN, PCN cells showed more abundant nuclear accumulation of Rela and Rela inhibition downregulated *Ccl2* in PCN cells (Extended Data 9e, f). Moreover, a chromatin immunoprecipitation (ChIP) assay showed that Rela occupied the *Ccl2* promoter in PCN but not in PN cells, consistent with the results of a previous report that RELA transactivates *CCL2*[30,31] (Extended Data 9g). In addition, a human ESCC tissue microarray analysis showed a positive correlation between CCL2 and RELA expression, supporting the pathological relevance of the RELA-CCL2-CCR2 axis to human ESCC tumorigenesis (Extended Data 9h, i).

Classification of ESCC patients, and PCN relevance

To extend our findings to the biology of human ESCC tumorigenesis, we analyzed single-cell transcriptomics of 69 ESCC patients (Fig. 6a). We integrated 69 human ESCC scRNA-seq datasets with the murine PCN scRNA-seq dataset. The unsupervised principal components analysis and Pearson correlation coefficient categorized 69 patient datasets into three groups (ESC1, 2, and 3). While ESC1, 2, and 3 constituted a comparable portion of patients (ESC1: 30.4%; ESC2: 33.3%; ESC3: 36.2%) with tumor cell heterogeneity as described previously[10], PCN belonged to ESC1 (Fig. 6b, c). Notably, the ESC1 displayed the lowest pathway score for TP53, CDKN2A, and NOTCH signaling, suggesting that ESC1 has consistent transcriptomic features of PCN EOs (Fig. 6b, d). Furthermore, ESC1 and ESC2 showed the highest score for the NF-κB pathway, which was correlated with *CCL2* expression (Fig. 6e, f). Consistent with our findings of the immune landscape of PCN tumors, tumor cells of ESC1 patients expressed a higher level of CD274/PD-L1, suggesting that PCN-type (ESC1) likely has immune escape potential through PD-L1-PD1-mediated T cell exhaustion (Fig. 6f and Extended Data 10a). Of note, despite the transcriptomic differences among the three groups of patients, we could not find categorical differences between the conventional tumor stage classification and ESC1-3 groups (Extended Data 10b).

We further identified the representative marker genes of each group of patients by DEG analysis. ESC1 showed higher expression of *B2M*, *RGS1*, and *THEMIS2*, which were reported to be highly expressed in several types of cancer[32-34]. On the other hand, *PERP*, *KRT15*, and *CD9* were highly expressed in ESC2, and *CD151*, *IGFBP4*, and *EID1* were enriched in ESC3 patients (Fig. 6g, Extended Data 10d). Although these markers are upregulated in various cancers[35-40], they have not been investigated in ESCC. ESCC patients with higher expression of *B2M*, *RGS1*, *THEMIS2*, *PERP*, and *CD151* experienced poor survival compared to those with lower expression (Extended Data 10d). A TCGA-based analysis also showed that B2M, RGS, PERP, KRT15, and CD9 were upregulated in ESCC compared to normal samples (Extended Data 10e). Immunostaining of the ESCC tissue microarray (n = 101) revealed a positive correlation between B2M and CCL2 expression and B2M and RELA expression (Fig 6h, i, and Extended Data 10f). Most patients (90.9%) showed relatively higher expression of B2M in tumor samples than in adjacent normal samples, and 22.7% (5 of 22, immunohistochemistry score ≥ 2.5) of patients showed robust expression of both CCL2 and B2M in the tumor tissue (Fig. 6k). Focusing on tumor samples alone, 49.0% (25 of 51) of tumor samples expressed high levels of B2M (immunohistochemistry score ≥ 2.5), a slightly higher proportion compared to the ESC1 from scRNA-seq. These data suggest that B2M is a prognostic marker of ESCC likely with the RELA-CCL2 axis activation.

Discussion

Understanding ESCC initiation may improve its early detection, diagnosis, and treatment. However, the biology of ESCC initiation remains elusive. Moreover, model systems that recapitulate ESCC initiation are not available. In this study, we found that the genetic inactivation of *TP53*, *CDKN2A*, and *NOTCH1* (PCN) induces esophageal neoplasia (cell-autonomous) and generates an immunosuppressive niche (non-cell-autonomous) to promote ESCC tumorigenesis.

To identify the multiple genetic factors responsible for ESCC initiation, we genetically engineered EOs with 32 different combinations of nine candidate tumor suppressor genes based on the genomics of ESCC patients. *Trp53*, *Cdkn2a*, and *Notch1* were found to be essential tumor suppressor genes associated with ESCC initiation. While PN and PCN EOs displayed the neoplastic features of ESCC *in vitro*, only PCN cells exhibited *in vivo* tumorigenicity, indicating the existence of PCN-driven immune evasion. Transcriptomic analysis identified that the CCL2-CCR2 axis remodels immune landscape in ESCC tumorigenesis, as confirmed by the loss-of-function studies (*Ccl2* KO and Ccr2 inhibitors) and a scRNA-seq analysis of the autochthonous mouse model. Finally, we used single-cell transcriptomics to stratify and identify a group of PCN-type ESCC patients with new biomarkers.

We selected organoids as a model system for multiple genetic manipulations, which overcomes the limitation of single-gene KO (e.g., CRISPR-based screening) and provides a better phenotype analysis (i.e., neoplasia). However, the pitfalls of the organoid model include the lack of a niche, including mesenchymal and immune cells, later complemented by our congenic transplantation approach (PN vs. PCN) (Fig. 4a-c). Although we determined that PCN was a minimal genetic combination for ESCC initiation, our approach did not test the impact of oncogenes (e.g., *SOX2*, *NFE2L2*, and *PIK3CA*) that were frequently hyperactivated in ESCC. Indeed, the introduction of *SOX2*, which was amplified in 40% of ESCC patients[29], or $Kras^{G12D}$ accelerated *in vivo* tumorigenesis[19] (Extended Data 8b). Therefore, the combinatorial effects of tumor suppressor genes and oncogenes in ESCC tumorigenesis remain to be determined.

Single-cell transcriptomics has deepened our understanding of ESCC tumorigenesis. An scRNA-seq analysis unraveled distinct cell lineage trajectories and higher cell plasticity generated by PCN (Fig. 2e-i). Whereas WT organoids bear a single root cell cluster that differentiates into other cell clusters, PCN organoids harbor multiple root cell clusters that generate complicated and interconnected cell lineage trajectories, somewhat recapitulating the multiple clonal expansion and tumor heterogeneity of human cancer (Fig. 2h, i). Furthermore, single-cell transcriptomics expanded our experimental findings to stratifying ESCC patients. Unlike bulk RNA-seq, scRNA-seq allowed us to analyze tumor cells, excluding infiltrated immune cells. Despite the inter-tumoral heterogeneity of patient tissue samples, ESCC patients were classified into three groups based on the single-cell transcriptional signatures of tumor cells. ESC1 ESCC patients were highly correlated with murine PCN. These well-matched results are striking because a single-cell transcriptomics-based comparison of models (mice and organoids) in ESCC patients not only identifies key genetic factors (PCN) for ESCC tumorigenesis but also validates model systems, which could be further improved by precise mouse-to-human gene conversion and more patient datasets. Nonetheless, inter-tumoral heterogeneity remains to be addressed based on our ESCC classification.

Tumor cells evade immune surveillance in various ways[41,42]. Despite several studies describing the immune landscape of ESCC based on single-cell transcriptomics[9,10,43], experimental demonstration of ESCC immune evasion was not achieved. In this study, our comprehensive approaches discovered that *CDKN2A* inactivation plays a key role in establishing tumor-favorable immune cell niche. scRNA-seq-based immune profiling of the mouse ESCC model showed decreased effector T cells and increased immune suppressive dendritic cells during ESCC development[44], consistent with our finding (Fig. 4j, k and Extended Data 6e, f). Similarly, human ESCC generates an immune-suppressive landscape with enriched T_{ex} or T_{reg} cells and upregulation of immunosuppression-related genes in TAMs and dendritic cells[10,11,43]. Notably, enriched T_{ex} cells and TAMs are commonly observed in mouse scRNA-seq-based immune profiling and PCN TME (Fig. 4j, k and Extended Data 6e, f), unveiling the immune niche of at least PCN-associated ESCC tumorigenesis.

We found that CCL2 is the key immune landscape remodeling factor for ESCC initiation, as represented by increased T_{ex} cells and M2-like macrophages and a reduced number of effector $Cd8^+$ T cells and dendritic cells. The CCL2-CCR2 axis modulates the TME to be favorable to tumor cells[45]. Given the high expression of CCR2 in monocytes, CCR2-mediated M2-like macrophage differentiation and polarization had been extensively studied in the context of tumor cell invasion and metastasis[25]. In addition, the CCL2-CCR2 axis was shown to induce ESCC immune evasion via TAM accumulation[28]. Moreover, CCR2 expressed in T cells modulates the recruitment of T_{reg} cells[46], hampering immune surveillance. Therefore, targeting CCL2 to enhance the immune response to tumor development may be a promising option for ESCC treatment. Indeed, targeting the CCL2-CCR2 axis has been actively tested in clinical trials in other types of cancer, with or without immune checkpoint inhibitors (ICIs)[26,45,47]. However, it should be noted that since CCL2 is a chemoattractant that recruits monocytes, systemic inhibition of CCL2-CCR2 might affect normal immune surveillance. For instance, impaired monocyte recruitment by *Ccr2* blockade could reduce the number of antigen-presenting cells, diminishing innate and adaptive immune responses that are critical for anti-cancer effects[48]. Our results showed that the transient inhibition of Ccr2 suppressed tumor growth, but monocyte-derived cell ($Cd209^+$ or $Cd68^+$) recruitment was decreased while the number of $Cd8^+$ T cells was not significantly increased compared to the TME of *Ccl2* KO tumors (Fig. 5 and Extended Data 8). Therefore, targeting CCL2 or the upstream molecules of CCL2 in tumor cells would minimize potential adverse effects.

Recent clinical trials showed promising outcomes of ICIs in patients with advanced ESCC[12-16]. Nonetheless, ESCC patient stratification is a significant hurdle. Although the datasets from the 69 patients analyzed in this study lacked ICI-related clinical information, PCN-type (ESC1) patients showed higher expression of PD-L1/CD274 than did other patients. Thus, in addition to blocking the CCL2-CCR2 axis, ICIs targeting PD-L1 or PD-1 can be effective in ESC1 patients, and B2M could be a biomarker for ESC1-type patients (Fig. 6g and Extended Data 10c). Therefore, our study also proposes a new biomarker-guided strategy for treating PCN-type ESCC patients with ICIs (PD-L1 or PD-1 inhibitors) or CCL2 blockers.

Together, this study identifies the inactivation of *TP53*, *CDKN2A*, and *NOTCH1* as the key genetic event leading to ESCC development accompanied by immune evasion and newly classifies ESCC patients for better outcomes of immunotherapy.

Author contributions

K.-P.K. and J.-I.P. conceived the experiments. K.-P.K. and J.-I.P. designed the experiments. K.-P.K., Y.H., B.K., S.Z., G.Z., J.Z., and S.J. performed the experiments. K.-P.K., H.N., and J.-I.P. analyzed the data. C.M., K.J.D., G.E., A.-K.R., and H.N. provided ESCC samples. K.-P.K. and J.-I.P. wrote the manuscript.

Acknowledgments

We are grateful to Kwon-Sik Park, Pierre D. McCrea, and Malgorzata Kloc for insightful comments, Ann Sutton (Research Medical Library, MD Anderson) for editing the manuscript, and the Herbert Irving Comprehensive Cancer Center for the shared resources (Biostatistics, Genomics, and Molecular Pathology, and 3D organoid/cell culture). This work was supported by the Cancer Prevention and Research Institute of Texas (RP200315 to J.-I.P.), the National Cancer Institute (CA193297 and CA256207 to J.-I.P; 5P30CA013696 and 5P01CA098101 to A.-K.R., H.N., K.D., G.E., C.M.), an Institutional Research Grant (MD Anderson to J.-I.P.), a Specialized Program of Research Excellence (SPORE) grant in endometrial cancer (P50 CA83639), and Radiation Oncology Research Initiatives. The core facilities at Baylor College of Medicine (Cytometry & Cell Sorting Core and Single Cell Genomics Core) were supported by CPRIT (RP180672, RP200504) and the National Institutes of Health (CA125123, RR024574).

Methods

Mice

C57BL/6, *Trp53*$^{floxed/floxed}$ (JAX no. 008462), and *Rosa26*$^{nT\text{-}nG}$ (JAX No. 023035) mice were purchased from the Jackson Laboratory. Mice were bred and housed in an air-conditioned room of the Division of Laboratory Animal Resources facility at The University of Texas MD Anderson Cancer Center on a 12-hr light/dark cycle. All animal procedures were performed in accordance with the guidelines of the Association for the Assessment and Accreditation of Laboratory Animal Care and institutionally approved protocols (MD Anderson Institutional Animal Care and Use Committee). The study was compliant with all relevant ethical regulations regarding animal research.

Wnt3A, R-Spondin1, and Noggin conditioned medium

The Wnt3A, R-Spondin1, and Noggin (WRN) conditioned medium was prepared as previously described[49]. In brief, L-WRN (ATCC CRL-3276) cells were cultured on a 10-cm plate with culture medium (Dulbecco's modified Eagle's medium [DMEM], 0.5 mg/mL G418, 0.5 mg/mL hygromycin B, 1% penicillin/streptomycin, and 10% fetal bovine serum [FBS]). After 10% of L-WRL cells had been seeded in culture medium (without G418 and hygromycin B) in 10-cm plates, cells were incubated for 3-4 days. The medium was replaced with 10 mL of fresh medium when the cells were 80%-90% confluent, and the cells were incubated for 24 hrs. The medium was collected, centrifuged at 1000 × g for 4 min, passed through a 0.22-μm sterile filter, and stored at −80 °C. Another 10 mL of fresh medium was added to the plates and collected after 24 hrs to make the second batch of conditioned medium using the same procedures. The first, second, and third batches of conditioned media were mixed before use to prepare a 100% WRN conditioned medium.

Esophageal tissue isolation

The cervical dislocation was performed after the 8- to 10-week-old mice had been euthanized via CO2 inhalation. The esophagi were collected in 10-cm Petri dishes with ice-cold phosphate-buffered saline (PBS) with 1% penicillin/streptomycin and swirled gently to remove blood. The esophagi were opened longitudinally and washed with cold PBS with 1% penicillin/streptomycin. The epithelial cell layer was peeled off using surgical tweezers and then dissected into 0.5-cm3 pieces with surgical blades. The minced esophagi were collected in a 15-mL conical tube and digested by 0.05% trypsin-EDTA at 37 ℃ for 60 min with frequent vortexing. After dissociation, 3× volume 10% FBS-supplemented DMEM was added to inactivate the dissociation enzyme, followed by vigorous pipetting. The suspension was passed through a 35-μm sterile cell strainer to collect a single-cell suspension. Finally, the cell suspension was spun down at 1000 rpm for 4 min at ambient temperature and resuspended in a 50% WRN-conditioned medium.

EO culture

The E-MEOM included 50% WRN medium (50% advanced DMEM/F12, 50% WRN conditioned medium, 1% penicillin/streptomycin, and 1 × GlutaMAX), 1 × B27, 50 ng/mL, 10 mM nicotinamide, 500 nM A83-01, 10 μM SB202190, 50 ng/mL EGF, and 10 μM Y-27632 ROCK inhibitor (first 3 days). Single-cell dissociated esophageal epithelial cells (1000 cells) were suspended in 8 μL of E-MEOM and 12 μL of pre-thawed Matrigel (Corning) on ice and seeded in the centers of a well to create a Matrigel dome. The plate was incubated at 37 °C for 10 min to solidify the Matrigel. Finally, 500 μL of E-MEOM was added and incubated at 37 °C with 5%

CO2. The medium was changed every 2-3 days. See Supplementary Table 1 for reagents information.

***Trp53KO* EO**

8- to 10-week-old *Trp53*$^{floxed/floxed}$ mice were euthanized to collect the esophagi, which were digested into a single-cell suspension and seeded in Matrigel to form EOs, as described above. After 7 days of culture, EOs were digested with 0.05% trypsin-EDTA at 37 °C for 45 min to create a single-cell suspension. Ad-CMV-EGFP or Ad-Cre-EGFP (University of Iowa) was added to the cell suspension at 1 × 103 pfu/cell. Cells were then suspended in Matrigel to generate new EOs. After 2 days of seeding, 10 μM nutlin3 was added for *Trp53*$^{floxed/floxed}$ cell selection. Two days later, the selection was performed by sorting the GFP+ cells after dissociating EOs into single cells, and the sorted cells were re-seeded for EO culture as described above. EOs were collected for genotyping 7 days after seeding. Wild-type, *Trp53*$^{floxed/floxed}$, and KO Trp53 alleles were amplified as 288 bp, 370 bp, and 612 bp, respectively. See Supplementary Table 2 for primer information.

CRISPR/Cas9-based gene KO in EOs

CRISPR/Cas9 system-mediated gene KO was described in a previous study[50]. In brief, WT or Trp53 KO EOs were digested with 0.05% Trypsin-EDTA to dissociate into single cells. Single cells were incubated with a virus-containing medium with polybrene for 1 hr with centrifugation (600 g) at 32 °C (see Supplementary Table 3 for sgRNA sequences). Cells were then incubated at 37 °C with 5% CO2 for 4 more hrs and embedded in Matrigel. The medium was replaced 2 days after infection with antibiotics (puromycin, blasticidin, or hygromycin) for selection. Gene KO was confirmed by genomic DNA PCR after lysis of the organoid using specific primer pairs (see Supplementary Table 2 for primer information).

Sox2 stable cell line establishment

Sox2 overexpressing PCN cells were established by Lenti-viral transduction. Briefly, pLenti-Sox2-GFP plasmid was prepared by amplifying Sox2 template from the pLV-tetO-Sox2 plasmid(addgene, #19765), and then transferred to pLenti CMV GFP plasmid (addgene, #17446) using High-Fidelity DNA polymerase (NEB). Viral transduction was performed as described in a previous study[50]. GFP fluorescence expressing cells were isolated using flow cytometry (BD FACSAria II Cell Sorter) and cultured again.

Organoid-forming efficiency and size analysis

After 7 days of organoid seeding in Matrigel, the size of the organoids was analyzed by measuring the volume under the microscope (ZEN software, ZEISS). To reduce the vulnerability of EOs, the measurements were performed after at least 3 passages after isolation from the KO experiments. All experiments included more than 50 organoids per group.

H&E, PAS, and immunofluorescence staining

All staining was performed as previously described[51]. 7 days after seeding, EOs were collected by dissociating Matrigel using ice-cold PBS and fixed in 4% paraformaldehyde at ambient temperature. For tumor tissue, excised tumors were washed with ice-cold PBS and fixed with formaldehyde at ambient temperature. After paraffin embedding, tumor tissue and organoid sections were mounted on glass slides. For H&E staining, sections were incubated in hematoxylin for 3-5 min and eosin Y for 20-40 s. For PAS staining, slides were immersed in the periodic acid

solution for 5 min at ambient temperature and then in Schiff's reagent for 15 min at ambient temperature, followed by hematoxylin solution for 60-90 s. After washing with tap water, slides were dried, and the coverslips were mounted with mounting media. For immunofluorescence, after blocking with 3% goat serum in PBS for 30 min at ambient temperature, sections were incubated with primary antibodies (Krt13 [1:250], MKi67 [1:250], Sox2 [1:250], Trp63 [1:250], cleaved caspase-3 [1:250], Rela [1:200], Pdcd-1/Pd-1 [1:200], perforin [1:200], Cd8 [1:200], Cd206 [1:200], Cd209 [1:200], Havcr2/Tim3 [1:200], Cd4 [1:200], Mpo [1:200], Ccl2 [1:200], Ccr2 [1:200], Cd3 [1:200], Cd68 [1:200], Cd11b [1:100], Ly6g [1:100], and Cd80 [1:200]) overnight at 4 °C and secondary antibody (1:250) for 1 hr at ambient temperature. Sections were mounted with ProLong Gold antifade reagent with DAPI (Invitrogen). Images were captured with the fluorescence microscope (Zeiss; AxioVision). See Supplementary Table 1 for antibody information.

Immunohistochemistry

ESCC cancer tissue microarray slides contained 144 samples from 55 patients (provided by Dr. Hiroshi Nakagawa). Immunostaining was performed as previously described[51]. In brief, paraffin-embedded tissue antigens were retrieved with a basic or citrate antigen retrieval buffer. After being blocked with goat serum in PBS, tissues were incubated with primary antibodies (CCL2/MCP-1 [1:200], RELA [1:200], and B2M [1:200]). The immunohistochemistry results were scored, analyzed, and visualized using the "gglot2" package in r and GraphPad Prism (v9.2.0).

Gene expression analysis

Organoids were harvested and lysed, and RNAs were extracted by TRIzol reagent. RNA was quantified with a Nanodrop 2000c Spectrophotometer (Thermo Scientific) and then converted to cDNA using SuperScript II reverse transcriptase with random hexamers. PCR amplification (StepOne Real-Time PCR System, Applied Biosystems) was performed using the following conditions: 95 °C for 10 min, 95 °C for 15 s (denature), and 60 °C for 1 min (anneal/extend) for 40 cycles; 95 °C for 15 s and 60 °C for 1 min; and 95 °C for 15 s (melting curve). qRT–PCR results were quantified by comparative 2−ΔΔCt methods (see Supplementary Table 2 for primer information). The results were expressed as average fold change in gene expression and were normalized to the expression of mHprt.

BrdU incorporation assay

Organoids were incubated with BrdU (10 μM) for 0.5 hr from 7-day cultured organoids. BrdU-containing medium was removed, and Matrigel-embedded organoids were harvested. EOs were washed three times with PBS and fixed with 4% PFA for 30 min at ambient temperature. After the paraffin-embedding process, slides were prepared for staining. Following the immunofluorescence staining method, the slides were incubated with BrdU antibody at 4 °C and fluorescence-conjugated secondary antibody. Samples were further stained with DAPI and mounted for imaging.

Organotypic co-culture

Four-well chamber slides (Falcon) were coated with 200 μg/mL Matrigel and stored at 4 °C until cell seeding. Stromal cells from esophagi were isolated from Rosa26nT-nG mice. We collected the stromal cell layer by peeling off the epithelial cells from the longitudinally cut esophagi. Stromal cells were dissociated with collagenase/dispase for 1 hr at 37 °C and passed through a 70-μm cell strainer. After counting, cells (1.2 × 105 cells/well) were seeded in the pre-warmed

Matrigel-coated chamber. Cells were grown with DMEM supplemented with 20% FBS for 14 days. Epithelial cells were collected from WT, PN, and PCN organoids and seeded on the stromal cell–cultured chamber slide. The medium was replaced every 2 days and cultured for 14 days. For staining, cells were fixed with 4% PFA for 30 min at ambient temperature and then incubated with Cdh1 (1:200) overnight at 4 °C. Cells were visualized using secondary antibodies conjugated with fluorescein (1:200) for 1 hr at ambient temperature, and the images were taken with a fluorescence microscope (ZEISS; AxioVision).

2D culture

EOs were digested with 0.05% trypsin-EDTA at 37 °C for 45 min to create a single-cell suspension in a 15-mL centrifuge tube. The cells were centrifuged down at 1000 rpm for 4 min at ambient temperature, suspended in DMEM + 10% FBS with 10 μM Y-27632, and seeded on a 24-well plate. Cells were passaged every 3-5 days. After the third passage, Y-27632 was removed from the culture medium. DMEM supplemented with 10% FBS and 10% DMSO was used to freeze cells and store them in liquid nitrogen.

Colony formation, cell growth, wound-healing assays

Cells (1 × 104) were seeded on a 60-mm dish, and the medium was replaced every 2 days. The cells' colony-forming ability was monitored, and they were fixed with methanol for 20 min. The fixative was removed, and the dishes were rinsed with distilled water. The colonies were stained with 0.05% crystal violet solution, and the dishes were dried after being washed three times with distilled water. For the cell growth analysis, 2 × 105 cells were seeded on the 60-mm dish. The cell number was measured every day using an automated cell counter (Biorad) after the cells were trypsinized and stained with Trypan Blue. For the wound-healing assay, 1 × 104 cells were seeded on the 24-well plate. A wound was made in the middle of the well after the cells were confluent, and the cells were imaged at 0, 8, and 20 hr.

Xenograft and allograft transplantation

Five-week-old BALB/c nude mice and C57BL/6 mice were maintained in the Division of Laboratory Animal Resources facility at MD Anderson. 2D-cultured PN, PCN, PCN-Ccl2 KO, PCNS, and PCNS-Ccl2 KO cells (3 × 106) were injected subcutaneously into the right dorsal flanks of mice, respectively. Tumor volume was calculated by measuring with calipers every 3-4 days (volume = (length × width2)/2). For Ccr2 inhibitor treatment, DMSO-dissolved BMS Ccr2 22 (1 mg/kg) and BMS-813160 (25 mg/kg) were injected intraperitoneally at day 17, 21, and 25. DMSO was injected as vehicle control. Mice were euthanized, and tumors were collected at day 57. The excised tumors were photographed and paraffin-embedded for immunostaining.

Library preparation for RNA-seq and scRNA-seq

For bulk RNA-seq, triplicated WT, PN, and PCN organoids were harvested 7 days after seeding and lysed using the RNeasy Plus Mini Kit (QIAGEN). Purified RNA was used for mRNA library preparation, sequenced by Illumina NovaSeq, and mapped to GRCm38/mm10 genome (Novogene).

For scRNA-seq, organoids from WT, PC, PN, and PCN were collected 7 days after seeding and digested with 0.05% trypsin-EDTA at 37 °C for 30 min. After trypsin had been inactivated with 10% FBS DMEM, a single-cell suspension was collected by passing cells through a 35-μm cell strainer. Each group was tagged with two CMO tags from the CellPlex kit (10× Genomics).

The tagged cells of each group were pooled together with the same number of cells after being counted. The cDNA library was prepared with the 10× Genomics 3' v2 kit and sequenced on an Illumina NovaSeq (Novogene), mapped to the GRCm38/mm10 genome, and demultiplexed using CellRanger. The resulting count matrices files were analyzed in R (Seurat) or Python (Scanpy).

Public scRNA-seq data preparation

Mouse datasets. Mouse 4-NQO-treated scRNA-seq datasets were obtained from the Genome Sequence Archive in BIG Data Center (Beijing Institute of Genomics, Chinese Academy of Sciences, http://gsa.big.ac.cn) under the accession number CRA002118. Mouse datasets were directly used for analysis as downloaded, with filtered matrices.

Human datasets. ESCC patient datasets were downloaded from the National Center for Biotechnology Information Sequence Read Archive under the accession numbers PRJNA777911 and PRJNA672851. ESCC patient datasets were downloaded and converted to fastq files using Sequence Read Archive tools and the parallel-fastq-dump package. The converted fastq files were used as input for the CellRanger (v6.1.2) pipeline. Analyses were conducted with CellRanger output files. The datasets from 11 patients (PRJNA777911) and tumor samples that had been previously sorted by Cd45− selection from 58 patients (PRJNA672851) were input to run CellRanger (v7.0.1). Patient ID and tumor grade are described in Supplementary Table 5.

scRNA-seq data analysis

integration and clustering. Organoid scRNA-seq data analysis. For pre-processing and clustering of scRNA-seq data, we used the Seurat R package and Python package Scanpy. For the organoid datasets, since two tags were used for one group of organoids, two datasets of the same genotype were integrated using "sctransform" in Seurat and "concatenate" in Scanpy and then annotated as WT, PC, PN, or PCN. The batch correction was performed in Scanpy using "Harmony". UMAP was used for dimensional reduction, and cells were clustered into 35 groups in Seurat. Each cluster was annotated based on marker gene information. Seurat was used for 4-NQO-treated mice datasets that had been sorted with $Cd45^-$ and $Cd45^+$ cells. Datasets were pre-processed, normalized separately, and annotated based on their marker gene expression. Disease status was annotated by converting the original implemented identity class, "X0W, X12W, X20W, X22W, X24W, X26W" to "Normal, Inflammation, Hyperplasia, Dysplasia, CIS, and Invasive cancer". Scanpy was used for human datasets preprocessing and integration. Each dataset was normalized separately and clustered by the "Louvain" algorithm. The clusters expressing a high level of PTPRC were removed to exclude immune cell contamination and then a total 69 patient datasets were integrated using the "concatenate" function. Batch effects were corrected using "Harmony". See Supplementary Table 6 for cell numbers in each cell cluster.

Proportion difference test. The differences between the clusters from the two datasets were tested using the scProportionTest package. The cluster difference between the two datasets was compared, and the significance was calculated from the *P* value and confidence interval for the magnitude difference via bootstrapping using the default parameter of permutation (n = 1000).

Trajectory inference. RNA velocity was used for cell lineage tracing and latent time inference based on the results of previous report[52]. Bam files produced from the CellRanger pipeline were used to create Loom files. Using the Python-based velocyto package, we produced Loom files that included unspliced and spliced read information. Two Loom files were generated for one genotype of organoid since the multiplex process had used two oligo tags for each genotype. The Loom files from the same genotype were merged using the "combine" function in the Loompy package. Cells were filtered, and dimensional reduction was performed following the default parameters using the scVelo and Scanpy packages. RNA velocity was calculated through a dynamical model, and cells were clustered using the "Louvain" algorithm. RNA velocity for all four datasets was performed with the same parameters (n_neighbors=10, n_pcs=50). Latent time analysis and a PAGA analysis were performed and plotted using the scVelo package.

Enrichr analysis. To perform Enrichr analysis of PCN- or PN-enriched genes, we performed DEG analysis from bulk RNA-seq and scRNA-seq results. DESeq2 package was used for bulk RNA-seq and PCN highly expressed 1516 genes were used as input gene list. For scRNA-seq data, 'FindMarkers' function was used to make DEG list from proliferating cell clusters of PCN and PN datasets. 100 genes highly expressed in PCN compared to PN were used for Enrichr analysis. The analysis was conducted following instructions as described[53]. See Supplementary Table 7 and 8 for DEG results.

fGSEA analysis. To perform a GSEA of PCN- or PN-enriched genes with human ESCC samples, we prepared a rank gene list from the DEG analysis of human ESCC vs. normal RNA-seq data (TCGA). The "glmLRT" method was used, and the fdr cut-off was 0.05. We collected 495 ESCC highly expressed genes and 258 normal highly expressed genes and filtered the not-assigned genes, resulting in 750 genes for analysis. The geneset was created from PCN or PN organoid highly expressed genes compared to WT organoid expressed genes from scRNA-seq data. To match the names of genes in the rank gene list and geneset, we converted mouse gene names to human gene names using the biomaRt package. In total, 2812 PCN and 1285 PN highly expressed genes were used to create genesets. The enrichment value was calculated and plotted with the fgsea package (permutation number = 2000).

Similarity test. The Scissor package was used to compare the transcriptome of the organoids scRNA-seq dataset and the bulk RNA-seq datasets of ESCC patients[54]. To compare scRNA datasets with ESCC patients, we downloaded ESCC patient data and normal sample HTSeq count data from the GDC data portal (TCGA-ESCA). After preparing the normal and ESCC gene expression matrix, we converted the human gene names to mice gene names using the biomaRt package. Not converted or assigned genes and duplicated gene names were removed from the matrix, and a Scissor analysis was performed with each scRNA dataset using the Cox regression model (alpha = 0.01). To compare scRNA-seq data with poor survival patient data, we downloaded ESCC patients' clinical metadata from the GDC data portal and binarized it into "poor survival" and "better survival". A Scissor analysis was conducted with scRNA datasets using the Cox regression model and 0.03 alpha value.

Regulon analysis. For the regulon analysis in organoids, we used the pySCENIC package[55]. The Loom file of each organoid dataset was used, and the regulon-based UMAP was redrawn after

binarization of each cell with regulons and an AUCell calculation. The AUCell of each regulon and cell cluster were combined to obtain a regulon specificity score, and we found the top regulons of each cluster. These processes were repeated five times in each organoid dataset (WT, PC, PN, and PCN). To find the Ccl2 relevance, we created modules of candidate transcription factors (Rela, Foxa1, Myc, Taf1, and Tcf12) in the PCN dataset. Transcription factors and their target genes were calculated and visualized using Cytoscape.

Cell-to-cell interaction analysis. CellChat and Squidpy packages were used for the cell-to-cell interaction inference. Epithelial cells, fibroblasts, and immune cell datasets were merged at each disease stage to generate gene expression matrices for the CellChat analysis. The merged Seurat objects were converted to an H5ad file for the Squidpy analysis. In the CellChat, the "CCL" pathway was specified for analysis. "Epithelial cell" was defined as a "source group" and "T cell", "myeloid cell", and "B cell" were selected as "target groups" in the Squidpy analysis (permutation = 100, *P* value threshold = 0.05).

Pathway score analysis. We used Scanpy with the "scanpy.tl.score_genes" function for the pathway score analysis. The analysis was performed with default parameters and the reference genes from the gene ontology biological process or the Kyoto Encyclopedia of Genes and Genomes database. The gene list for the score analysis is shown in Supplementary Table 9.

Chromatin immunoprecipitation assay

A chromatin immunoprecipitation (ChIP) assay was performed as previously described[56]. In brief, PCN cells were harvested after being crosslinked with 1% formaldehyde for 10 min at ambient temperature, followed by quenching with glycine (final concentration = 0.125M). Cells were lysed and subjected to sonication (15 rounds of 30 sec on and 30 sec off, Bioruptor 300 [Diagenode]). Lysates were cleared and incubated with pre-conjugated RELA antibody-Dynabeads overnight at 4 °C. Immunoprecipitates were washed and eluted, followed by incubation with RNase. After RELA protein degradation with Proteinase K, ChIP DNA was purified using a PCR purification kit (QIAGEN, #28106). Putative RELA binding sites of CCL2 promoter regions (region 1: -673~-683, region 2: -1803~-1813, and region 3: -2035~-2045) and non-specific regions (-9708~-9838) as a negative control were subjected to ChIP-qPCR. The primers for ChIP-qPCR are listed in Supplementary Table 2.

Kaplan-Meier analysis

The overall survival of ESCC patients based on gene expression was determined using a publicly available database (Kaplan-Meier Plotter, http://kmplot.com/analysis). Eighty-one ESCC patients were divided into "high" and "low" groups by the median expression of the gene of interest. The overall survival of the patient groups was compared for 70 months, and the log-rank *P* value is marked in the figure.

Statistical analysis

The Student's *t*-test was used for comparisons of two groups ($n \geq 3$), and one-way analysis of statistical variance evaluation was used for comparisons of at least three groups ($n \geq 3$). *P* values < 0.05 were considered significant. Error bars indicate the standard deviation (s.d.). All experiments were performed three or more times independently under identical or similar conditions.

Data availability

Bulk and scRNA-seq data are available via the Gene Expression Omnibus (GSE 213929; http://https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE213929).
(Log-in token for reviewers:)

Code availability

The code used to reproduce the analyses described in this manuscript can be accessed via GitHub (https://github.com/jaeilparklab/ESCC_project_1) and is available upon request.

References

1 Pennathur, A., Gibson, M. K., Jobe, B. A. & Luketich, J. D. Oesophageal carcinoma. *The Lancet* **381**, 400-412, doi:10.1016/s0140-6736(12)60643-6 (2013).

2 Wang, G. Q. *et al.* Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population. *Gut* **54**, 187-192, doi:10.1136/gut.2004.046631 (2005).

3 Bejarano, L., Jordao, M. J. C. & Joyce, J. A. Therapeutic Targeting of the Tumor Microenvironment. *Cancer Discov* **11**, 933-959, doi:10.1158/2159-8290.CD-20-1808 (2021).

4 Hinshaw, D. C. & Shevde, L. A. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res* **79**, 4557-4566, doi:10.1158/0008-5472.CAN-18-3962 (2019).

5 Keskin, D. B. *et al.* Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. *Nature* **565**, 234-239, doi:10.1038/s41586-018-0792-9 (2019).

6 Friedl, P. & Alexander, S. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell* **147**, 992-1009, doi:10.1016/j.cell.2011.11.016 (2011).

7 Baba, Y. *et al.* Tumor immune microenvironment and immune checkpoint inhibitors in esophageal squamous cell carcinoma. *Cancer Sci* **111**, 3132-3141, doi:10.1111/cas.14541 (2020).

8 Luan, S. *et al.* Advances in Drug Resistance of Esophageal Cancer: From the Perspective of Tumor Microenvironment. *Front Cell Dev Biol* **9**, 664816, doi:10.3389/fcell.2021.664816 (2021).

9 Chen, Z. *et al.* Dissecting the single-cell transcriptome network underlying esophagus non-malignant tissues and esophageal squamous cell carcinoma. *EBioMedicine* **69**, 103459, doi:10.1016/j.ebiom.2021.103459 (2021).

10 Zhang, X. *et al.* Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. *Nat Commun* **12**, 5291, doi:10.1038/s41467-021-25539-x (2021).

11 Zheng, Y. *et al.* Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun* **11**, 6268, doi:10.1038/s41467-020-20019-0 (2020).

12 Wang, Z. X. *et al.* Toripalimab plus chemotherapy in treatment-naive, advanced esophageal squamous cell carcinoma (JUPITER-06): A multi-center phase 3 trial. *Cancer Cell* **40**, 277-288 e273, doi:10.1016/j.ccell.2022.02.007 (2022).

13 Sun, J. M. *et al.* Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. *Lancet* **398**, 759-771, doi:10.1016/S0140-6736(21)01234-4 (2021).

14 Doki, Y. *et al.* Nivolumab Combination Therapy in Advanced Esophageal Squamous-Cell Carcinoma. *N Engl J Med* **386**, 449-462, doi:10.1056/NEJMoa2111380 (2022).

15 Yamamoto, S. & Kato, K. JUPITER-06 establishes immune checkpoint inhibitors as essential first-line drugs for the treatment of advanced esophageal squamous cell carcinoma. *Cancer Cell* **40**, 238-240, doi:10.1016/j.ccell.2022.02.009 (2022).

16 Okada, M. *et al.* Three-Year Follow-Up and Response-Survival Relationship of Nivolumab in Previously Treated Patients with Advanced Esophageal Squamous Cell

Carcinoma (ATTRACTION-3). *Clinical Cancer Research* **28**, doi:10.1158/1078-0432.Ccr-21-0985 (2022).

17 Gao, Y. B. *et al.* Genetic landscape of esophageal squamous cell carcinoma. *Nat Genet* **46**, 1097-1102, doi:10.1038/ng.3076 (2014).

18 Song, Y. *et al.* Identification of genomic alterations in oesophageal squamous cell cancer. *Nature* **509**, 91-95, doi:10.1038/nature13176 (2014).

19 Zheng, B. *et al.* A new murine esophageal organoid culture method and organoid-based model of esophageal squamous cell neoplasia. *iScience* **24**, 103440, doi:10.1016/j.isci.2021.103440 (2021).

20 in *PDQ Cancer Information Summaries* (2002).

21 Busslinger, G. A. *et al.* Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution. *Cell Rep* **34**, 108819, doi:10.1016/j.celrep.2021.108819 (2021).

22 Kinker, G. S. *et al.* Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat Genet* **52**, 1208-1218, doi:10.1038/s41588-020-00726-6 (2020).

23 Popivanova, B. K. *et al.* Blockade of a chemokine, CCL2, reduces chronic colitis-associated carcinogenesis in mice. *Cancer Res* **69**, 7884-7892, doi:10.1158/0008-5472.CAN-09-1451 (2009).

24 Hao, Q., Vadgama, J. V. & Wang, P. CCL2/CCR2 signaling in cancer pathogenesis. *Cell Commun Signal* **18**, 82, doi:10.1186/s12964-020-00589-8 (2020).

25 Qian, B. Z. *et al.* CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* **475**, 222-225, doi:10.1038/nature10138 (2011).

26 Tu, M. M. *et al.* Inhibition of the CCL2 receptor, CCR2, enhances tumor response to immune checkpoint therapy. *Commun Biol* **3**, 720, doi:10.1038/s42003-020-01441-y (2020).

27 Yao, J. *et al.* Single-cell transcriptomic analysis in a mouse model deciphers cell transition states in the multistep development of esophageal cancer. *Nature Communications* **11**, doi:10.1038/s41467-020-17492-y (2020).

28 Yang, H. *et al.* CCL2-CCR2 axis recruits tumor associated macrophages to induce immune evasion through PD-1 signaling in esophageal carcinogenesis. *Mol Cancer* **19**, 41, doi:10.1186/s12943-020-01165-x (2020).

29 Wu, Z. *et al.* Reprogramming of the esophageal squamous carcinoma epigenome by SOX2 promotes ADAR1 dependence. *Nat Genet* **53**, 881-894, doi:10.1038/s41588-021-00859-2 (2021).

30 Jin, J. *et al.* CCL2: An Important Mediator Between Tumor Cells and Host Cells in Tumor Microenvironment. *Front Oncol* **11**, 722916, doi:10.3389/fonc.2021.722916 (2021).

31 Ueda, A. *et al.* NF-kappa B and Sp1 regulate transcription of the human monocyte chemoattractant protein-1 gene. *J Immunol* **153**, 2052-2063 (1994).

32 Zhang, H. *et al.* B2M overexpression correlates with malignancy and immune signatures in human gliomas. *Sci Rep* **11**, 5045, doi:10.1038/s41598-021-84465-6 (2021).

33 Huang, W. C. *et al.* Novel function of THEMIS2 in the enhancement of cancer stemness and chemoresistance by releasing PTP1B from MET. *Oncogene* **41**, 997-1010, doi:10.1038/s41388-021-02136-2 (2022).

34 Li, S. L. *et al.* High expression of regulator of G-protein signalling 1 is associated with the poor differentiation and prognosis of gastric cancer. *Oncology Letters* **21**, doi:ARTN 322 10.3892/ol.2021.12584 (2021).

35 Wang, V. M. *et al.* CD9 identifies pancreatic cancer stem cells and modulates glutamine metabolism to fuel tumour growth. *Nat Cell Biol* **21**, 1425-1435, doi:10.1038/s41556-019-0407-1 (2019).

36 Dasgupta, A. *et al.* Anticachectic regulator analysis reveals Perp-dependent antitumorigenic properties of 3-methyladenine in pancreatic cancer. *Jci Insight* **7**, doi:ARTN e153842 10.1172/jci.insight.153842 (2022).

37 Rao, X., Wang, J., Song, H. M., Deng, B. & Li, J. G. KRT15 overexpression predicts poor prognosis in colorectal cancer. *Neoplasma* **67**, 410-414, doi:10.4149/neo_2019_190531N475 (2020).

38 Turczyk, L. *et al.* Tetraspanin CD151 regulates expression of fibroblast growth factor receptor-2 (FGFR2) in ductal carcinoma in situ cell line. *Eur J Cancer* **49**, S131-S131 (2013).

39 Xiao, Y., Zhu, S., Yin, W., Liu, X. F. & Hu, Y. IGFBP-4 expression is adversely associated with lung cancer prognosis. *Oncology Letters* **14**, 6876-6880, doi:10.3892/ol.2017.7014 (2017).

40 Miyake, S. *et al.* Cells degrade a novel inhibitor of differentiation with E1A-like properties upon exiting the cell cycle. *Mol Cell Biol* **20**, 8889-8902, doi:Doi 10.1128/Mcb.20.23.8889-8902.2000 (2000).

41 Spranger, S. & Gajewski, T. F. Impact of oncogenic pathways on evasion of antitumour immune responses. *Nat Rev Cancer* **18**, 139-147, doi:10.1038/nrc.2017.117 (2018).

42 Togashi, Y., Shitara, K. & Nishikawa, H. Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. *Nat Rev Clin Oncol* **16**, 356-371, doi:10.1038/s41571-019-0175-7 (2019).

43 Dinh, H. Q. *et al.* Integrated single-cell transcriptome analysis reveals heterogeneity of esophageal squamous cell carcinoma microenvironment. *Nat Commun* **12**, 7335, doi:10.1038/s41467-021-27599-5 (2021).

44 Yao, J. *et al.* Single-cell transcriptomic analysis in a mouse model deciphers cell transition states in the multistep development of esophageal cancer. *Nat Commun* **11**, 3715, doi:10.1038/s41467-020-17492-y (2020).

45 Fei, L., Ren, X., Yu, H. & Zhan, Y. Targeting the CCL2/CCR2 Axis in Cancer Immunotherapy: One Stone, Three Birds? *Front Immunol* **12**, 771210, doi:10.3389/fimmu.2021.771210 (2021).

46 Loyher, P. L. *et al.* CCR2 Influences T Regulatory Cell Migration to Tumors and Serves as a Biomarker of Cyclophosphamide Sensitivity. *Cancer Res* **76**, 6483-6494, doi:10.1158/0008-5472.CAN-16-0984 (2016).

47 Flores-Toro, J. A. *et al.* CCR2 inhibition reduces tumor myeloid cells and unmasks a checkpoint inhibitor effect to slow progression of resistant murine gliomas. *Proc Natl Acad Sci U S A* **117**, 1129-1138, doi:10.1073/pnas.1910856117 (2020).

48 Peters, W., Dupuis, M. & Charo, I. F. A mechanism for the impaired IFN-gamma production in C-C chemokine receptor 2 (CCR2) knockout mice: role of CCR2 in linking

the innate and adaptive immune responses. *J Immunol* **165**, 7072-7077, doi:10.4049/jimmunol.165.12.7072 (2000).

49 Miyoshi, H. & Stappenbeck, T. S. In vitro expansion and genetic modification of gastrointestinal stem cells in spheroid culture. *Nat Protoc* **8**, 2471-2482, doi:10.1038/nprot.2013.153 (2013).

50 Ko, K. P., Zhang, J. & Park, J. I. Establishing transgenic murine esophageal organoids. *STAR Protoc* **3**, 101317, doi:10.1016/j.xpro.2022.101317 (2022).

51 Jung, Y. S. *et al.* Deregulation of CRAD-controlled cytoskeleton initiates mucinous colorectal cancer via beta-catenin. *Nat Cell Biol* **20**, 1303-1314, doi:10.1038/s41556-018-0215-z (2018).

52 Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology* **38**, 1408-1414, doi:10.1038/s41587-020-0591-3 (2020).

53 Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, e90, doi:10.1002/cpz1.90 (2021).

54 Sun, D. *et al.* Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nat Biotechnol* **40**, 527-538, doi:10.1038/s41587-021-01091-3 (2022).

55 Van de Sande, B. *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* **15**, 2247-2276, doi:10.1038/s41596-020-0336-2 (2020).

56 Kim, M. J. *et al.* PAF remodels the DREAM complex to bypass cell quiescence and promote lung tumorigenesis. *Mol Cell* **81**, 1698-1714 e1696, doi:10.1016/j.molcel.2021.02.001 (2021).

Figure Legends

Extended Data 1 | Candidate genes KO in EOs. a, Oncogrids of nine candidate genes were generated based on the mutation and copy number alterations. Eighty-six patients with ESCC were analyzed from the TCGA database. **b,** Mutation types of candidate genes were analyzed in the patients' data from cBioportal and the TCGA. The genes harboring more than 50% of truncation or frameshift mutations are shown. **c,** Gene expression of FAT4 and KMT2C from ESCC patients (TCGA, $n = 95$) and normal (TCGA, $n = 11$) samples are shown with box plots. Data are shown as means ± SEM. **d,** Genetic status of each organoid line was designed as shown in the table. **e,** Bright-field images of organoids are shown. Images were taken on day 8 of passage 3. Scale bar = 50 μm. **f,** Hematoxylin-and-eosin-stained sections from KO organoids. Scale bar = 50 μm.

Fig. 1 | Genetic ablation of *Trp53* and *Notch1* induces esophageal hyperplasia and de-differentiation. a, Schematic structure of EOs. **b,** The volume of each organoid is displayed as a heatmap. **c,** H&E staining showing the structures and morphologies of each organoid. Scale bar = 50 μm. **d, e,** Organoids were stained with Mki67, Sox2, and DAPI (**d**) and the number of Mki67-stained cells was quantified (**e**). WT: $p53^{flox/flox}$, N: *Notch1* KO, C: *Cdkn2a* KO, CN: *Cdkn2a* KO + *Notch1* KO, P: $Trp53^{del/del}$, PN: $Trp53^{del/del}$ + *Notch1* KO, PC: $Trp53^{del/del}$ + *Cdkn2a* KO, PCN: $Trp53^{del/del}$ + *Cdkn2a* KO + *Notch1* KO. Scale bar = 20 μm. $^{****}P < 0.0001$. **f,** Trp63 staining of WT, PC, PN, and PCN organoids. Scale bar = 20 μm. **g,** BrdU-incorporated organoid cells were stained with anti-BrdU antibody, and the total cell nuclei were stained with DAPI. Scale bar = 10 μm. **h,** Number of BrdU-positive cells in an organoid of WT, PN, and PCN is shown in bar plots. $^{****}P < 0.0001$. **i,** Bright-field images of WT, PN, and PCN organotypic cultures were shown. Scale bar = 50 μm. **j,** Organotypic cultured WT, PN, and PCN cells were fixed and stained with E-cadherin. Red fluorescence was visualized to show stromal cells derived from nT-nG mice. Scale bar = 50 μm. **k, l,** Colony formation ability (**k**) and cell growth rate (**l**) of PN and PCN were evaluated at different time points.

Supplementary Fig. 1 | Establishment of CRISPR/Cas9-based KO organoids and validation. a, Representative images of mouse EO growth of $p53^{flox/flox}$ are shown. Images were taken from day 3 to day 9 of passage 3. Scale bar = 20 μm. **b,** KO efficiency of each sgRNA was validated with immunoblot. Three different sgRNA constructs of each gene were tested with transient transfection. **c,** Genomic DNA of each organoid was sequenced to verify CRISPR/Cas9-induced gene editing. Pink boxes show the sgRNA sequences of each gene.

Extended Data 2 | *Notch1* KO suppresses organoid differentiation and induces cell migration. a, b, Organoids were stained with Mki67, Sox2 (**a**), and Krt13 (**b**). PA: $p53^{del/del}$ + *Ajuba* KO, PCF4: $p53^{del/del}$ + *Cdkn2a* KO + *Fat4* KO, PCA: $p53^{del/del}$ + *Cdkn2a* KO + *Ajuba* KO, WT: $p53^{flox/flox}$, N: *Notch1* KO, C: *Cdkn2a* KO, CN: *Cdkn2a* KO + *Notch1* KO, and P: $Trp53^{del/del}$. Scale bar = 20 μm. **c,** Morphologies and internal structures of DAPT-treated P and PC organoids were compared with PN and PCN, respectively. DAPT was treated with different doses; low dose = 5 mM, high dose = 10 mM. Scale bar = 50 μm. **d,** A reversible Notch1 inhibition effect was shown in the bright-field microscopic images. Morphologies of P and PC were rescued after DAPT

942 removal in the second passage. DAPT (10 mM) was treated for 8 days for each passage. Upper
943 panel, magnification, ×50, Scale bar = 200 μm. Lower panel, magnification, ×100, Scale bar = 100
944 μm. **e,** Schematic overview of organotypic culture experiment. **f,** Bright-field images of 2D-
945 cultured cells. Images were taken on day 3 of passage 3. Scale bar = 100 μm. **g,** Wound closure
946 rates of PC, PN, and PCN cells were evaluated by analyzing images taken at 0, 8, and 20 hrs after
947 scratch with ImageJ software. **h,** Alcian Blue-PAS staining results of organoids are shown. Large
948 intestine tissue was used as a positive control. Scale bar = 50 μm.
949

950 **Fig. 2 | Single-cell transcriptomic analysis of genetically engineered esophageal organoids. a,**
951 Schematic overview of single-cell RNA-seq procedure. Four organoids (WT, PC, PN, and PCN)
952 were multiplexed for library preparation and demultiplexed after sequencing. **b, c,** UMAP of four
953 integrated datasets (WT, PC, PN, and PCN). **d,** The proportion of each cluster was compared to
954 the same cluster of the WT dataset. Cell types were annotated on the right side of the plots.
955 Statistical significance between the two groups was assessed by the permutation test. **e,** RNA
956 velocity–based UMAP projection based on the cell type in each dataset. **f,** Cell trajectory inference
957 by RNA velocity shown with streamline. **g,** A latent time analysis is shown in the velocity-based
958 UMAP. **h,** PAGA analysis of RNA velocity–based cell clusters showing the direction of cell
959 lineage on the UMAP. The size of the circle corresponds to the cell number. **i,** Representative
960 marker genes of root cell clusters are indicated with simplified lineage directions.
961

962 **Extended Data 3 | Integration and annotation of scRNA datasets of organoids. a,** Integrated
963 UMAP of WT, PC, PN, and PCN datasets (Seurat package). **b,** Heatmap of each cluster from the
964 integrated dataset. **c,** Cells were clustered by 35 subtypes and annotated as four cell types based
965 on the marker genes. **d,** Marker gene expression of each cell type shown in the dot plot. **e,** The
966 proportion comparison of each cell cluster from the PC dataset and WT.
967

968 **Fig. 3 | Transcriptomes of PN and PCN recapitulate the ESCC phenotype. a,** Dot plot showing
969 cancer stem cell marker gene expression in each cell cluster of the dataset. **b, c,** ESCC phenotype–
970 associated cells were marked as $ESCC^+$ cells in the Scissor-based UMAP projection (**b**), and the
971 proportion of $ESCC^+$ cells is shown in the bar plot (**c**). **d, e,** Poor survival of ESCC patient–
972 associated cells is displayed as Poor $survival^+$ cells in the UMAP (**d**), and the proportion was
973 analyzed (**e**). **f,** A gene set enrichment analysis was performed to compare the gene signature of
974 ESCC to each geneset of the PN or PCN dataset. **g,** Heatmap clustering of bulk RNA-sequencing
975 from WT, PN, and PCN organoids. **h,** An Enrichr analysis was performed with RNA-seq result.
976 PN-enriched genes or PCN-enriched genes compared to WT were analyzed using Bioplanet and
977 REACTOME databases; the TOP5 features of each database are shown. **P* value. **i,** Dot plots of
978 PN or PCN highly expressed genes were compared to the gene sets of the Kyoto Encyclopedia of
979 Genes and Genomes (KEGG) pathways. Dot size represents the gene number count, and colors
980 show the adjusted *P* value.
981

982 **Fig. 4 | *In vivo* tumorigenicity of PCN with pro-tumorigenic TME. a,** Tumor growth of PN and
983 PCN was monitored after transplantation into C57BL/6 mice. **b,** H&E staining showing tumor

cells and the TME. Red arrowheads, mitotic cells; black arrowhead, blood vessel; dotted circle, inflammatory cells. Scale bar = 100 μm. **c,** Transplanted tumor cells were stained with ESCC cancer stem cell marker Trp63 and proliferative cell marker Mki67. Scale bar = 100 μm. **d,** An Enrichr analysis with a bulk RNA-seq result. DEGs were created with PCN and PN datasets, and PCN highly expressed genes were input for a Reactome analysis. **e,** Volcano plot from a DEG analysis of PCN and PN from a scRNA analysis. **f,** Feature plot of Ccl2 expression from WT, PN, and PCN scRNA-seq results. Split UMAPs were displayed with each dataset. **g-i,** Tumor growth measured in PCN and PCN-*Ccl2* KO cell–transplanted mice. **j,** Pdcd1^{+}, perforin^{+}, Cd8^{+}, Cd206^{+}, Cd209^{+}, and Mki67^{+} cells in randomly chosen 630× magnified images were analyzed and plotted. n.s. = not significant. **k,** Pdcd1/Pd-1, perforin, Cd8, Cd206, Cd209, and Mki67 staining of PCN and PCN-*Ccl2* KO–derived tumors. Scale bar = 20 μm.

Extended Data 4 | Ccl2 relevance in immune response. a, Tumor-infiltrated immune cells of PCN-transplanted tumor tissues were stained with Havcr2/Tim3, Pdcd1/Pd-1, Cd8, Cd4, Mpo, and Cd209/Dc-sign. Nuclei were stained with DAPI. Scale bar = 50 μm. Dotted circle, inflammatory cells. **b,** An Enrichr analysis from PCN highly expressed genes in proliferating cells from the scRNA-seq results of PN and PCN datasets. DEGs were created between proliferating cells of PCN and PN, and the relevant pathways were analyzed with BioPlanet and the KEGG databases. **c,** An Enrichr analysis with bulk RNA-seq results. PCN highly expressed genes compared to PN were analyzed using gene ontology biological process (GOBP) database. **d,** Dot plots showing the gene expression of *Ccl2*, *Cxcl1*, and *Cxcl2* in a different cell type and datasets of organoids from scRNA-seq results. **e,** *Ccl2* expression in the PC dataset displayed on UMAP. **f,** Dot plot from scRNA-seq showing *Ccl2* gene expression in different genotypes of organoids. **g, h,** Cxcl1 and Cxcl2 gene expression in different genotypes of organoids shown on a dot plot (**g**) and UMAP (**h**). **i,** PCN-transplanted tumors were stained with Ccl2 and Ccr2 antibodies. Scale bar = 50 μm.

Extended Data 5 | TME comparison of PCN- and PCN-*Ccl2* KO–transplanted tumors. a, Sanger sequencing results of PCN and PCN-*Ccl2* KO cells. Pink box showing the sgRNA sequences for the *Ccl2* gene. **b,** Growth rates of PCN and PCN-Ccl2 KO cells. **c,** H&E staining of PCN- and PCN-*Ccl2* KO–derived tumors. Scale bar = 100 μm. **d,** Pdcd1/Pd-1, Cd8, Cd3, Cd206, Cd209, Cd68, Cd11b, Ly6g, cleaved-Caspase3 (cCas3), and Mki67 staining in PCN- and PCN-*Ccl2* KO cell–derived tumors. Samples were counterstained with DAPI. Scale bar = 100 μm. **e,** Randomly chosen 630× magnified images were analyzed and plotted. n.s. = not significant. **f,**. Higher magnification (630×) images of Cd68, Cd80, Cd11b, and Ly6g staining. Scale bar = 20 μm.

Fig. 5 | CCL2-CCR2-induced immune evasion during ESCC development. a, Epithelial cells' Ccl2 expression of each disease status is shown with dot plots. **b,** Dot plot showing the Ccr2 expression in each type of immune cell based on the disease status. **c,** Ligand-receptor interactions predicted between epithelial cells and immune cells (using the Squidpy package). Ccl2- and Ccr2-interacting proteins are displayed with different disease stages. **d,** Ligand-receptor interactions related to the CCL pathway were predicted in sub-clusters of immune cells and epithelial cells. CCL pathway-related ligand-receptor interactions at different disease status (normal, hyperplasia,

and cancer in situ) were calculated and visualized with the circle plot using the CellChat package. **e,** Significant interactions of the CCL pathway in epithelial cells with T_{ex} cell, MDSC, and macrophage sub-clusters were plotted using a chord diagram. The predicted genes were annotated. Directions from the ligand (epithelial cell) to the receptor (immune cell) are indicated. **f,** Tumors from Sox2-overexpressed PCN (PCNS) cell–transplanted mice were monitored and measured after injection of Ccr2 inhibitors (Ccr2 22 and BMS-813160). DMSO was used as vehicle control. Reagents were intraperitoneally injected three times on days 17, 21, and 25. ***$P < 0.001$, ****$P < 0.0001$. **g, h,** Vehicle-, Ccr2 22-, and BMS-813160-treated tumors were stained with Pdcd1/Pd-1, perforin, Cd8, and Cd206 antibodies (**h**), and the results were analyzed (**g**). Scale bar = 20 μm, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$, n.s. = not significant.

Extended Data 6 | Unsupervised clustering and annotation of epithelial cells and immune cells of mouse ESCC model datasets. a, Heatmap showing the marker genes of each cluster of epithelial cells. **b,** scRNA-seq data of 4-NQO-treated mouse esophagus epithelial cells were projected with UMAP and clustered by disease status. Normal: 0 week of 4-NQO treatment, inflammation: 12 weeks of 4-NQO treatment, hyperplasia: 20 weeks of 4-NQO treatment, dysplasia: 22 weeks of 4-NQO treatment, cancer *in situ*: 24 weeks of 4-NQO treatment, invasive cancer: 26 weeks of 4-NQO treatment. **c,** UMAP projection of cells with cell type annotations based on the marker gene expression. **d,** Heatmap showing the marker genes of the immune cell sub-clusters. **e,** Immune cells of each disease status are clustered by cell types and displayed with UMAP. **f,** Detailed sub-clusters of immune cells are annotated by marker gene expression using the CellKb database.

Extended Data 7 | Ccl2-Ccr2 interaction in epithelial and immune cells of the mouse ESCC model. a, Trp53, Cdkn2a, and Notch signaling pathway genes expression were displayed at the different disease stages after 4-NQO treatment by dot plot. **b,** ESCC patients-associated cells were assessed using the Scissor package and displayed with red dots in the UMAP. Blue dots represent normal cell-associated cells, and gray dots are background cells. **c,** ESCC-associated cell proportions of each disease status are displayed with bar plots. **d, e,** Feature plot (**d**) and dot plot (**e**) of Ccr2 expression in immune cells. Plots were visualized on the basis of the disease status. **f,** Ccr2 expression displayed in the dot plot for the sub-clusters of T cells. **g,** T_{ex} cell marker gene expression in the immune cell clusters were displayed with a dot plot. **h,** Ligand-receptor interactions related to the CCL pathway were predicted in sub-clusters of immune cells and epithelial cells. Interactions in the CCL pathway in each disease status (inflammation, dysplasia, and invasive cancer) were calculated and visualized with the circle plot. **i,** Significant interactions of the CCL pathway in epithelial cells with T_{ex} cell, MDSC, and macrophage sub-clusters were plotted in inflammation, dysplasia, and invasive cancer status using a chord diagram. The predicted genes were annotated for ligands in epithelial cells and receptors in immune cells.

Extended Data 8 | Immune landscape regulation by Ccr2 inhibitors during tumorigenesis. a, qRT-PCR results showing the relative mRNA expression of Ccl2 in PCN and Sox2-overexpressed PCN (PCNS) cells. n.s. = not significant. **b, c,** Tumorigenicity comparison in PCN and PCNS cells. **d,** Cell growth of PCNS was assessed after treatment with DMSO or 1 μM of Ccr2 inhibitors (Ccr2

22 and BMS-813160) by measuring the cell number. **e,** H&E staining of vehicle-, Ccr2 22-, and BMS-813160-treated tumors. **f,** Pdcd1/Pd-1, Cd8, Mki67, Cd206, Cd68, and Cd209 staining images in vehicle-, Ccr2 22-, and BMS-813160-treated tumors. Scale bar = 100 μm. **g, h,** Images of high magnification (630×) from Cd8 and Cd209 staining results in vehicle-, Ccr2 22-, and BMS-813160-treated tumors (**g**) and quantitation of $Cd209^+$ cells (**h**). Scale bar = 20 μm. $^{**}P < 0.01$, n.s. = not significant.

Extended Data 9 | Rela/NF-κB is an upstream regulator of *Ccl2*. a, Venn diagram showing the overlapped transcription factors of the UCSC ChIP seq database-, PN-, and PCN-specific regulons. **b,** Schematic representation of human CCL2 promoter binding sites and transcription factors predicted in PCN regulons. **c,** Rela regulon projection in the regulon-based UMAP of WT, PC, and PCN datasets. The color shows the regulon specificity score of the cells. **d,** Network analysis of transcription factors related to human Ccl2 gene expression. Target genes of transcription factors were refined by iRegulon, and the Ccl2 connection to Taf1 and Rela was displayed. **e,** Rela immunofluorescence images of 2D-cultured PN and PCN. Nuclei were stained with DAPI. Scale bar = 10 μm. The proportion of cells with nuclear-accumulated Rela was evaluated in PN and PCN cells. **f,** qRT-PCR results showing Ccl2 expression in the PCN cells treated with different doses of NF-κB inhibitor for 6 hrs. **g,** ChIP assays showing binding activity of Rela to Ccl2 promoter in PN and PCN cells. Putative Rela binding sites (a, b, and c) and non-binding sites in the distant region (d) were analyzed with eluted DNA fragment amplification by PCR. **h,** Human ESCC tissue microarray slides from 112 samples were stained with RELA and CCL2 antibodies. Scale bar = 50 μm (lower magnification) and 20 μm (higher magnification). **i,** Correlations between RELA and CCL2 are displayed with a heatmap. The Pearson correlation coefficient (r) and *P* value (*p*) are displayed. *n* = number of samples.

Fig. 6 | Classification of ESCC patients and PCN relevance. a, Tumor epithelial cells of 69 ESCC patients were integrated with PCN organoid cells and displayed with UMAP. **b,** Correlation matrix heatmap of ESCC patients and PCN-integrated datasets. The dendrogram showing the distance of each dataset on the basis of principal component analysis, and the Pearson correlation is displayed with a color spectrum. Groups of patients were categorized by dendrogram and correlation. Pathway scores of TP53, CDKN2A, and NOTCH were displayed on the top. PCN was excluded for pathway scores. **c,** UMAP showing each group of patients and PCN dataset. **d,** Pathway scores of TP53, CDKN2A, and NOTCH were displayed in each group of patients using a dot plot. **e,** Dot plot showing the NF-κB score genes in each group of patients. **f,** CCL2**,** Cd274 (PD-L1) and PDCD1LG2 (PD-L2) gene expression in each group of patients. **g,** Representative marker gene expressions of each group were visualized in each patient. Genes were selected from the DEG analysis using the Wilcoxon method. **h, i,** B2M and CCL2 staining results from human ESCC tissue microarray samples (**h**) and the correlation of the results (**i**). Scale bar = 50 μm (lower magnification) and 20 μm (higher magnification). Pearson correlation coefficient (r) and *P* value (p) are displayed. *n* = number of samples. **j, k,** Heatmap of B2M staining results in tumor samples (**j**) and B2M, CCL2, and RELA staining results in ESCC patients with adjacent normal and tumor paired samples (**k**). IHC scores displayed from 1 (lowest expression) to 3 (highest expression).

Extended Data 10 | Characterization of ESCC patient subgroups. a, CD274 (PD-L1) expression in each patient shown with a dot plot. **b,** Stacked bar plot with each group of patients and their tumor grade. **c,** Expression of marker genes of each patient group. **d,** Expression of marker genes from each group was assessed by survival duration. A Kaplan-Meier plot was drawn by splitting ESCC patients from the TCGA database by the median value of each gene expression. **e,** Expression levels of marker genes of each group were measured in ESCC patients of the TCGA database and shown with box plots. NT: normal tissue of EAC and ESCC, TP: tumor from ESCC patients. **f,** Heatmap display of the correlation matrix of human ESCC tissue microarray sample staining results of B2M, CCL2, and RELA.

Extended Data 1

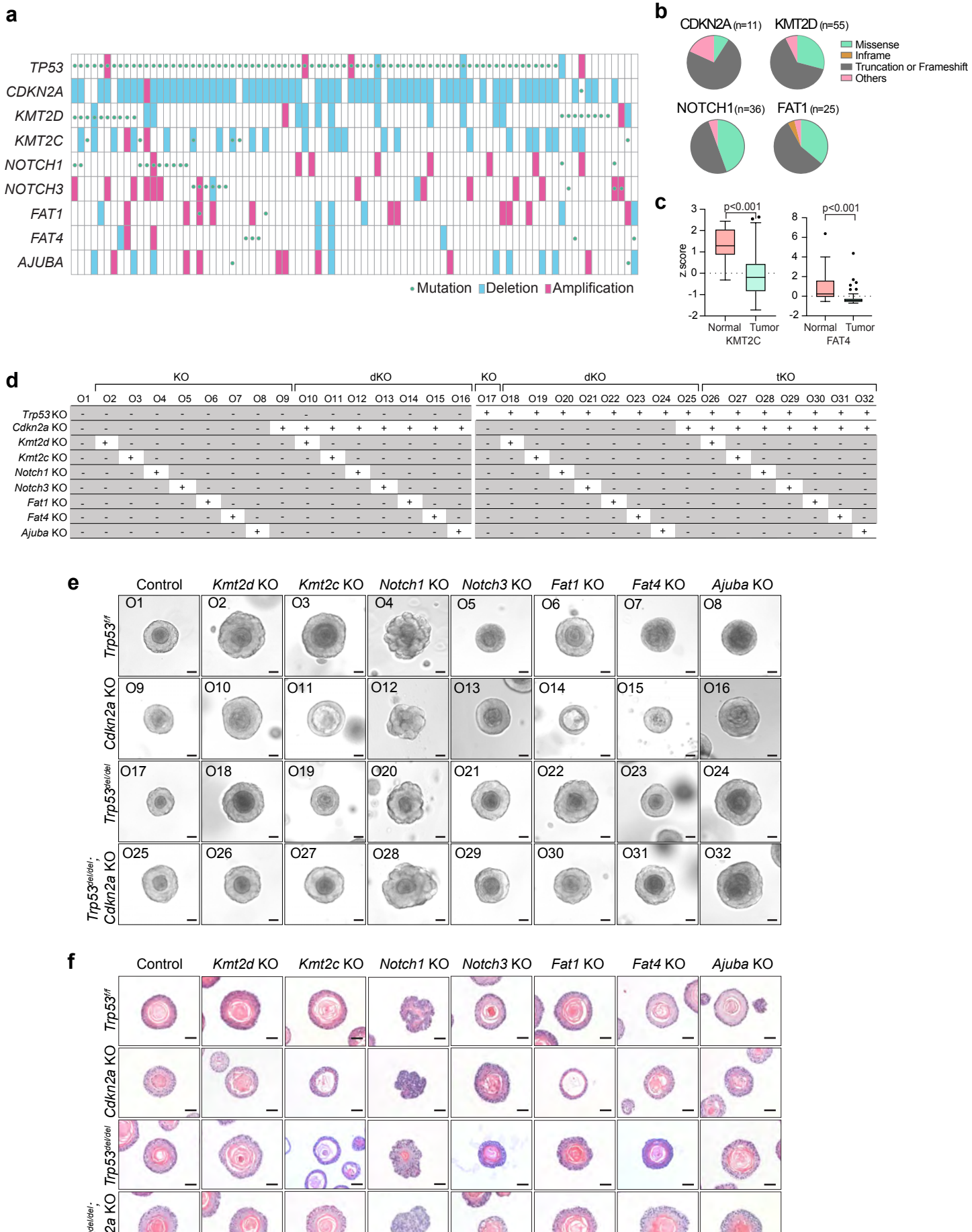
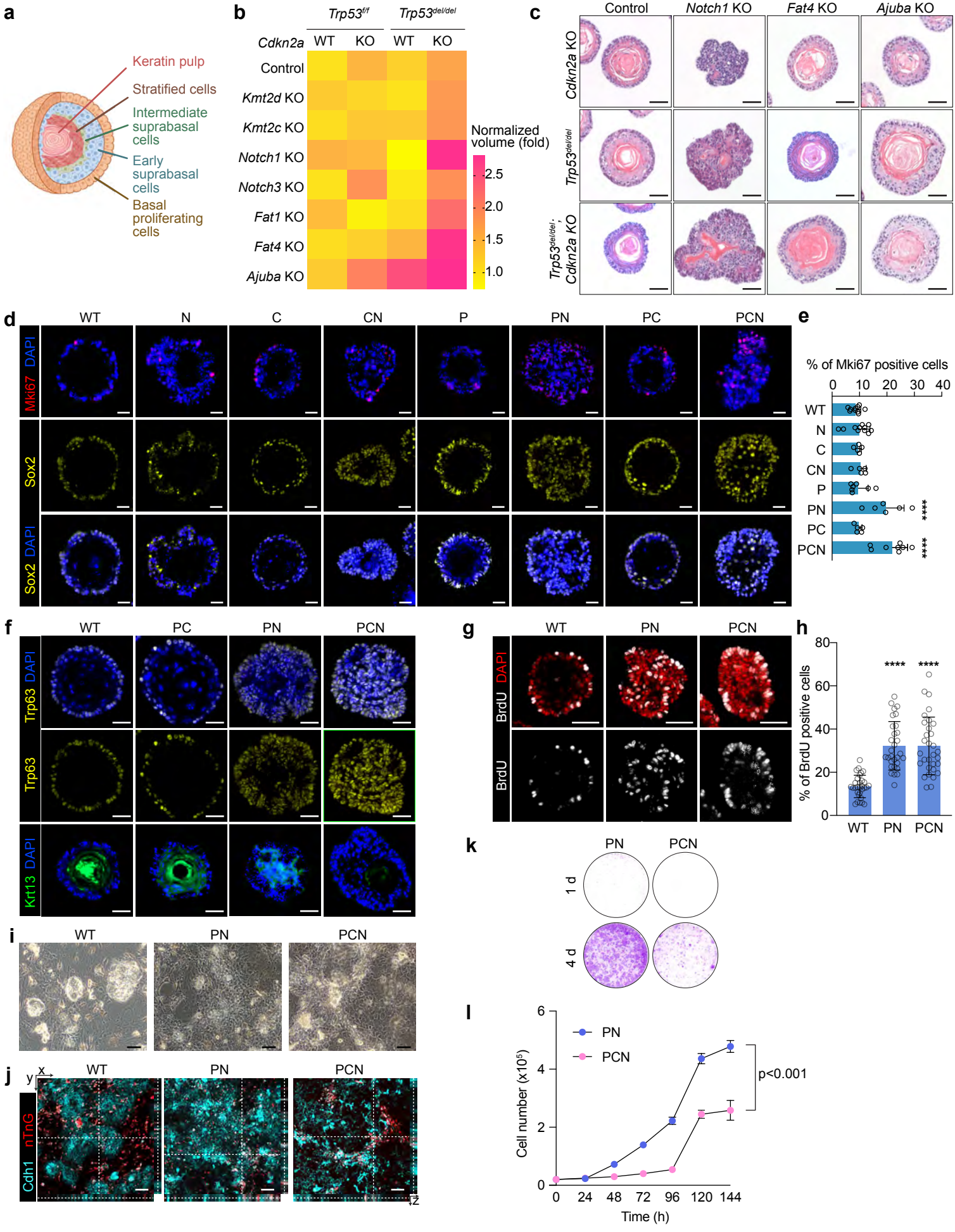
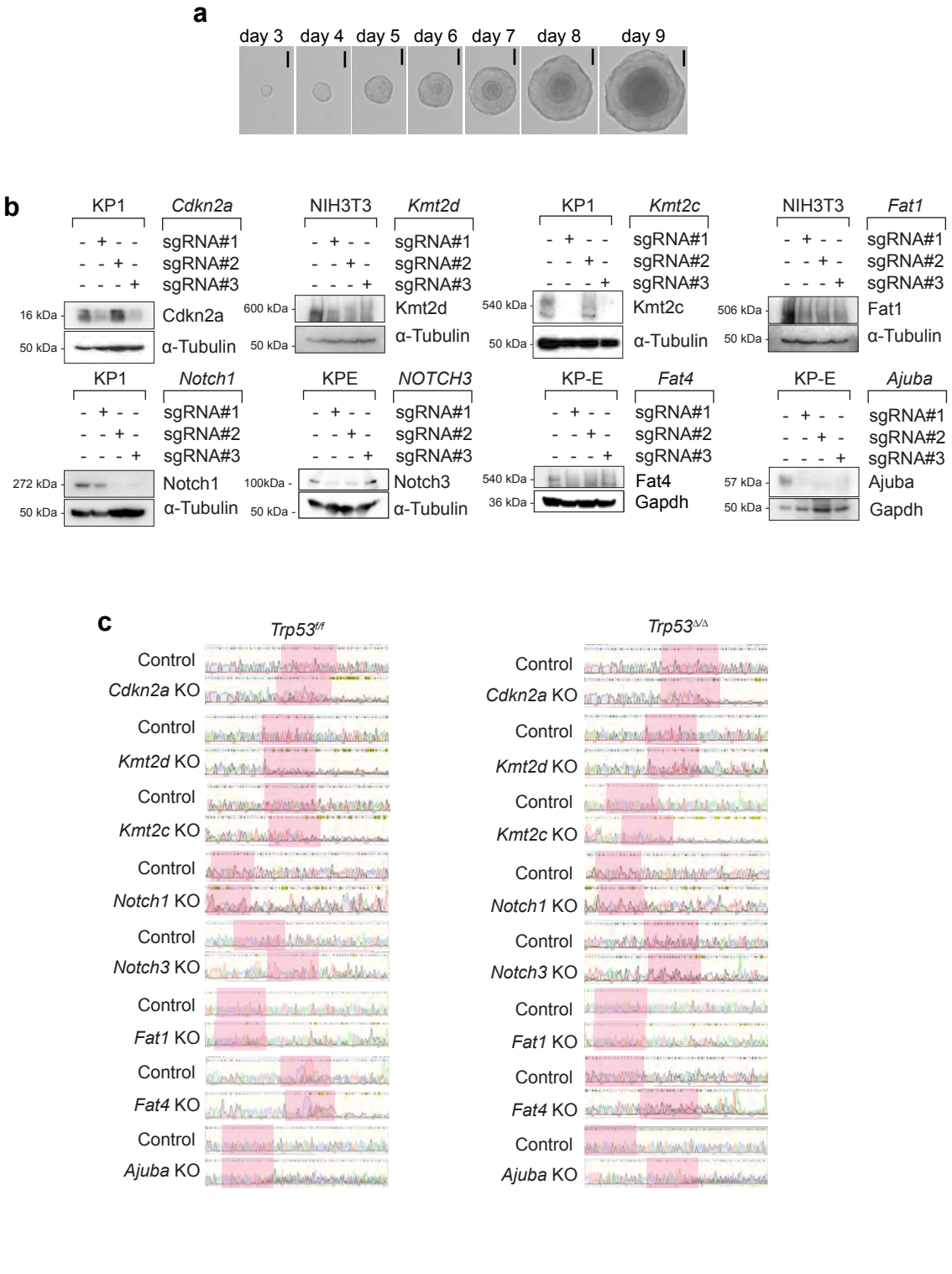


Figure 1



Supplementary Figure 1



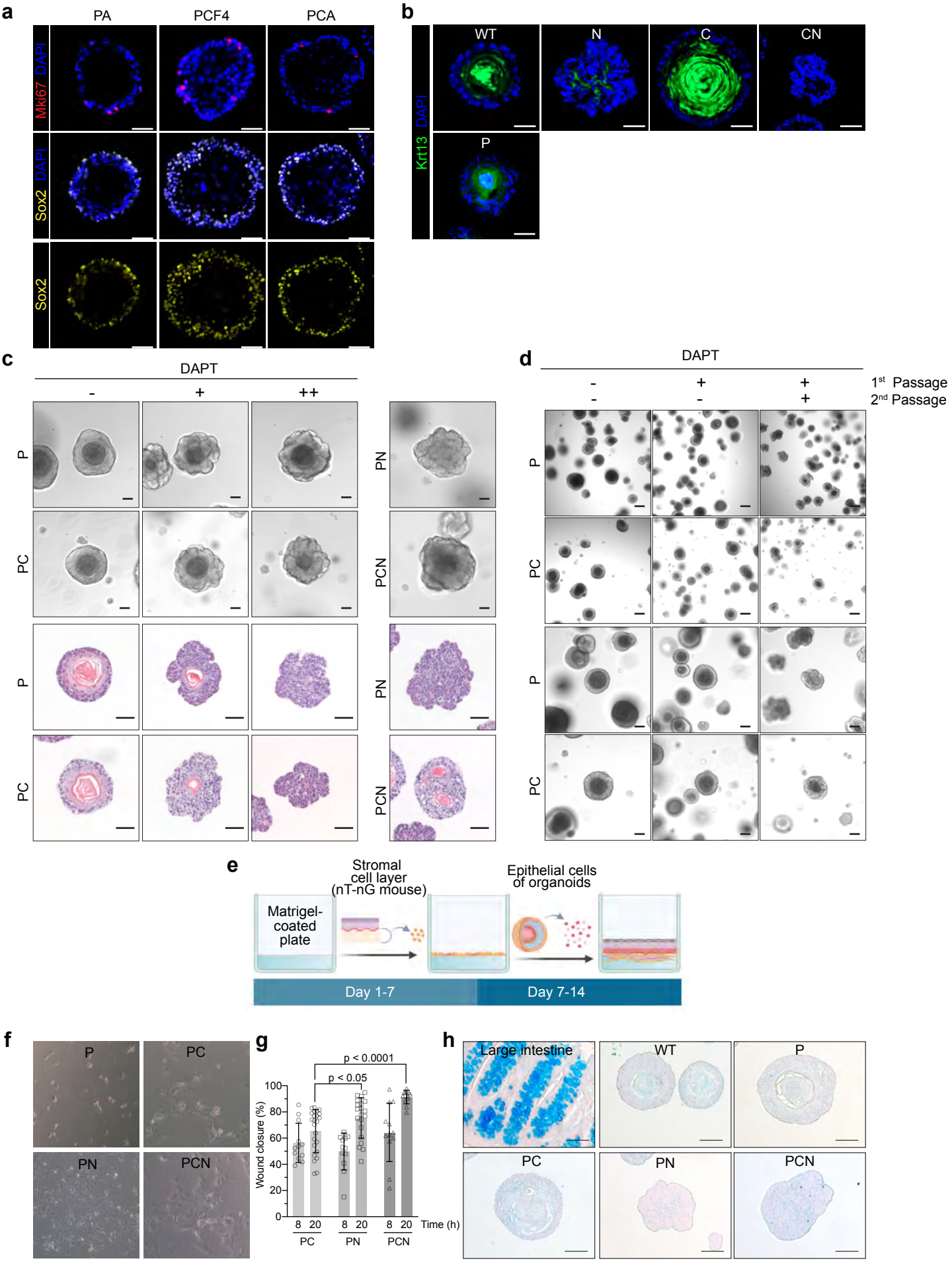
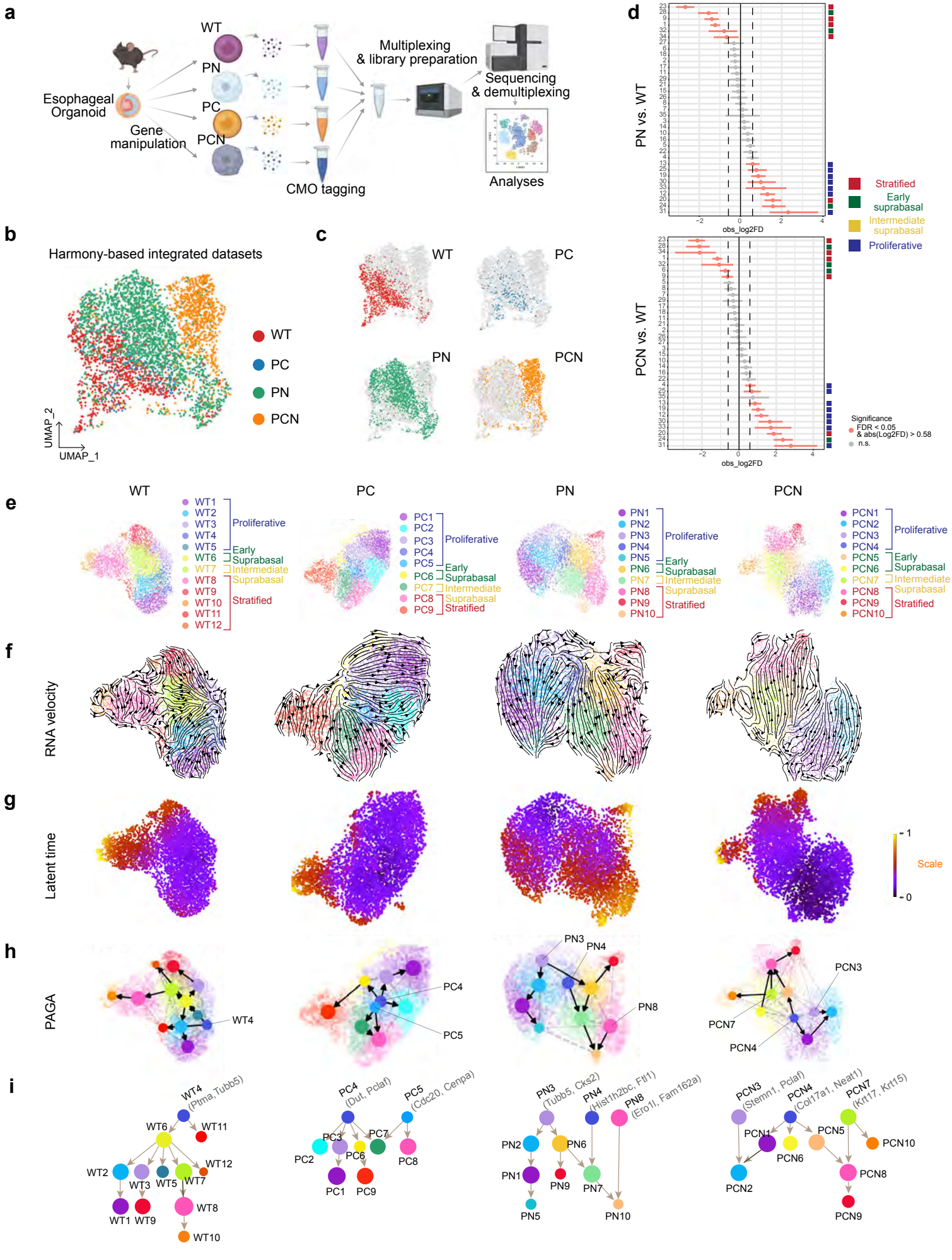


Figure 2



Extended Data 3

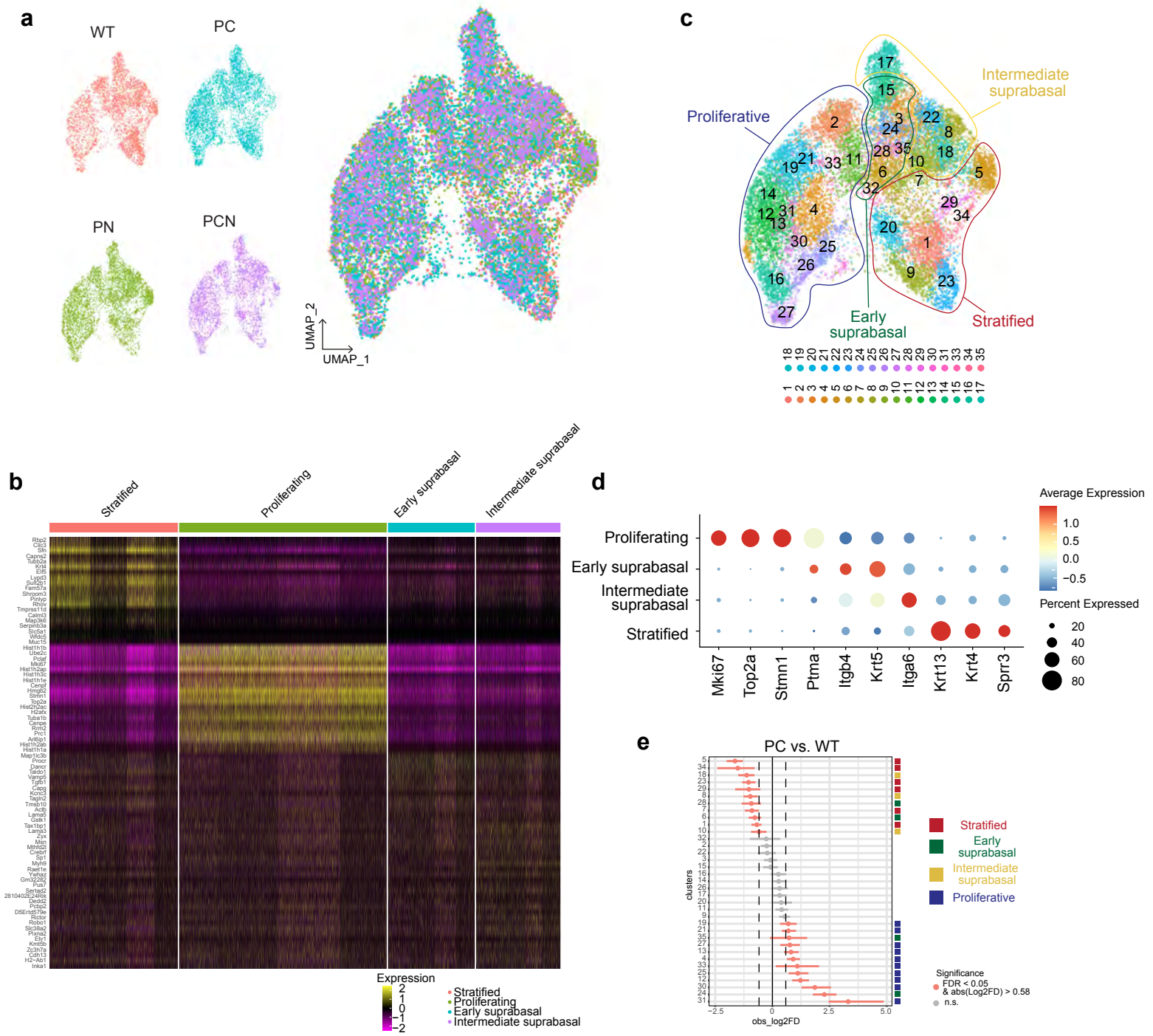


Figure 3

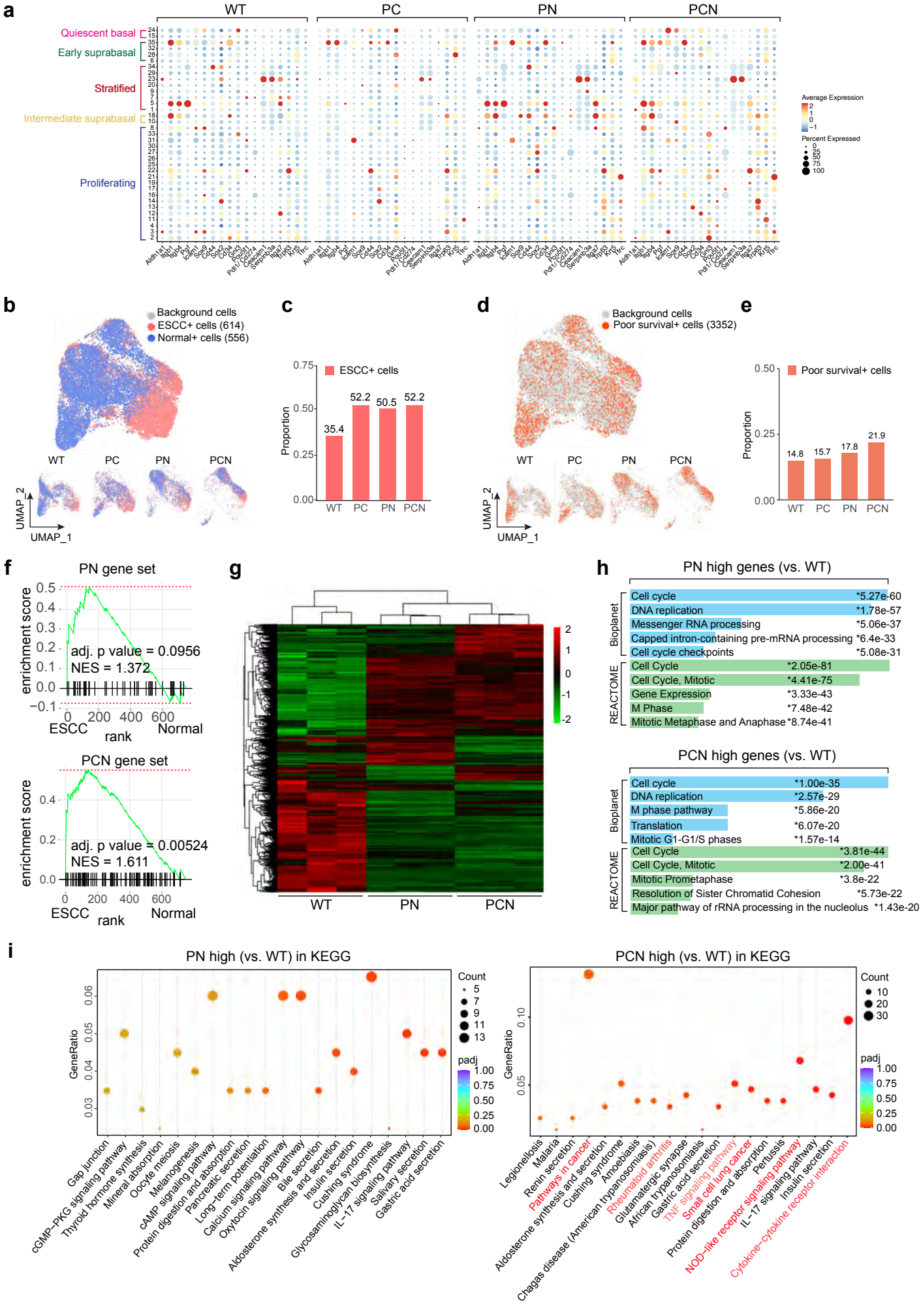
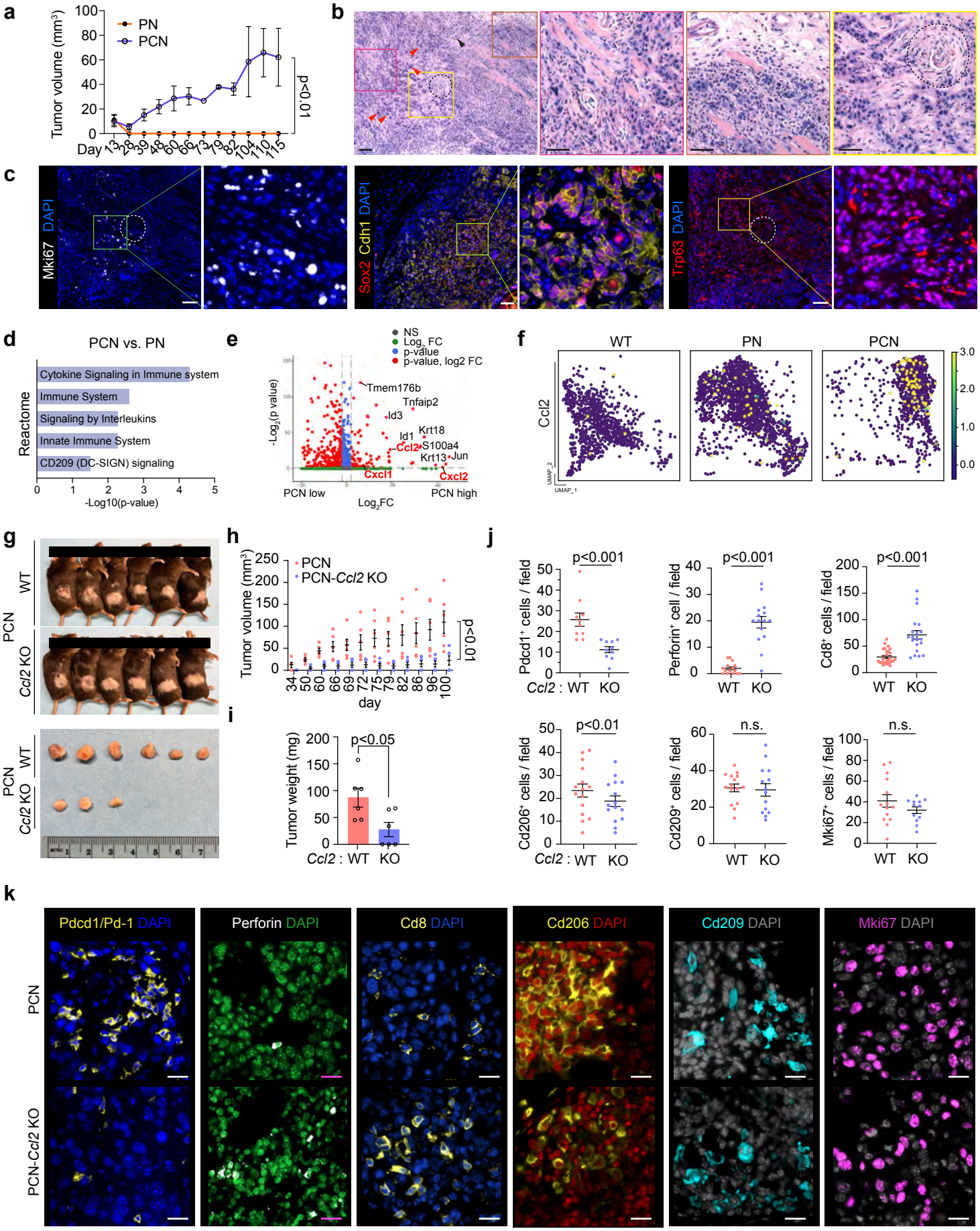
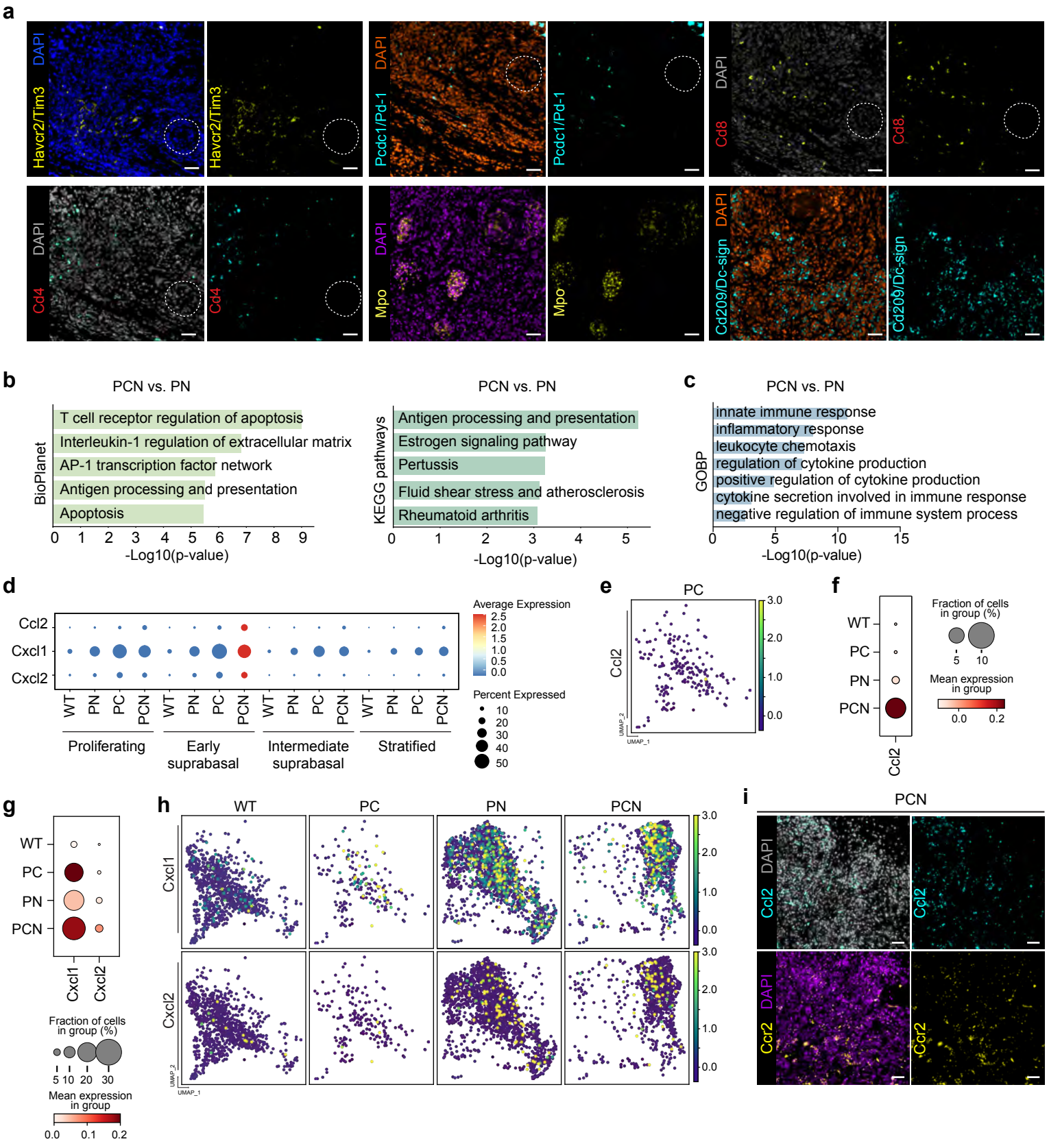


Figure 4



Extended Data 4



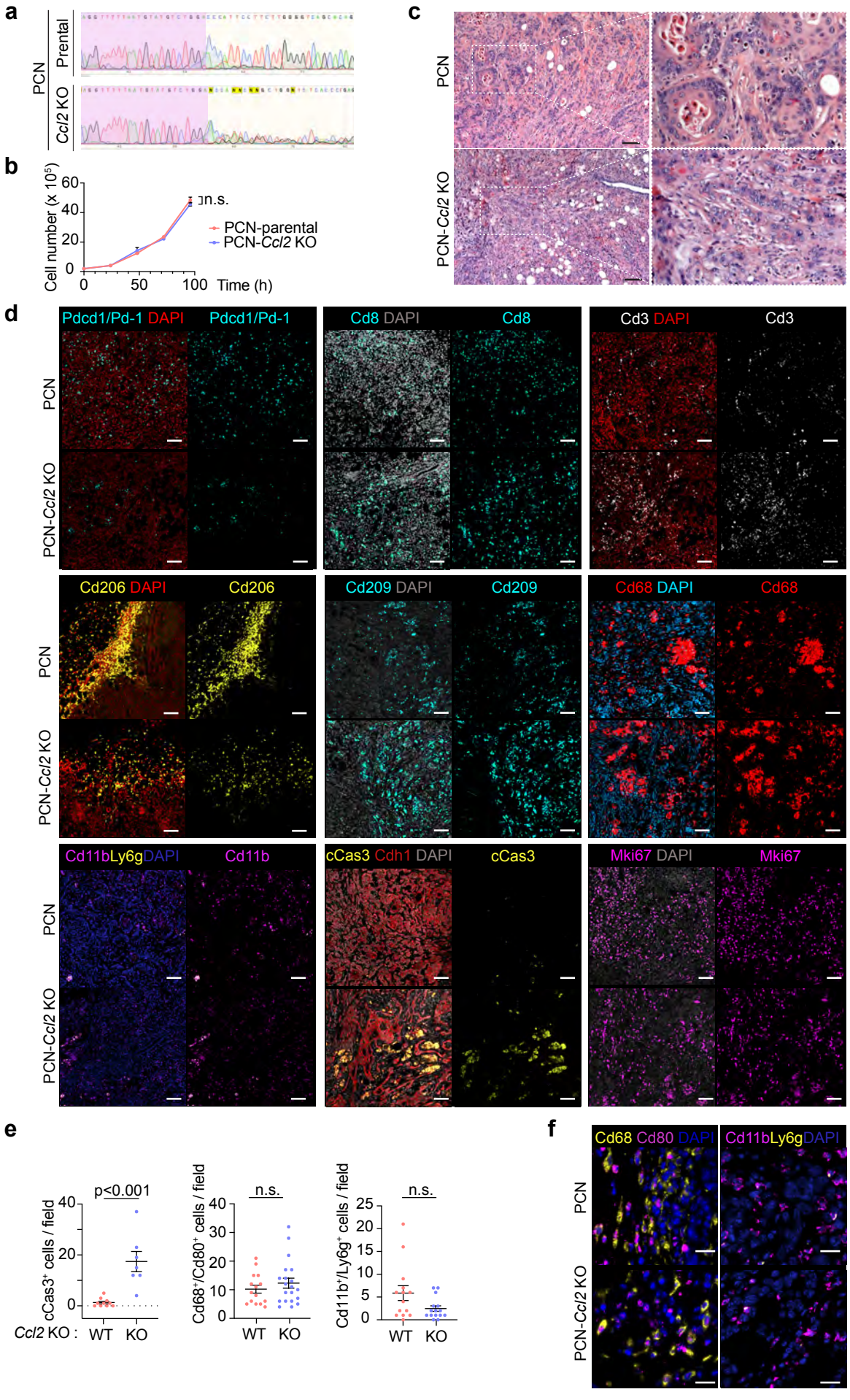
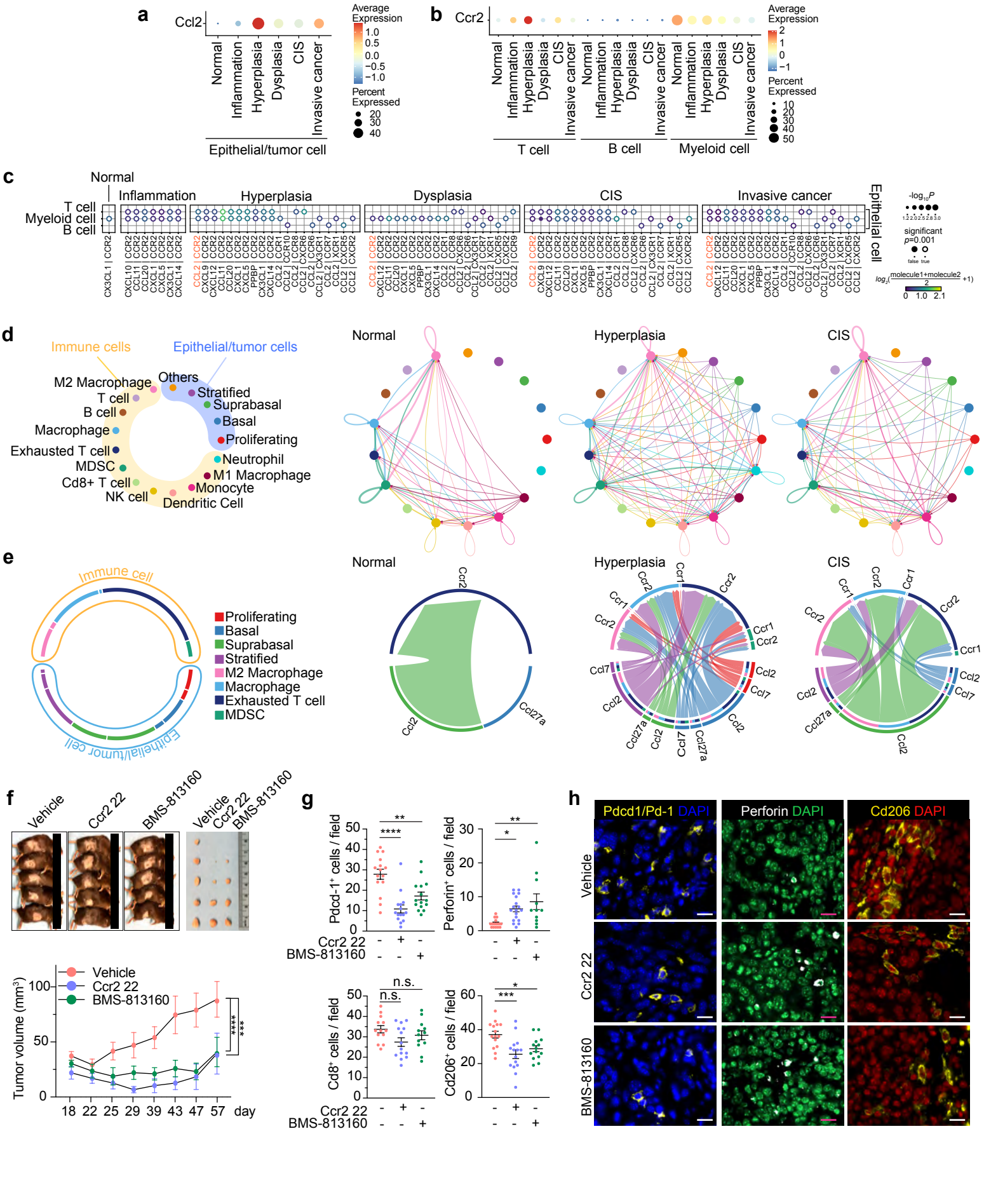
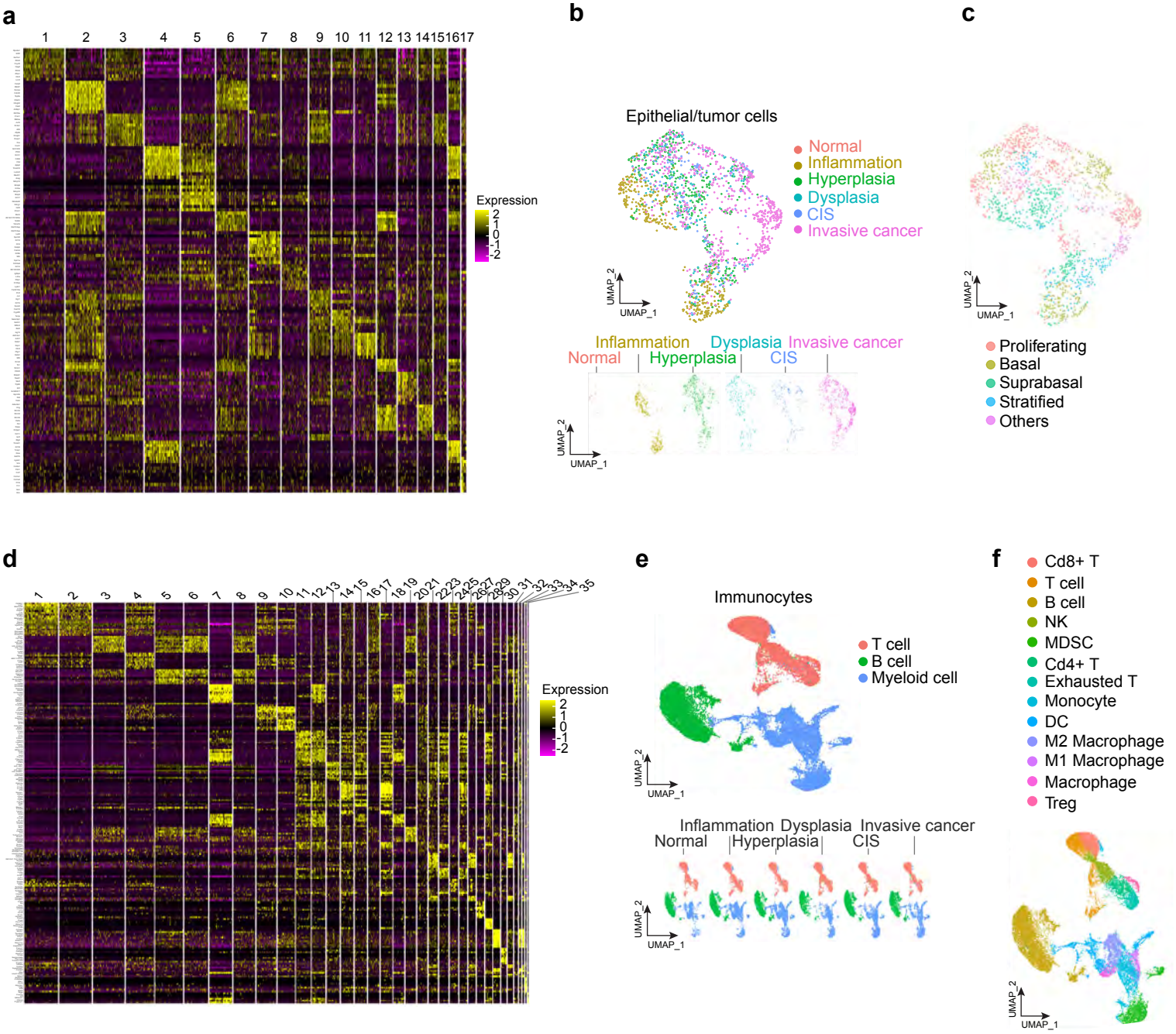


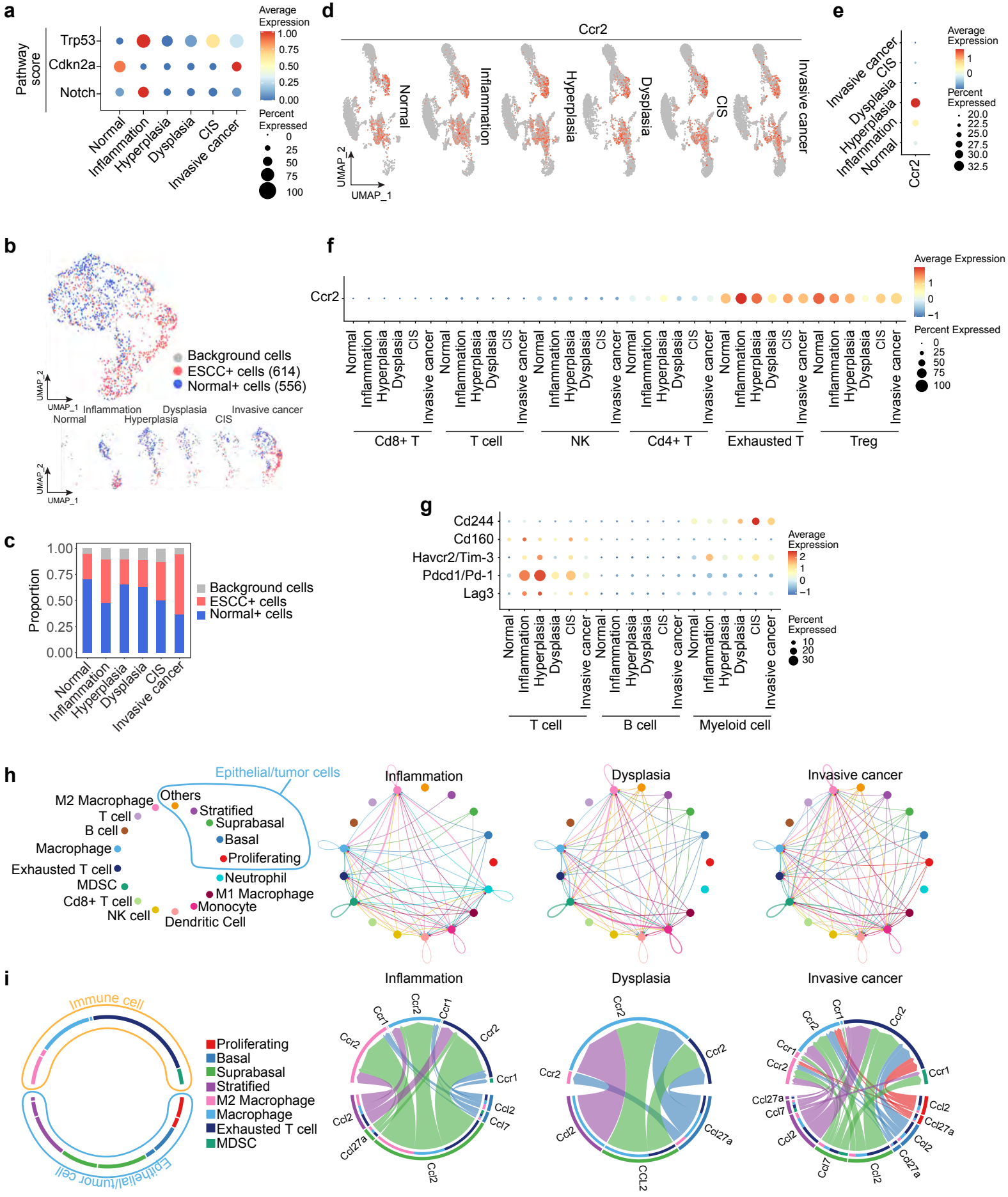
Figure 5



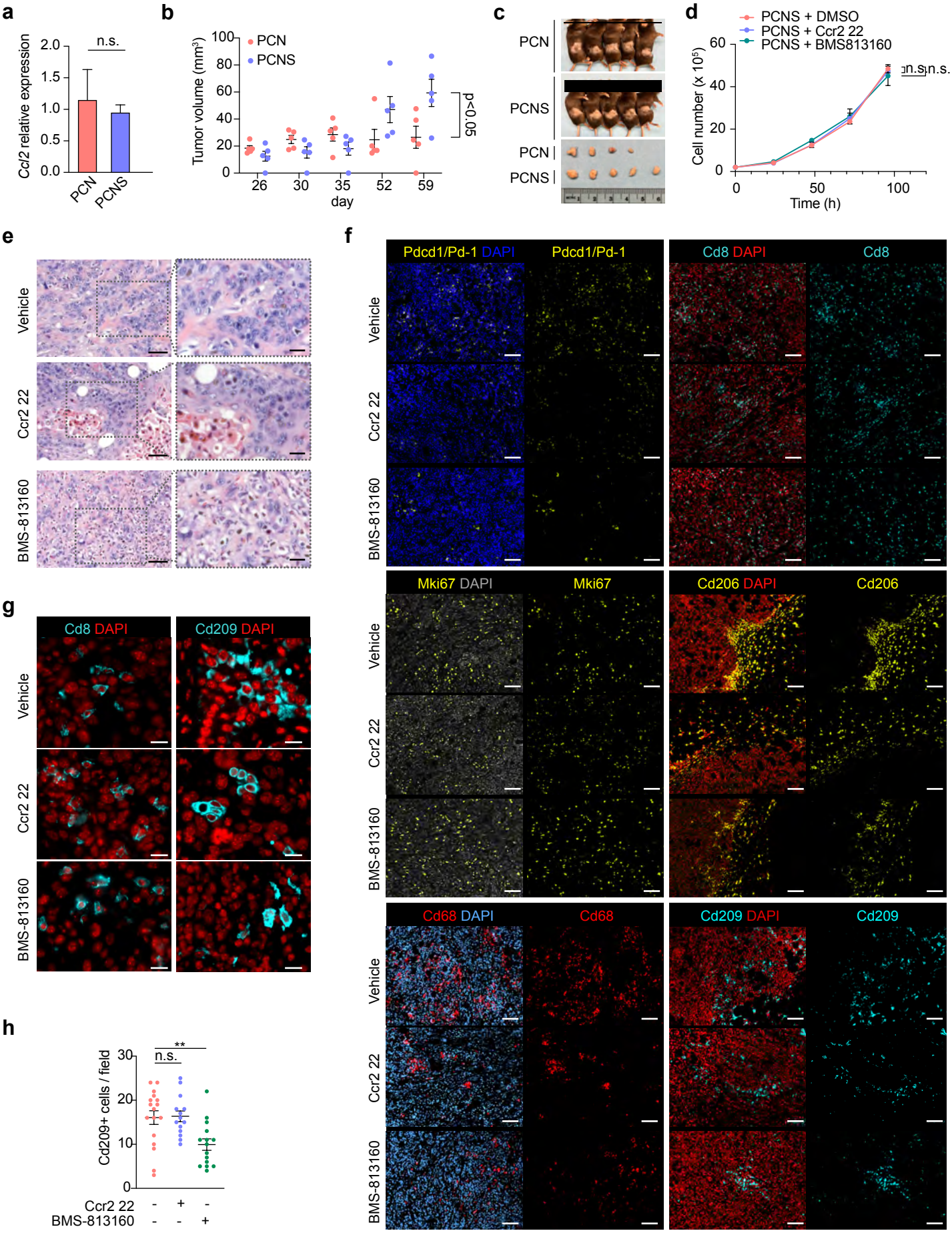
Extended Data 6



Extended Data 7



Extended Data 8



Extended Data 9

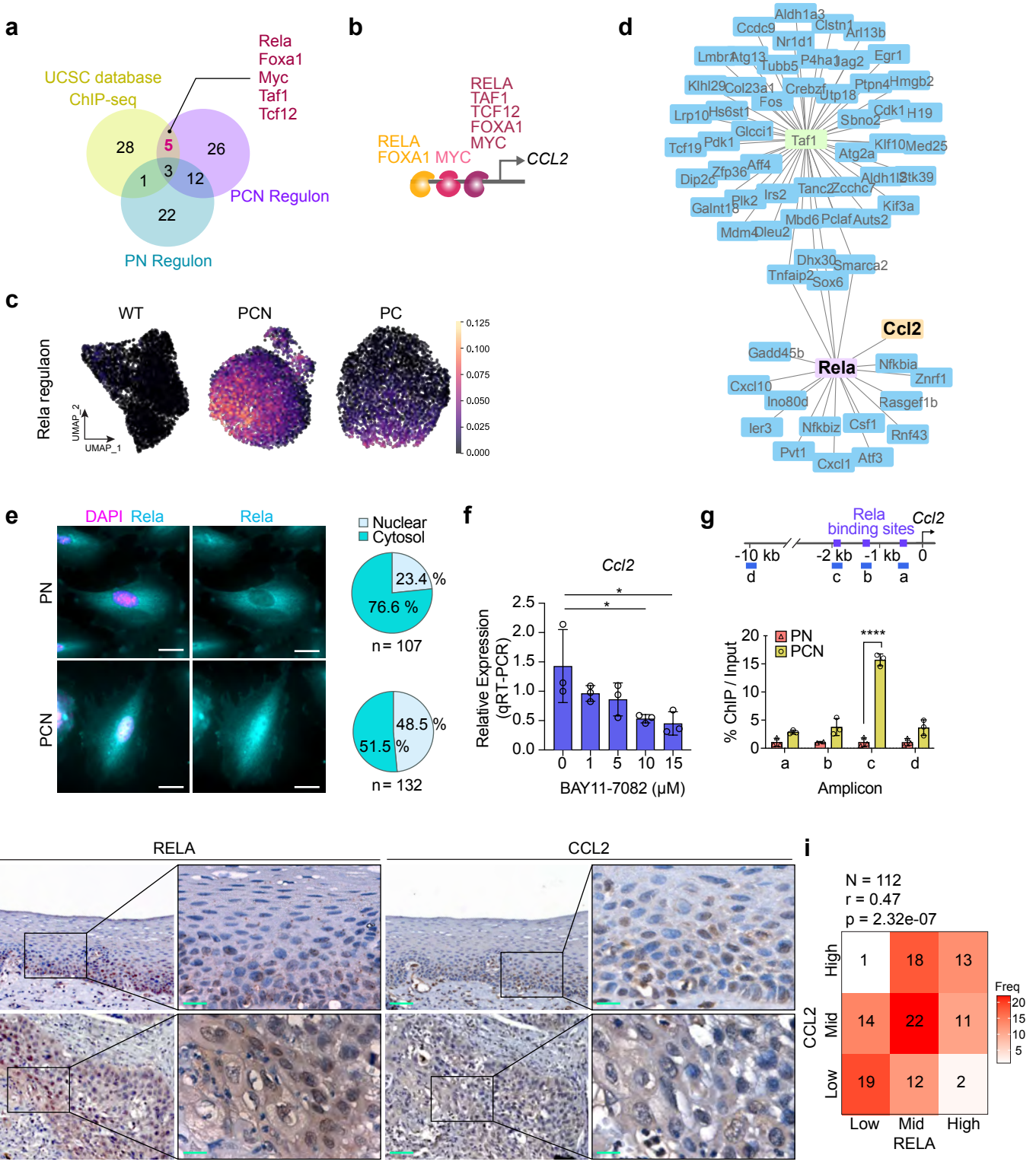


Figure 6