RESEARCH

Dissecting small cell carcinoma of the esophagus ecosystem by single-cell transcriptomic analysis

Hao-Xiang Wu[1,2,3,7†], Yu-Kun Chen[1,2,4,7†], Ying-Nan Wang[1,2,7†], Jia-Ying Chen[1,2,7†], Shu-Jing Xiang[1,2,4], Ying Jin[1,2], Zi-Xian Wang[1,2], Chun-Yu Huang[5], Lu-Ping Yang[5], Ye He[1,2], Wen-Long Guan[1,2], Long Bai[6], Yan-Xing Chen[1,2,4], Min Wang[1,2], Chao-Ye Wang[1,2,4], Run-Jie Huang[1,2,4], Yue Huang[1,2], Jin-Ling Zhang[1,2], Zhi-Da Lv[1,2,3], Si-Qi Yang[1,2], Rui-Hua Xu[2,7,8*], Qi Zhao[2,4,7,8*] and Feng Wang[2,7,8*]

Abstract

Small cell carcinoma of the esophagus (SCCE) is an aggressive and rare neuroendocrine malignancy with poor prognosis. Here, we firstly performed single-cell transcriptional profiling derived from 10 SCCE patients, with normal esophageal mucosa, adjacent non-malignant tissue and tumors from esophageal squamous cell carcinoma (ESCC) as reference. We observed enrichment of activated regulatory T cells and an angiogenesis-induced niche existed in SCCE compared with ESCC, revealing an immune suppressive and vessel-induced tumor microenvironment (TME) in SCCE. Totally, we identified five TME ecotypes (EC1 ~ 5). Notably, EC1 was highly enriched in SCCE, associating with molecular subtyping and survival outcomes. To dissecting heterogeneity of epithelium in SCCE, we constructed eight transcriptional metaprograms (MPs) that underscored significant heterogeneity of SCCE. High expression of MP5 was linked to neuroendocrine phenotype and poor clinical survival. Collectively, these results, for the first time, systematically deciphered the TME and epithelial heterogeneity of SCCE and provided evidences that SCCE patients might benefit from anti-angiogenesis therapy.

[†]Hao-Xiang Wu, Yu-Kun Chen, Ying-Nan Wang and Jia-Ying Chen contributed equally to this work.

*Correspondence:
Rui-Hua Xu
xurh@sysucc.org.cn
Qi Zhao
zhaoqi@sysucc.org.cn
Feng Wang
wangfeng@sysucc.org.cn

[1] Department of Medical Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, China
[2] Research Unit of Precision Diagnosis and Treatment for Gastrointestinal Cancer, Chinese Academy of Medical Sciences, Guangzhou 510060, People's Republic of China
[3] Department of Clinical Research, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, China

[4] Bioinformatics Platform, Department of Experimental Research, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, People's Republic of China
[5] Department of Endoscopy, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, China
[6] VIP Inpatient Department, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, China
[7] Guangdong Esophageal Cancer Institute, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dongfeng Road East, Guangzhou 510060, China
[8] Department of Medical Oncology, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-Sen University Cancer Center, Sun Yat-Sen University, Guangzhou, People's Republic of China

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 2 of 19

Introduction

Esophageal cancer is the sixth most common causes of cancer-related death worldwide in 2020, especially in southern and south-central Asia, accounting for approximately 544,000 deaths in the same year [1]. Small cell carcinoma of the esophagus (SCCE) is the most common extrapulmonary small cell carcinoma, accounting for 0.8%–2.4% of all esophageal carcinomas [2–5]. Most patients are frequently diagnosed with SCCE at late stage with positive neuroendocrine (NE) markers and die within 2 years, experiencing a median survival of only 8–13 months [6, 7].

To enhance our understanding of the molecular patterns of SCCE, we further performed a comprehensive genomic profiling of SCCE for the first time in 2018, which revealed key genomic alterations and offered insights into its pathogenesis. Additionally, we demonstrated the genomic similarities between SCCE and esophageal squamous cell carcinoma (ESCC) [8]. We also used ultra-deep whole transcriptome sequencing to analyze the expression profile of nine paired SCCE samples and compared the transcriptome with public transcriptomic data sets of normal esophageal mucosa and other cancer types. This study further characterized the immune microenvironment in SCCE and provided evidence that several patients with SCCE might benefit from immune checkpoint blockade therapy [9]. However, a lack of understanding regarding the ecosystem of SCCE continues to pose significant challenges in managing this deadly disease and hinders the development of improved treatment strategies.

To further define the heterogeneity of tumors and their associated ecosystem, in this study, we performed single-cell transcriptional profiling coupled with T cell receptor (TCR) sequencing of tumor samples from 10 SCCE patients, comparing to single-cell transcriptomes of tumors and adjacent nonmalignant tissue from 11 ESCC patients. This study, for the first time, elucidates the complete TME of SCCE, dissecting the heterogeneity of epithelial cells on single-cell resolution and provides the evidence that application of anti-angiogenesis therapy may benefit patients with SCCE.

Result

High resolution single-cell landscape of SCCE

To elucidate heterogenic cellular compositions between SCCE and ESCC, we applied scRNA-seq analysis to 10 tumor samples collected from patients diagnosed with SCCE in SYSUCC, combining 21 samples with published scRNA-seq data from normal esophageal mucosa, ESCC and matched nonmalignant adjacent tissue (NAT) (Fig. 1A, S1 A- 1B; Table S1). All the patients with SCCE enrolled from SYSUCC were male with tumor lesions on

mid or lower thorax (100%), ranging from 55 ~ 75 years old (Figure S1D). Nearly half of the patients had wine and cigarette addiction. To avoid any artificial bias, all the samples were collected by experienced endoscopists with multiple biopsies, among which we also performed single-cell V(D)J profiling of T cells on seven samples. To validate the major findings of our study, we also collected an independent cohort with bulk RNA-seq data ($n = 38$) [10].

After initial quality filtration, we obtained 197,858 high-quality single cells from four distinct histological subtypes and classified them into ten clusters with canonical marker expression, including B cell, dendritic cell (DC), endothelial cell, epithelial cell, fibroblast, mast cell, myeloid cell, neutrophil, plasma cell, and T/NK cell (Fig. 1B, S1 C). To better understanding the heterogeneity of TME, we analyzed cellular compositions among different tissue subtypes and found that epithelial cells were predominantly enriched in normal epithelial mucosa with few immune and stromal compositions. Between ESCC and SCCE, epithelial cells, myeloid cells and neutrophils mainly distributed in SCCE compared to ESCC, while B cell, T cell, mast cell and fibroblasts were enriched in ESCC (Fig. 1C-D, S1E-F). Together, these observations indicated an immune desert phenotype in SCCE.

Suppressive tumor microenvironment formation in SCCE

T cells play a vital role in anti-tumor immunity. Analyzing the function and composition of the T cells can help us predict the response of immune check point inhibitor (ICI) and explore other therapeutic strategies. A comprehensive atlas of T/NK cells was constructed on 41,906 cells, which was divided into four major clusters, including CD4+T, CD8+T, NK, NKT (Fig. 2A, S2 A-C). Among CD8+T, we identified eight subclusters according to classical cell markers and established functional signatures, including two clusters of effector memory T cells (Tem, c08, 10), two clusters of effector T cells (Teff, c11 - 12), stress-related T cells (Tstr, c09), proliferating T cells (Tpro, c15). Among them, CD8_c08_Tem (ANXA1) and CD8_c11_Teff (CRTAM) showed high enrichment in NAT of ESCC (Fig. 2B). We also identified a cluster of CD8 T cell featured by CXCL13, highly expressing immune check point molecules, such as CTLA4, LAG3, LGALS1, HAVCR2 and a subset of resident memory T cells (Trm, c13) characterized by high expression of GZMK, tissue resident markers (Fig. 2C, S2 F). Consist with biomarker expression, CD8_c14_Tex (CXCL13) showed the highest T cell dysfunction score, indicating an exhausted state (Fig. 2F).

In order to explore evolutionary relationship between each CD8 +T clusters, we applied slingshot to all

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 3 of 19
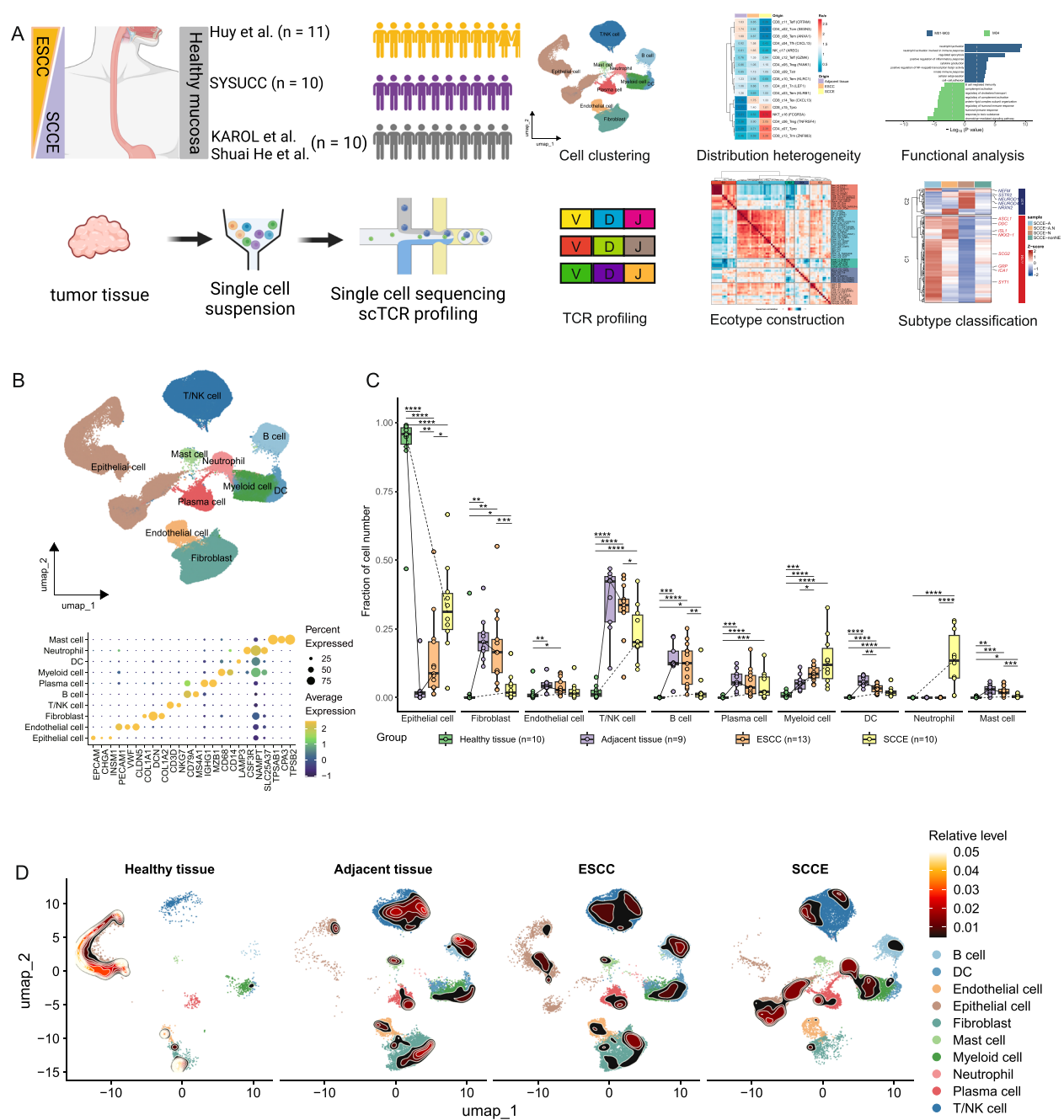
Fig. 1 Single cell landscape of immune, stroma and epithelial cells of esophageal comparative cohorts. **A** Scheme of the overall study design. **B** UMAP visualization of the major cell clusters including 197,858 high-quality single cells from 42 samples collected from 31 candidates (10 HCs, 11 pts with ESCC, 10 pts with SCCE). **C** Box and whisker plots showing the fraction of cell types originating from four groups in each major cell cluster with plot center, box and whiskers corresponding to median, interquartile range (IQR) and 1.5 × IQR, respectively (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$). **D** UMAP plots, as shown in (B), showing TME cell clusters and cell distribution density (shown as thermodynamic chart) across different tissue groups

the CD8 T cells and identified two distinct trajectories (Fig. 2D). Both of them started at CD8_c08_Tem (ANXA1), followed by CD8_c09_Tstr, CD8_c10_Tem (KLRC1), and branched at CD8_c13_Trm (ZNF683). One

branch turned to exhausted state and terminated at CD8_c14_Tex (CXCL13), while the other one sustained cytotoxicity and differentiated into effector T cells, including CD8_c12_Teff (GZMK) and CD8_c11_Teff (CRTAM). As

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 4 of 19

expected, two esophageal carcinomas, ESCC and SCCE, both aligned with branch one harboring an exhausted T cell state, while NAT of ESCC with branch two sustaining cytotoxicity during evolution (Fig. 2E-F). The evolutionary trajectory in SCCE was validated via TCR-seq analysis. CD8_CXCL13 T cells had high proportion of clone expansion in SCCE (Figure S2D). In clone overlap analysis, CD8_Trm_ZNF683 showed higher clone overlaps with CD8_c14_Tex (CXCL13) than CD8_c12_Teff (GZMK), suggesting an exhausted evolutionary lineage in SCCE (Figure S2E). Nevertheless, CD8_c14_Tex (CXCL13) showed higher enrichment in ESCC than SCCE. As previously reported, tumor-reactive CXCL13 +T cells can response to ICI therapy during anti-cancer immunity, for which CXCL13 +T cells were correlated with favorable responses to ICI therapy [11]. We then examined immune check point expression in four tissue subtypes and found that ESCC had much higher immune check point expression than SCCE relatively, indicating that ESCC might benefit more from ICI therapy than SCCE in clinical practice (Fig. 2G, S2G).

Among CD4+T cell, we identified seven subclusters, including a clusters of naïve T cell (Tn, c01); central memory T cell (Tcm, c02); effector memory T cell (Tem, c03); follicular helper T cell (Tfh, c04); proliferating T cell (Tpro, c07) (Fig. 2A). Notably, we discovered two clusters of regulatory T cells (Treg, c05 - 06), one of which highly expressed FOXP3, IKZF2, IL2RA and costimulatory molecules preferentially distributing in SCCE. Next, we subclustered all the Treg into five subtypes and applied slingshot to explore the evolution trajectory (Fig. 2H). Treg_c1 highly expressed naïve markers (such as LEF1, CCR7) locating at the beginning of two evolutionary branches. Treg_c4 was characterized by high expression of HLA molecules recognized as transitional state, locating at the center of the trajectory. And then one branch turned to Treg_c2, which highly expressed FOXP1, a critical regulator of Treg cell quiescence and homeostasis. The other branch went into Treg_c5 characterized by high expression of FOXP3, CTLA4, and co-stimulatory molecules and became activated (Fig. 2I-J). Interestingly, we found that transitional and activated Treg highly enriched in SCCE than ESCC, suggesting that Treg contributed a lot to suppressive TME formation in SCCE (Fig. 2K-L).

Taken together, both esophageal carcinomas showed a suppressive TME with distinct characteristics. ESCC showed significant enrichment of exhausted CD8+ T cell with high immune molecule checkpoint expression, while SCCE enriched more activated Treg cells.

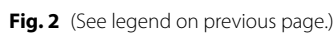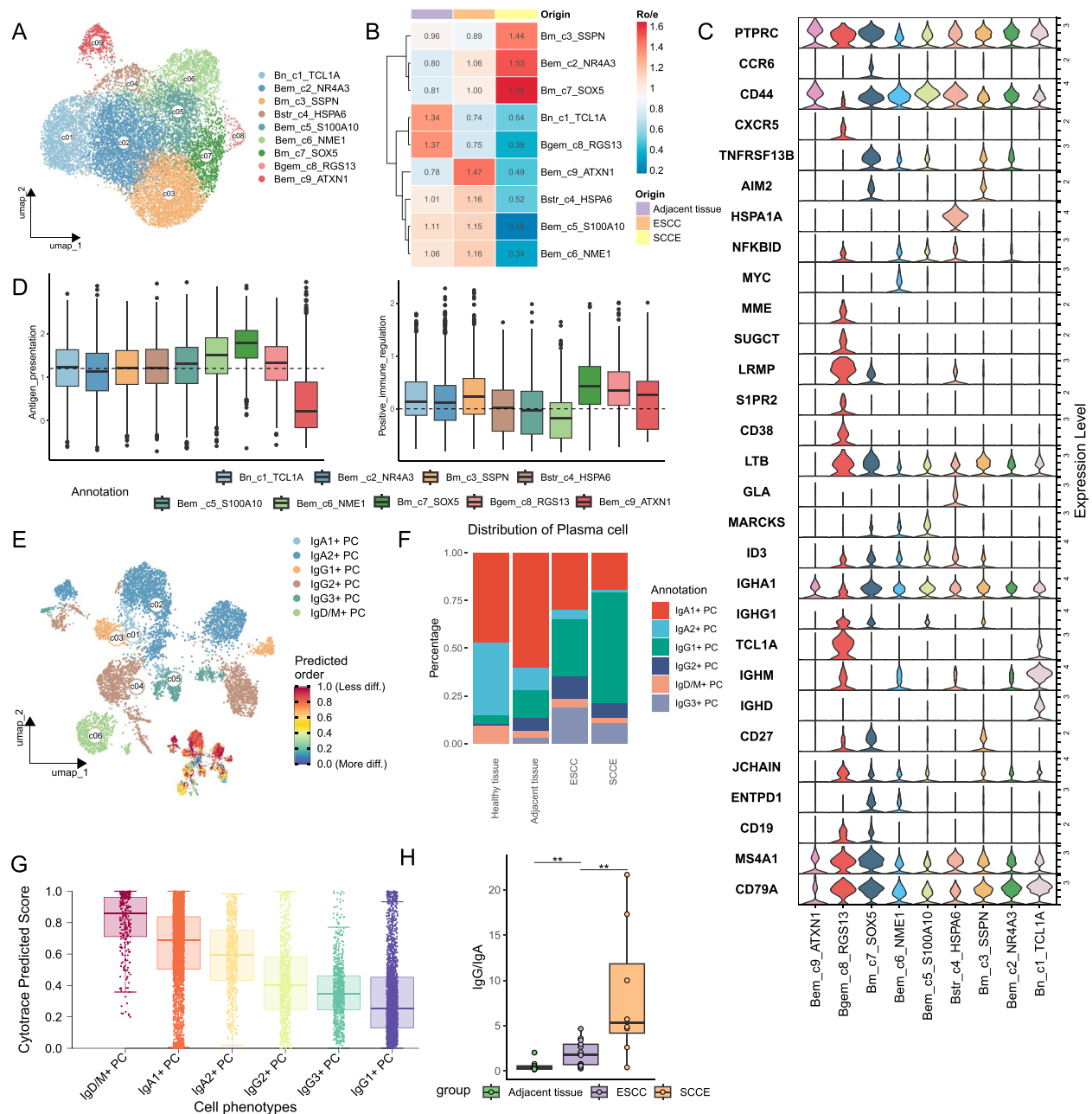## Highly differentiated B and plasma cells in SCCEs

To further dissect cellular compositions of B-lineage cells in SCCEs, we divided all the B cells into nine subclusters (Fig. 3A-B, S3 A). We identified a cluster of naïve B cell (Bn) featured by expression of TCL1 A, which mainly enriched in NAT of ESCC; stress-related B cell (Bstr) expressing HSPA6; germinal center B cell (Bgem) characterized by RGS13, SUGCT, MME etc. Memory B cells were identified by CD45RA and CD27. CD45RA +CD27- B cells were taken as early memory B cells (Bem), while CD45RA +CD27 +B cells (such as B_c3_SSPN, B_c7_ SOX5) were considered as late memory B cells (Bm), showing high positive immune regulation and antigen presentation score (Fig. 3C-D). Interesting, the abundance and composition of B cells varied greatly between SCCE and ESCC. Late memory B cells showed high enrichment in SCCE, while early memory B cells in NAT and ESCC (Figure S3B). The same distribution tendency had been also found in plasma cell (Fig. 3E, S3D). All the plasma cells were divided into six subclusters according to their secreting antibody (Fig. 3E-F, S3 C). The cellular differentiation state was inferred by Cytotrace analysis, suggesting IgD/M +PCs were the least differentiated, followed by lgA +PCs and IgG +PCs. IgG1 +PCs had the lowest cytotrace predicted scores, indicating the highest differentiated state (Fig. 3G). In normal gastrointestinal and respiratory tissue, lgA antibodies contribute to maintenance of mucosal immunity. However, during tumorigenesis, lg class switching to IgG predominantly occurs in tumor region, initiating antibody mediated immune response [12]. To further quantify the lg switch process, we calculated lgG/IgA score among different tissues, and found that SCCE had the highest differentiated plasma cells, indicating a more active B cell mediated immune response in SCCE. (Fig. 3H, S3E).

(See figure on next page.)

Fig. 2 T cell characterization across different tissue groups in esophageal comparative cohorts. **A** UMAP visualization of T cells separated into 17 subsets. **B** Heatmap displaying the tissue distribution preference of T and NK cells through ratio of observed to expected cell numbers (Ro/e). **C** Bubble heatmap displaying expression levels of selected markers in T and NK cells. **D** Evolutionary lineage of CD8 T cells inferred by slingshot. **E** The same as D) but displayed in different tissue origin facets. **F** Expression of cytotoxic and dysfunction signatures among CD8 T cell subsets. **G** Relative expression of immune check points in CD8, CD4, NK subsets among different tissue origins, shown in heatmap. **H** UMAP visualization of Treg cells separated into 5 subsets. **I** Evolutionary lineage of Treg cells inferred by slingshot. **J** Expression levels of selected markers in Treg cells. **K** Boxplot comparison of Treg cell clusters among different tissue origins. **L** The composition of Treg cells among different tissues, shown in stack plot

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 5 of 19

Fig. 2 (See legend on previous page.)

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 6 of 19

Fig. 3 B and plasma cells characterization across different tissue groups in esophageal comparative cohorts. **A** UMAP visualization of B cells separated into 9 subsets. **B** Heatmap displaying the tissue distribution preference of B cells through ratio of observed to expected cell numbers (Ro/e). **C** Expression levels of selected markers in B cells, shown in violin plot. **D** Expression of antigen presentation and positive immune regulation signatures among B cell subsets. **E** UMAP visualization of Plasma cells separated into 6 subsets. **F** The composition of different plasma cell clusters among different tissues, shown in stack plot. **G** Cell differentiation inferred by Cytotrace software. High cytotrace predicted score suggested that the cellular clusters have high stemness. **H** The comparison of IgG/IgA between different tissue origins

Vessel-induced niche formation in SCCEs

Myeloid cells were predominantly distributed in SCCE. In addition to DC and mast cell, we discovered five clusters of macrophages with distinct function (Fig. 4A). Among them, Macro_c01_FCN1 and Macro_c03_HSPH1

showed much more enrichment in SCCE with high angiogenesis-related markers and signatures, while Macro_c04_LYVE1 enriched in ESCC with high phagocytosis feature (Fig. 4B, S4 A-B). Next, we explored functional state between two different tissue-derived macrophages,

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 7 of 19

and discovered that macrophages from SCCE highly expressed S100 A8, 9, 12, FCN1, VCAN and IL1B, which corelated to angiogenic phenotype and aggressive characteristics of macrophages [13]. Nevertheless, macrophages in ESCC upregulated HLA molecules, CXCL1/3, C1Q and APOE showing that macrophages in ESCC might played an important role in antigen presentation and T cell recruitment (Fig. 4C). Pathway analysis also revealed significant enrichment of neutrophil activation and NFκB signaling pathway in SCCE, whereas complement activation and lipid metabolism-related pathway enriched in ESCC (Figure S4 C). We further explored potential origin of macrophages from two tissue lineages through calculating tissue-resident score [14, 15]. Macro_c04_LYVE1 got the highest ranks among three macrophages, indicating that macrophages of ESCC might originate from tissue residence. (Figure S4D).

For nonimmune cell, we first dissected the endothelial cells into two major clusters, including vessel and lymph endothelial cells. The endothelial cells were further clustered into seven subclusters according to their topmarkers, among which Endo_c2_CACNA1 C and Endo_c3_ESM1 were mainly enriched in SCCE (Fig. 4D-E). Moreover, Endo_c3_ESM1 was featured by ESM1 with upregulated high Tip-like signature (Fig. 4F, S4E). ESM1 was reported as a target and modulator of VEGF signaling in endothelial cells, participating in angiogenesis, inflammation, and vascular permeability [16–19]. Then, Further, we performed differential expression analysis between endothelial clusters enriched in SCCE and ESCC. CACNA1 C+ E02 and ESM1 + E03 expressed high levels of PGF, CCL2, CCL14, CXCL1, MMP1, AKR1 C, involved in angiogenesis, inflammation, matrix remodeling and metabolic reorganization [20]. In contrast, FGF7 + E04 highly expressed PIEZO2, a mechanically activated cation channels, suggesting that endothelial cells in SCCE might contribute more to formation of TME with the highest vessel disorganization
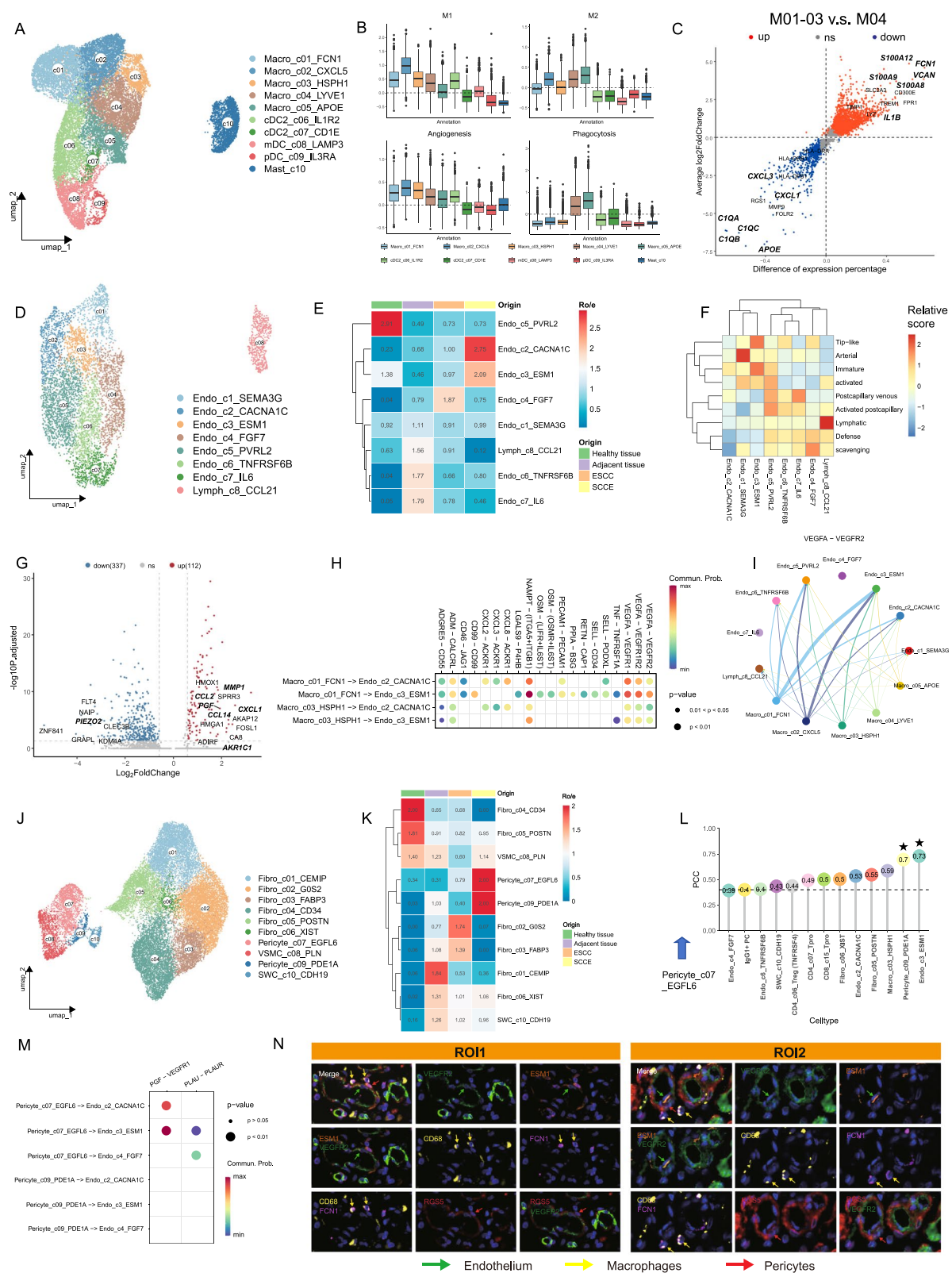
scores among all the tissue lineages (Fig. 4G, S4H). In the contrast, lymph endothelial cells were mainly distributed in ESCC (Figure S4 F). Pathway analysis also revealed that lymphangiogenesis and lymph vessel development pathways were more enriched in ESCC (Figure S4G). Cell–cell communication analysis suggested specific interactions between macrophage subclusters and endothelial cells were mainly related to angiogenesis signaling. Therein, FCN1 + macrophages had strong cross talk with ESM1 + endothelial cells via VEGF-VEGFR interaction (Fig. 4H-I). Focusing on ligand-receptor pair expression, we found classical angiogenic receptor, KDR, ITGB1, showed the highest expression level in Endo_ESM1, and multiple pairs including, KDR-VEGFA, FGF1-FGFR1, IGF2-IGF2R mediate crosstalk between Macro_c01 ‑ 03 and Endo_c2 ‑ 3 (Figure S5 A-B). As reported recently, cross talk of macrophages and endothelial cells in TME played an important role in angiogenesis, along with matrix remodeling and stabilization of neo-vessel by pericytes [21, 22].

As for stromal cells, we discovered two clusters of pericytes highly expressing RGS5 enriched in SCCE, and significantly upregulated focal adhesion and cell junction pathway (Fig. 4J-K, S4I-J). Additionally, we found that frequency of EGFL6 + pericyte_c07 positively correlated with proportions of ESM1 + Endo_c3, indicating a close connection among matrix cellular compositions (Fig. 4L). Consistently, EGFL6 + pericytes showed close interaction with ESM1 + endothelial cells through PGF-VEGFR1 and PLAU-PLAUR signaling pathway (Fig. 4M, S5 C-F).

In addition, neutrophils were exclusively enriched in SCCE, which were dissected into six subclusters and three subclusters were found highly enriched in inflammation related pathways with relatively high differentiation scores, including TNF-signaling, interferon alpha/gamma response (Figure S6 A-D). The results indicated that neutrophils that preferred to accumulate in SCCE might greatly contribute to the local inflammation.

(See figure on next page.)

Fig. 4 Myeloid and stromal characterization across different tissue groups in esophageal comparative cohorts. **A** UMAP visualization of myeloid cells separated into 10 subsets. **B** Expression of M1, M2, angiogenesis, phagocytosis signatures among different myeloid subsets. **C** Differentiated expressed genes between M01-M03 enriched in SCCE and M04 enriched in ESCC. The x-axis represents the difference of cellular proportion between two originated macrophages, while y-axis refers to average log2 FoldChange of each gene. **D** UMAP visualization of endothelial cells separated into 8 subsets. **E** Heatmap displaying the tissue distribution preference of endothelial cells through ratio of observed to expected cell numbers (Ro/e). **F** Relative expression of endothelium-related functional signatures among different endothelium subsets, shown in heatmap. **G** Differentiated expressed genes between E02-E03 enriched in SCCE and E04 enriched in ESCC. The x-axis represents log2foldchange of each gene between two originated endothelial subsets, while y-axis refers to -log10P.adjusted. **H** Cell–cell communication probability between M01, 03 and E02, 03 was calculated by cellchat. **I** The VEGFA-VEGFR2 signaling pathway between macrophage and endothelial cell clusters, scaled by weight. **J** UMAP visualization of stromal cells separated into 10 subsets. **K** Heatmap displaying the tissue distribution preference of stromal cells through ratio of observed to expected cell numbers (Ro/e). **L** Correlation coefficient between cell proportions of Pericyte_c07_EGFL6 and other TME cell subsets. **M** Cell–cell communication probability between pericyte_c07, c09 and E02, 03, 04 was calculated by cellchat. **N** Two representative regions of interest in multicolor IHC staining of SCCE patient. Magnification 40X

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 8 of 19

Fig. 4 (See legend on previous page.)

Wu *et al. Molecular Cancer*　　　(2025) 24:142

Page 9 of 19

Considering that subclusters involved in angiogenesis mentioned above were all enriched in SCCE, we tried anti-angiogenic agents on SCCE patients in later-line treatment. Surprisingly, a SCCE patient suffered from disease progression for lung metastasis aggravation reached disease control after addition of anlotinib (Figure S7). To encapsulate our findings, we proposed a vessel-induced niche in SCCE formed by FCN1 + macrophages, ESM1 + endothelial cells, and RGS5 + pericytes. Immunofluorescence revealed a scene of neovascularization accompanied by perivascular angio-induced macrophage infiltration, providing a promising new target for SCCE therapy and deepening our understanding to biological properties of SCCE (Fig. 4N).

Cellular compositions and ecotypes of TME cells

After dissecting the cellular compositions among major clusters above, we next explored the similarity and functional relationship between 66 TME cell subset. Firstly, unsupervised hierarchical clustering was performed to quantify similarity between different cellular subsets. Myeloid and lymphatic cells combined and formed a major branch, while stromal cells formed the other. Notably, neutrophils predominantly enriched in SCCE, showing more similarity with T, B lymphocytes than myeloid cells (Fig. 5A). To better understand how these cell subsets orchestrated together to form cohesive ecotypes in the TME and to what extent the ecotypes contributed to clinical benefit, we performed unsupervised clustering analysis on pairwise spearman correlation coefficients between different cellular subsets. Our analysis identified five ecotypes (EC1-EC5) with different cellular composition and tissue distribution, showing transcriptional heterogeneity across different esophageal tissues (Fig. 5B). EC1 enriched in SCCE, which mainly contained tumor-associated neutrophils and T cell clusters, including all the neutrophil subsets, resident T cell (CD8_c13_Trm), exhausted T cell (CD8_c14_Tex), proliferating T cell, and NKT. Upregulation of neutrophil degranulation and cellular chemotaxis pathway indicated that EC1-dominant patients exhibited a neutrophil-activation state. Enrichment of mature B cell subsets, including B_c3_SSPN, B_c7_SOX2 and upregulation of complement activation pathway suggested humoral immune-response dominant phenotype of EC2, which mainly enriched in ESCC and related adjacent normal tissue. EC5 distributed in both ESCC and SCCE, while EC3, 4 scattered in different tissue lineages (Fig. 5C). Moreover, analysis of ecotype EC1 further revealed its correlation with pathological subtypes and clinical outcomes. In other SCCE cohort, EC1 showed positive correlation with better clinical benefit. Moreover, in SCCE-A subtype, the patients with high EC1 signature also had better overall survival
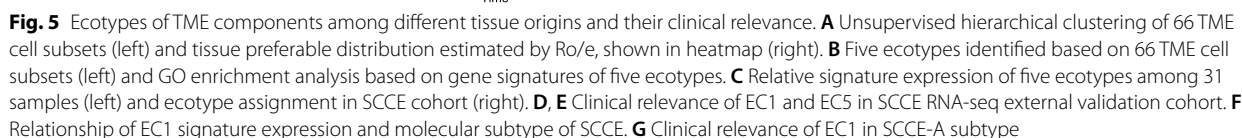
(Fig. 5D-G). As reported previously, inflammatory neutrophil response corelated with better clinical benefit in immuno-chemotherapy in SCLC [23]. Several inflammatory neutrophil subsets enriched in EC1, including Neu_c1_CCL3L1, Neu_c4_ELL2, Neu_c5_IFIT3, might perform anti-tumor function in TME.

Overall, single-cell based ecotypes provided us with a brand-new perspective towards the TME and predictive biomarkers in SCCE.

Dissecting heterogeneity of SCCE malignant cells

SCCE had higher proportions of epithelial cells compared to ESCC, showing an immune exclusive phenotype. To analyze the expression heterogeneity of malignant compositions among esophageal carcinomas, we firstly performed differential expression analysis and found epithelium of SCCE highly expressed NE markers, including NEUROD1, NEUROG3, GHRH, while epithelium of ESCC upregulated expression of growth factors, and markers of squamous epithelial cells (Figure S8 A). Consistently, pathway analysis revealed that NE, signal transduction as well as cell cycle related pathways were mainly enriched in SCCE, while keratinization and defense related signaling pathways were upregulated in ESCC (Figure S8B). Further, we compared the cell–cell communication between two esophageal carcinomas. The epithelium of SCCE had more interaction with immune compartment, including T/NK and myeloid cells. In the contrast, the epithelium of ESCC showed much more crosstalk between stromal components and immune cells (Figure S8 C-D).

To dissect the heterogeneity of epithelial cells of SCCE, we employed infercnv and calculated CNV_score of each epithelial cell. Then, k-means clustering was used to separated normal epithelial cells from malignant ones (Figure S8E-F). Among small cell NE carcinomas, advances had been made in molecular subtyping of SCLC based on transcriptional factors (TFs) and regulators [24]. Considering the cytomorphologic and transcriptional similarity of SCCE and SCLC, we examined the expression of several NE TFs, including ASCL1, NEUROD1/2, YAP1, POU2 F3, and two newly defined markers, HNF4G, FOXA3, in SCCE (Fig. 6A). Unsurprisingly, ASCL1 and NEUROD1 were exclusively expressed in epithelial cells of SCCE, consistent with that reported by Wang ZY et al. [25]. Due to the inherent sparsity of single cell data, we constructed 'pseudobulk' profiles from single cell data and applied the traditional TFs-based classification framework. Both the two methods, relative expression of NE transcription factors and unsupervised clustering, identified two molecular subtypes in SCCE, including SCCE-A (*n* = 9) and SCCE-N (*n* = 1) (Figure S9 A-B). To verify the subtyping results recalled by pseudobulk

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 10 of 19

Fig. 5 Ecotypes of TME components among different tissue origins and their clinical relevance. **A** Unsupervised hierarchical clustering of 66 TME cell subsets (left) and tissue preferable distribution estimated by Ro/e, shown in heatmap (right). **B** Five ecotypes identified based on 66 TME cell subsets (left) and GO enrichment analysis based on gene signatures of five ecotypes. **C** Relative signature expression of five ecotypes among 31 samples (left) and ecotype assignment in SCCE cohort (right). **D**, **E** Clinical relevance of EC1 and EC5 in SCCE RNA-seq external validation cohort. **F** Relationship of EC1 signature expression and molecular subtype of SCCE. **G** Clinical relevance of EC1 in SCCE-A subtype

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 11 of 19

analysis, we scored each epithelial cell with established A/N signatures and designed the overall transcriptional subtypes according to cellular proportions. The results totally agreed with that recalled by pseudobulk method (Figure S9 C-D). To further dissect transcriptional heterogeneity on single cell level, we subclustered all the malignant epithelial cells into four clusters according to the expression of ASCL1 and NEUROD1 (Figure S9E). The epithelial cells that expressed ASCL1 and NEUROD1 were considered to be epithelium with neuroendocrine phenotype (SCCE-NE), among which the epithelial cells expressing both ASCL1 and NEUROD1 were assigned as SCCE-A.N. Interestingly, we also discovered a cluster of epithelial cells with no expression of NE related markers (SCCE-nonNE), distributing in all the patients in our cohort (Fig. 6B). Pathway analysis showed that immune and inflammation-related signaling pathways were highly enriched in nonNE epithelial cells (Figure S9 F). SCENIC analysis identified high expression of TFs related to cell cycle and DNA repair in nonNE components, including BRCA1, E2 F1, NFKB1 (Figure S9I). Among NE compositions, both SCCE-A and SCCE-N showed upregulation of neuroendocrine signaling pathways, while SCCE-A enriched in cell–cell adhesion and epidermal differentiation pathways and SCCE-N enriched in cell cycle, protein metabolism related pathways (Figure S9G). Among our SCCE cohort, SCCE-nonNE distributed in each patient evenly, while SCCE-N exclusively concentrated in patient 09 and SCCE-A.N in patient 10, indicating tremendous heterogeneity in SCCE patients (Fig. 6C). To infer the evolutionary trajectory of SCCE, we applied cytotrace to delineate the developmental order of SCCE and found that SCCE-A.N showed the highest stemness while SCCE-N exhibited a high degree of differentiation (Figure S9H).

Considering that all the classification methods of NE tumors were based on TF transcription, we tried to develop a new classification framework on the single cell-level. We identified gene modules for each patient with preferentially expressed 30 genes and subsequently aggregated 120 gene modules into eight recurrent metaprograms (MPs) (Fig. 6E, S10 A-C, Table S6). Gene set enrichment analysis (GSEA) was employed to analyze the share and unique biological function of each MP.
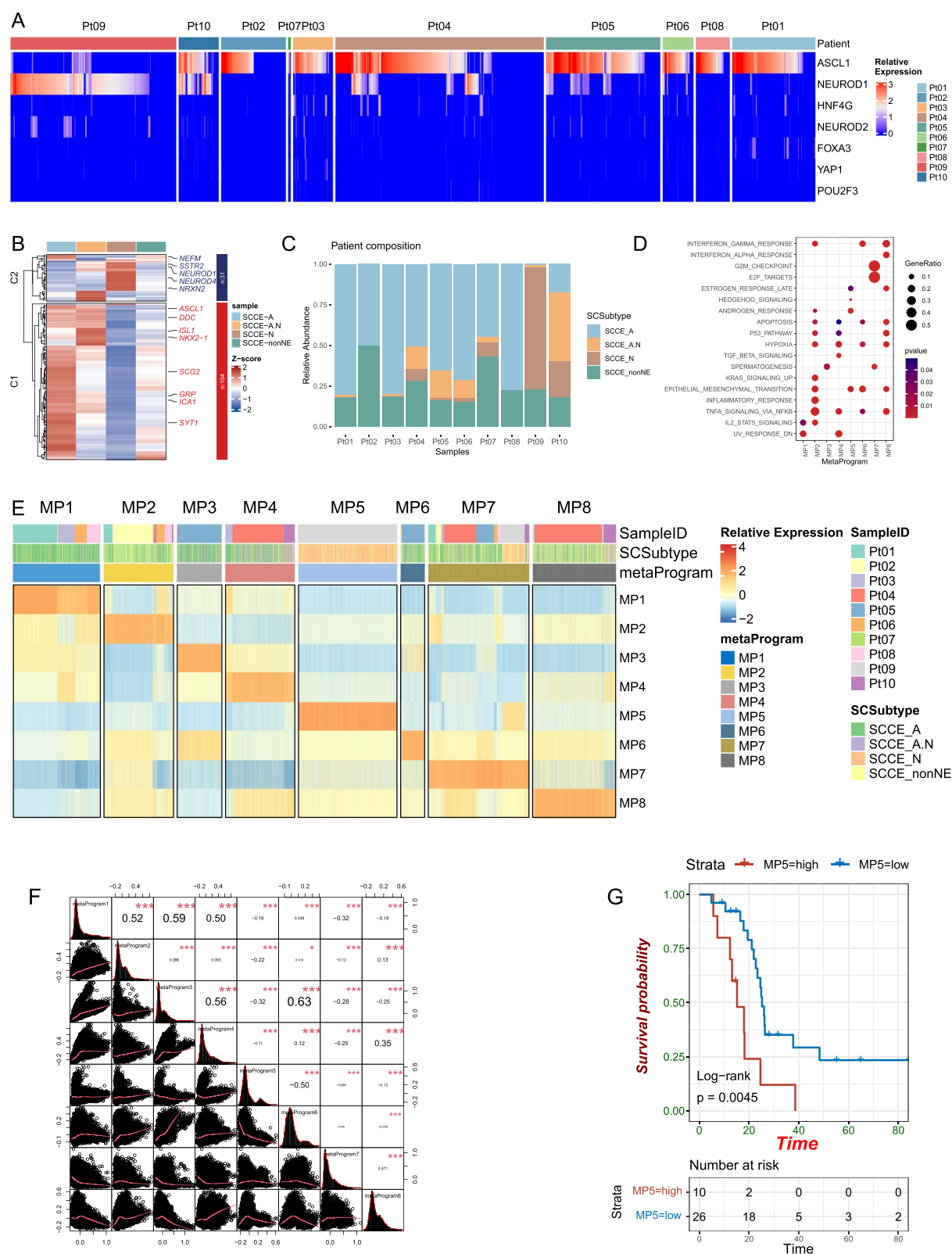
MP7 was uniquely enriched in hallmarks of cell cycle and proliferation (i.e., G2M_checkpoint). MP5 was predominantly enriched for neuroendocrine related pathways (i.e., androgen_response and estrogen_response_late). Both MP1 and MP4 showed stress-related characteristics (i.e., UV_response), while MP4 also enriched in apoptosis and hypoxia pathways. MP2 showed high enrichment in KRAS signaling and epithelial-mesenchymal transition pathways. MP6 and MP8 were enriched for interferon response pathway, while MP8 also showed high enrichment in p53 and estrogen response pathways (Fig. 6D). To further explore the relationship of different MPs, we performed correlation analysis to infer pairwise interaction and found 5 significant co-occurring program pairs and 1 exclusive program pair. MP1 was significantly co-occurred with MP2, 3, 4, indicating all these MPs had stress-related characteristics. MP5 was highly exclusive with MP6, suggesting neuroendocrine phenotype might negatively correlated with inflammation and apoptosis pathway, exhibiting a malignant phenotype (Fig. 6F). We then investigated the distribution of different MPs among our cohort and found that MP5 was exclusively distributed in Patient09 and SCCE-N subtype, consistent with the neuroendocrine phenotype of MP5 (Figure S10D-E). We then validated the survival prediction ability in external RNA-seq cohort [10] and discovered that MP5 significantly correlated with shorter OS in SCCE patients ($p < 0.01$), showing a better predictability than signature based on bulk RNA-seq (Fig. 6G).

Discussion

In this study, we compared the difference of TME between ESCC and SCCE and provided a comprehensive landscape of SCCE on single-cell resolution with ESCC as reference. We comprehensively characterized the immune and stromal compositions in the tumor microenvironment among SCCE and ESCC, including their transcriptional feature, cellular compositions, developmental lineages, as well as cellular ecotypes. Notably, we observed a more immunosuppressive ecosystem in SCCE compared to ESCC, characterized by decreased infiltration of effector CD8 + T cells, an enrichment of activated Treg cells, highly differentiated B and Plasma cells, and a specific pro-angiogenic

(See figure on next page.)

Fig. 6 Heterogeneity of SCCE defined by classical TFs expression and metaprograms. **A** Expression of neuroendocrine markers on single cell level from ten patients, including ASCL1, NEUROD1, HNF4G, NEUROD2, FOXA3, YAP1, POU2P3. **B** Expression of signatures of SCCE-A/N among four SCCE molecule subtypes, shown in heatmap. **C** Composition of different subtypes of epithelial cells in each SCCE patient. **D** Gene set enrichment with signature genes of each MP and the significantly enriched gens sets from MSigDB HALLMARK collection are shown. **E** Scaled signature scores of each MP (rows) across all individual malignant cells (columns). Cells are ordered based on the strength of the GM signature score. **F** Relationship of each pair of metaprogram. **G** Clinical relevance of MP5 in SCCE RNA-seq external validation cohort

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 12 of 19

Fig. 6 (See legend on previous page.)

Wu *et al. Molecular Cancer* (2025) 24:142

Page 13 of 19

niche induced by the crosstalk among FCN1 + tumor-associated macrophages, ESM1 + endothelial cells, and EGFL6 + pericytes. We constructed five ecotypes based on the non-epithelial components of the tumor microenvironment, each exhibiting distinct cell-compositional features and revealing their clinical significance in SCCE. Furthermore, we identified eight essential expression programs that highlight the significant heterogeneity of SCCE and fully depict their biological relevance.

Our analyses revealed that SCCE showed an exclusive phenotype with less immune cell infiltration compared to ESCC, indicating SCCE might less benefit from immunotherapy than ESCC. However, PD-L1 in tumor infiltrated immune cells and CD8 states were also demonstrated as independent predictors of OS in SCCE [2, 10]. Previous research conducted by our team revealed that 44.1% of SCCE patients had positive PD-L1 expression [9]. In our analysis of TME, SCCE showed higher CD8_c14_CXCL13 infiltration than healthy tissue, which was considered as immune reactive T cells accounting for response of immunotherapy. On the whole, the usage of immunotherapy in SCCE is expectable and more clinical data is needed to answer this question. What's more, we also found high differentiated B cell and lgG PC were highly enriched in SCCE. Shift from lgA to lgG PCs has been reported widely along with tumor progression compared to normal/adjacent normal tissue [12]. In gastrointestinal tissue, including SCCE, the normal mucosal immunity is mainly maintained by abundant lgA antibodies. The isotype of PC skewed from lgA to lgG in tumors not only associates with anti-tumor response, but also the impair of local mucosal immunity.

Another interesting finding in our research is that SCCE has more angiogenic activity in TME than ESCC, including the enrichment of tip-like endothelial cells, Endo_c3_ESM1, and vessel induced niche, comprising Macro_c01_FCN1 and Pericytes. Esm1 is a secreted protein thought to play a vital role in angiogenesis and inflammation. And thirty percent of VEGF-induced vascular permeability could be explained by Esm1 [16]. In advanced SCLC, SALUTE Trial showed that the addition of bevacizumab to cisplatin or carboplatin plus etoposide for treatment of extensive-stage SCLC improved PFS [26]. Anlotinib, a multi-target vessel tyrosine kinase inhibitor, improved PFS and OS in advanced SCLC patients relapsed within 3 months after second-line treatment and was well tolerated [27]. Considering the biological similarity among neuroendocrine carcinoma, anti-angiogenesis therapy might also benefit SCCE patients. In the article, we also provided a case in our center that a patient reached disease control after addition of anlotinib in second line therapy. These findings

offered us a promising direction for experimental application of anti-angiogenesis therapy in SCCE.

A more comprehensive comparison between SCCE and ESCC by ecotypes through unsupervised hierarchical clustering identified EC1 that was correlated with better survival in SCCE. EC1 showed high enrichment in neutrophil activation and chemotaxis related pathways. Although tumor-infiltrated neutrophils had much evidence in promoting tumor progression and metastasis [28], neutrophil response also showed positive correlation with benefit of immunotherapy in SCLC [23], indicating a heterogenous biological function of neutrophils in TME [29]. Previously R.D. Xue et al. had reported that inflamed tumor associated neutrophils (TANs), CCL4+ TANs might recruit macrophages through CCL4-CCR5. Y.C Wu discovered a cluster of neutrophils with high expression of HLA-DR, linking to a key anti-tumor effect across a wide range of cancer. Interestingly, SCCE showed exclusive enrichment of neutrophils than ESCC, indicating that neutrophils might played anti-tumor role to some extent in SCCE. However, further evidence and mechanistic research were needed to confirm this hypothesis.

Typical classification framework of neuroendocrine carcinoma was based on transcriptional markers, including ASCL1, NEUROD1, YAP1, POU2 F3. In our study, we also identified a cluster of epithelial cells expressing no neuroendocrine marker distributing in nearly all the patients, indicating a high heterogeneity among epithelial cells of SCCE. Here, we tried to dissect heterogeneity of SCCE through construction of metaprograms (MPs). MPs referred to the recurrent transcriptomic patterns across tumors of the same cancer type, reflecting the fundamental aspects of tumor biology [30]. An integrated analysis identifying the metaprograms of SCCE at single-cell resolutions revealed that stemness (MP7) and high proliferative ability were shared characteristics among SCCE. Neuroendocrine phenotype (MP5) was exclusive to Pt09, only one patient assigned to SCCE-N in our cohort.

Limitations of our study include the small sample size and lack of further experimental validation. Due to the rarity of SCCE in clinical practice, small sample size limited the generality in the construction of MPs of SCCE. Only one patient in our cohort was assigned as SCCE-N, which limited our analysis on this subtype. In addition, no suitable experimental model of SCCE hindered us from experimental validation or further exploration on possible target of SCCE. Despite these limitations, our work offered a comprehensive single-cell atlas for SCCE in multiverse dimensions highlighted by its ecotypes and metaprograms of epithelial cells. In conclusion, We provided valuable resources along with biological insights at

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 14 of 19

single-cell resolution to better understand the ecosystem of SCCE and to develop potential treatment strategies for patients with this condition.

Method

Human specimens

This study followed the principles according to the Declaration of Helsinki and was approved by the ethic committees in Sun Yat-sen University Cancer Center (B- 2020–311 -02), with written informed consents obtained from all patients. 10 patients were pathologically diagnosed with SCCE in Sun Yat-sen University Cancer Center. Independent review was conducted by experienced pathologists and radiologists to confirm disease diagnosis and the disease stages. Samples of primary tumor were obtained through biopsy.

The remaining two cohorts and validated bulk RNA-seq cohort ($n = 36$) were obtained from two published studies [10, 31, 32]. The detailed information of the three dataset was summarized in Table S1.

ScRNA-seq data generation

Fresh tissues were placed in a 6 cm dish and rinsed with 5 mL of DPBS and using a surgical blade to cut the tissue into < 3 mm^3 pieces, then rinse again with 5 mL of DPBS. After washing, remove the supernatant using a Pasteur pipette and further mince the tissue. Transfer the minced tissue to a 15 mL centrifuge tube containing the prepared enzyme solution, mix well, and place on a shaker at 37 °C, 200 rpm for 30 min, gently pipetting twice during the process. Once complete digestion is achieved (no obvious tissue clumps), remove the centrifuge tube. Filter through 70 μm and 40 μm mesh (Falcon,352,340) into a new 15 mL centrifuge tube and centrifuge at 400 g for 5 min at 4 °C with the brake on. After centrifugation, slowly remove the supernatant and resuspend the pellet with 2 mL of RBC Lysis buffer, disperse the cell clumps using a wide-bore pipette tip. After incubating on ice for 5 min, add 8 mL of DPBS containing 0.1% BSA and centrifuge again at 400 g for 5 min at 4 °C with the brake on. Remove the supernatant, resuspend the pellet with 200 μL of DPBS (0.1% BSA) and then remove debris and dead cells. The obtained cells are stained with AOPI and counted. Adjust the cell concentration to 700–1400 cells/μL and proceed with the 10X Chromium Single-cell 3' and 5'assay (Chromium Single Cell 3′Library & Gel Bead Kit v3 PN- 1000121; Chromium Next GEM Single Cell 5'Kit v2 PN- 1000263; Chromium Next GEM Single Cell 5'HT Kit v2 48 rxnsPN- 1000356), with 21,000 cells for tissue. All subsequent procedures were conducted following standard protocols provided by the manufacturer. The purified libraries were sequenced using illumina novaseq6000 and BGI T7 sequencer, generating 150-base pair (bp) paired-end reads.

Then, immune repertoire measurements were conducted with joining (VDJ) library construction (Chromium Single Cell V(D)J Amplification Kits Human 16 rxns TCR PN- 1000252, 10X Genomics) as per the manufacturer's guidelines. All subsequent procedures were conducted following standard protocols provided by the manufacturer. The purified libraries were sequenced using an BGI T7 sequencer, generating 150-base pair (bp) paired-end reads.

ScRNA-seq data processing

Low-quality cells were filtered out if cells had fewer than 500 genes expressed, or larger than 6000 genes, and >10% unique molecular identifiers (UMIs) linked to mitochondrial genes. The gene expression matrices of the remaining cells were generated with log normalization and linear regression using the NormalizeData and ScaleData functions of the Seurat package [33].

(v.5.0.1). Doublet cells were recognized by function doubletDetect imported from R package, DoubletFinder [34] (v.2.0.3) and were removed carefully after examination. The remaining cells that passed the filtering criteria were considered as single cells. For visualization, the dimensionality of each dataset was further reduced using UMAP with the Seurat function Run-UMAP. The principal components (PCs) used to calculate the embedding were determined by the elbow point accounting for more than 90% variance and ranging from 10 to 50. We constructed a shared nearest neighbour (SNN) based on expression profiles and applied an unsupervised graph-based clustering algorithm called the original Louvain to cluster single cells. We first performed unsupervised clustering with the parameter "resolution =0.8". We characterized major cell lineages with high expression of CD3D and CD3E (T cells), NKG7 and FCGR3 A (NK cells), MS4 A1 and CD19 (B cells), MZB1 and JCHAIN (plasma cells), TPSAB1 and CPA3 (mast cells), CD14 and S100 A8 (myeloid cells), CD1 C and CLEC9 A (Dendritic cell), CSF3R and IL1R2 (Neutrophils), EPCAM and KRT18 (Epithelial cell), PECAM1 AND VWF (Endothelial cell), COL1 A1 and DCN (Fibroblast). To characterize the subsets of major lineages, a second round of dimension reduction and unsupervised clustering was conducted, following the strategy described above. The detailed top markers of each clusters were summarized in Table S2.

Batch effect evaluation and correction

Statistical assessment of possible batch effects was performed using benchmarking framework, scIB [35] (single-cell integration benchmarking). Among three

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 15 of 19

common approaches, including fastmnn, rpca, harmony, we assess the efficacy through bio-conservation and batch correction two aspects. Finally, we employ Fastmnn for actual batch correction with higher comprehensive performance.

TCR analysis

The Cellranger software (v 7.1.0, 10X Genomics Inc.) provided by the 10X Genomics platform was utilized to align the data to the human VDJ reference genome (GRCh38). To minimize noise, only assembled chains that met the criteria of being highly confident, full length, and having a valid barcode were aligned and retained. If two or more cells had the same identical α/β chain pair, the α/β chain pair were identified as clonal TCRs and these T cells were considered to originate from the same clonotypes, identified as clonal cells. The R package scRepertoire (v 1.12.0) [36] was used to perform overlap analysis between different clusters to find the evolutionary lineage of T cells.

Tissue distribution analysis

To measure the preference of cell types in different esophageal carcinoma, we utilized the Ro/e index. This index is calculated as the ratio of observed to expected cell numbers. The observed cell numbers represent the actual counts, while the expected cell numbers are determined through the chi-square test [37].

$$R_{o/e} = \frac{Observed}{Expected}$$

Evolutionary trace

R package slingshot (v 2.10.0) was used to perform trajectory inference for single cell-data [38]. We figured out the major evolutionary pathway of T cell under guidance of TCR clonetype overlapping analysis.

Differential expression and Gene Ontology enrichment analysis

The significantly overexpressed marker genes for each cluster was identified using the FindAllMarkers() function of Seurat. Genes with adjusted P value <0.05 by Wilcoxon rank-sum test were defined as cluster-specific signature genes. For two interesting clusters, we employed FindMarkers() and Wilcoxon test to determine the significantly expressed gene between two clusters. Genes with adjusted P value <0.05 and pct >0.1 were considered as differentially expressed genes that were further used for volcano plot visualization and GO enrichment analysis with the clusterProfiler package (v.4.12.0) [39]. GO terms with adjusted P values <0.05, were considered as significant.

Cell differentiation analysis

We used CytoTRACE software to predict developmental potential based on cellular transcriptional diversity [40]. In our study, CytoTRACE helped us to evaluate stemness of different cellular clusters.

Signature score analysis

To discover the functional states of T, B, Myeloid, Endothelial, Neutrophil cells, we collected a list of published gene signatures including the cytotoxic score, T cell dysfunctional score (T cell) [41]; positive immune regulation, antigen presentation (B cell); M1, M2, angiogenesis, phagocytosis (Macrophage) [42]; Tip-like, arterial, immature, activated, postcapillary venous, activated postcapillary, lymphatic, defense, scavenging (Endothelial cell) [43]; neutrophil-related functional signatures (Neutrophil) [44], and cancer hallmark gene sets downloaded from the Molecular Signature Database. The signature scores of these gene signatures and pathways in cells of each cell subsets were calculated through function AddModuleScore in Seurat.

Multiplexed immunohistochemistry

Multiplexed immunohistochemistry (mIHC) was performed on 4-μm-thick formalin-fixed, paraffin-embedded whole tissue sections using standard primary antibodies sequentially, in conjunction with the TSA 7-color kit (abs50015‐100 T, Absinbio, Shanghai). The process was followed by DAPI staining. For instance, deparaffinized slides were incubated with an anti-CD68 antibody (#76437 CST) for 30 min, followed by treatment with an anti-rabbit/mouse horseradish peroxidase (HRP)-conjugated secondary antibody (abs50015‐02, Absinbio, Shanghai) for 10 min. Labeling was then developed for exactly 10 min using TSA 520, according to the manufacturer's instructions. Afterward, slides were washed in TBST buffer and transferred to a preheated citrate solution (90 °C) before undergoing heat treatment in a microwave at 20% of maximum power for 15 min. The slides were then cooled to room temperature in the same solution. Between each step, they were thoroughly washed with Tris buffer. This process was repeated for the subsequent antibodies and fluorescent dyes in the following order: anti-CD68/TSA 570, anti-FCN1/TSA 480, anti-ESM1/TSA 620, anti-RGS5/TSA 690, and anti-VEGFR2/TSA 520. Finally, each slide was treated with two drops of DAPI (abs47047616, Absinbio, Shanghai), washed in distilled water, and manually coverslipped. The slides were air-dried and imaged using the Vectra Polaris™ Automated Quantitative Pathology Imaging System

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 16 of 19

(Akoya Biosciences). Image analysis was performed using Indica Halo software.

Hierarchical relationships among TME subsets

We performed unsupervised hierarchical cluster analysis, as described in a paper previously, to clarify the phenotypic relationship among these TME subsets identified in our study. Generally, the dendrogram was drawn based on Pearson correlation coefficients with average PCA space (Seurat function RunPCA) for each subset [42].

Ecotype identification in esophageal carcinoma

To examine the cellular compositions of different TME ecosystems in esophageal cancer among Asian population, we investigated the co-existence patterns of different TME cells in healthy control, normal tissues adjacent to the ESCC, ESCC, SCCE. Pairwise spearman correlation values between the normalized frequency of any two clusters across different samples were calculated using the corr function (Table S3). The values were clustered using the ComplexHeatmap (v.2.16.0) [45] R package with ward.D2 method and euclidean distance. As a result, we identified five ecotypes in esophageal carcinoma and defined gene signatures for each ecotype by combining top 10 marker genes of all clusters through function FindAllMarkers in the corresponding ecotype (Table S4). For each patient in single cell cohort, cluster-normalized frequencies of related clusters from the same ecotype were summed and the most abundant ecotype was designated as the dominant ecotype for this patient. For bulk RNA-seq cohort, we assigned ecotype to each patient according to the highest ssgsea score of five ecotypes. The biological function of each ecotype was evaluated through GO enrichment analysis and prognostic relevance.

Cell–cell communication analysis

Potential cell–cell interaction between any two cell subsets was predicted by Cellchat (v2.1.0) [46], a computational framework based on published cell–cell interaction ligand-receptor pairs. The core functions, including 'computeCommunProb', 'computeCommunProbPathway', 'aggregateNet', were applied with standard parameters. 'netVisual' function was used to visualize the significant cell–cell interaction pairs (L-R pairs).

Identifying neoplastic cells from normal esophageal epithelial cells

Copy number variations (CNV) signal in each epithelial cell was recalled by infercnv (v1.20.0) based on single cell sequencing data with a default length of sliding window. The method employed for smoothing was pyramidal. The epithelial cells from normal esophageal mucosa were used as reference, while epithelial cells from SCCE were used for observations. CNV_score for each cell was calculated for each epithelial cells by summing the squares of these inferred changes to one at each genomic locus. 'Kmeans' clustering method was applied to all the epithelial cells with fixed seed and the cells clustered with healthy control were considered as normal epithelial cells. In our study, over 90% of epithelial cells in SCCE were identified as malignant ones for high cnv, suggesting that SCCE had a more aggressive phenotype.

Calling intrinsic subtype on scRNA-seq based on classical neuroendocrine transcriptional factors

To apply classical TF classification framework to SCCE scRNA-seq data, we construct pseudo-bulk expression matrix for each tumor. First, we defined robust subtypes through three approaches: 1) detect the relative expression of related transcriptional factor in each pseudo-bulk tumor sample; 2) hierarchical clustering of the pseudo-bulk data with 36 RNA-seq tumor samples published in Nature Comm, 2021, employing the signature list of A/N subtypes; 3) Assign A/N subtype to each cell in scRNA-seq through AddModuleScore function. An overall tumor subtype was determined by the majority cell subtype. All the three methods indicated the same result of classification. To dissect heterogeneity of epithelial cells on single cell level, we divided the epithelium of SCCE into 4 clusters according to the expression of ASCL1 and NEUROD1 through kmeans with fixed seed, among which the cluster harbored low expression of ASCL1 and NEUROD1 was identified as non-neuroendocrine components.

Identify metaprograms of small cell carcinoma of esophagus

For total 10 samples of SCCE with >100 epithelial cells, non-matrix factorization (NMF v0.26) [47] was performed separately on the identified malignant cells of each sample. Starting from scaled matrix of each sample, all negative values were set to zero. Only the genes with variance of expression >0.5 and recognized as highly variable genes were kept for further analysis. The 'nsNMF' method was applied for ranks between 2 and 25. To define transcriptomic gene modules individually, function *extractFeature* was employed to extract the top 30 genes from each NMF program (Table S5). A total of 120 gene modules were obtained from 10 samples, and pairwise pearson correlation values between scores of any two modules across cells were calculated using the corr function. To obtain recurrent metaprograms among SCCE, unsupervised hierarchical cluster analysis was applied to the matrix of correlation coefficient. The optimal number of clusters was determined by gap statistics. Totally, eight metaprograms were identified and gene

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 17 of 19

signatures of each metaprogram were extracted as previously described. All the mitochondrial, ribosomal genes and non-coding genes were excluded [48].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12943-025-02335-5.

Additional file 1. Extended Figure 1. Supplementary data for single cell atlas of esophageal comparative cohort. **A** Efficacy assessment of different batch removal method including harmony, rpca, fastmnn through two dimensions. **B** UMAP visualization of single cell landscape across each individual. **C** Expression featureplots of selected markers. **D** Summary of clinical information of ten SCCE patients from SYSUCC. **E** The composition of major clusters among each patient, shown in stack plot. **F** The composition of major clusters across different tissue origins, shown in pie chart. Extended Figure 2. Supplementary data for T cells. **A** Composition of T cell subclusters among each patient. **B** Heatmap of selected genes on different expression levels across different T and NK clusters. **C** UMAP plots showing T cell clusters and cell distribution density (shown as thermo-dynamic chart) across different tissue groups. **D** CloneType distribution among each T cell subcluster (Large, 20<x≤ 100; Medium, 5<x≤ 20; small, 1<x≤ 5; single, 0<x≤ 1). **E** Clone overlap analysis between different T cell subsets. **F** Differentiation analysis between CD8_c14_Tex (CXCL13) and other CD8 T cells. **G** Expression of selected markers across T and NK clusters (i.e CXCL13, LAG3, CTLA4, TIGIT). Extended Figure 3. Supplementary data for B and plasma cells. A Expression levels of selected markers in B cells, shown in dotplot. **B** Boxplot comparison of functional B cell subclusters among different tissue origins. **C** Expression levels of selected markers in plasma cells, shown in heatmap. **D** Heatmap displaying the tissue distribution preference of plasma cells through ratio of observed to expected cell numbers (Ro/e). **E** Comparison of cytotrace predicted score of plasma cells among different tissue origins. Extended Figure 4. Supplementary data for myeloid and stromal cells. **A** Heatmap displaying the tissue distribution preference of myeloid cells through ratio of observed to expected cell numbers (Ro/e). **B** Boxplot comparison of myeloid cell clusters (Macro_c01, c03, 04) among different tissue origins. **C** GO enrichment analysis based on DEGs between M01-M03 enriched in SCCE and M04 enriched in ESCC. **D** Comparison of tissue-resident macrophage score between selected subsets of macrophages. **E** Expression of Tip-like signature among different endothelial subsets. **F** Comparison of proportion of lymph endothelial cells across different cell origin lineages. **G** GO enrichment analysis based on DEGs between E02-E03 enriched in SCCE and E04 enriched in ESCC. **H** Relative expression of vessel disorganization score among among different tissue origins, shown in boxplot. **I** Expression levels of selected markers in stromal cells, shown in dotplot. **J** GO enrichment analysis of pericyte_c07_EGFL6. Extended Figure 5. Supplementary data for cell-cell communication among myeloid and stroma subsets. **A** Heatmap showing the expression of angiogenesis related ligand-receptor pairs highly expressed in endothelial and myeloid subsets. **B** Heatmap showing the expression of angiogenesis related ligand-receptor pairs across different tissue origins. **C-D** The PLAU-PLAUR, PGF-VEGFR1 signaling pathway between pericyte and endothelial cell clusters, scaled by weight. **E-F** Relative expression of angiogenesis related L-R pairs in stromal subsets. **G** Graphical illustration of vessel-induced niche in SCCE. Extended Figure 6. Supplementary data for neutrophils. **A** UMAP visualization of neutrophils separated into 6 subsets. **B** Relative expression of neutrophil-related functional signatures among different neutrophil subsets, shown in heatmap. **C** Comparison of cytotrace predicted score of neutrophils among different tissue origins. **D** Expression of HALLMARK collection gene sets among different clusters of neutrophils. Extended Figure 7. Supplementary data for a clinical case. Treatment time line and image examination for a patient with SCCE. Extended Figure 8. Supplementary data for epithelial cells. **A** Differentiation analysis between Epithelial cells between SCCE and ESCC. **B** GO enrichment analysis based on DEGs between Epithelial cells between SCCE and ESCC. **C-D** Communication between different major cell clusters between ESCC and SCCE, shown in count and weight methods. **E** K-means was applied to CNV_level and identified the malignant epithelial cells. **F** Relative expression of CNV level among different kmean celusters. Extended Figure 9. Supplementary data for SCSubtype of epithelial cells. **A** Relative expression of five classical TFs markers among each patient as 'pseudobulk' mode. **B** Clustering of ten SCCE patients as pseudobulk with 42 external RNA-seq validation cohort. **C** Composition of different subtypes of epithelial cells in each patient. **D** Expression of selected markers in each patient. **E** Kmean clustering among all the malignant cell in SCCE based on the expression of ASCL1, and NEUROD1. **F** GO enrichment analysis based on DEGs between SCCE-NE cells and SCCE-nonNE cells. **G** GO enrichment analysis based on DEGs between SCCE-A cells and SCCE-N cells. **H** Comparison of cytotrace predicted score of epithelial cells among differnet tissue. **I** Predicted TFs in different epithelial subtypes through SCENIC pipeline. Extended Figure 10. Supplementary data for MPs of epithelial cells. **A** Flowchart of identification of MPs in epithelial cells. **B** Concurrent metaprograms among SCCE patients. **C** Boxplot displaying the relative expression of MPs in each patient. **D** Composition of the MPs among each patient. **E** Composition of the MPs among each SCSubtype among SCCE

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committees of Sun Yat-sen University Cancer Center (No. B2020 - 311–01). Written informed consent was obtained from all patients in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Wu *et al. Molecular Cancer*   (2025) 24:142

Page 18 of 19

References

1. Shah MA, Altorki N, Patel P, Harrison S, Bass A, Abrams JA. Improving outcomes in patients with oesophageal cancer. Nat Rev Clin Oncol. 2023;20(6):390–407.
2. Zhang C, Zhang G, Xue L, Zhang Z, Zeng Q, Wu P, et al. Patterns and prognostic values of programmed cell death-ligand 1 expression and CD8 + T-cell infiltration in small cell carcinoma of the esophagus: a retrospective analysis of 34 years of National Cancer Center data in China. Int J Surg. 2024;110(7):4297–309.
3. Zhu J, Wang Y, Sun H, Zhang Y, Zhang W, Shen W, et al. Surgery versus radiotherapy for limited-stage small cell esophageal carcinoma: a multicenter, retrospective, cohort study in China (ChiSCEC). Int J Surg. 2024;110(2):956–64.
4. Chen S, Wang XM, Wu F, Huang C, Gao TT, Zhang ZW, et al. Primary small cell carcinoma of the esophagus in a large multicenter cohort: prognostic factors and treatment strategies in the modern era. Int J Radiat Oncol Biol Phys. 2023;117(2):e286–7.
5. Xu L, Li Y, Liu X, Sun H, Zhang R, Zhang J, et al. Treatment strategies and prognostic factors of limited-stage primary small cell carcinoma of the esophagus. J Thorac Oncol. 2017;12(12):1834–44.
6. Brenner B, Tang LH, Klimstra DS, Kelsen DP. Small-cell carcinomas of the gastrointestinal tract: a review. J Clin Oncol. 2004;22(13):2730–9.
7. Ji A, Jin R, Zhang R, Li H. Primary small cell carcinoma of the esophagus: progression in the last decade. Ann Transl Med. 2020;8(7):502.
8. Wang F, Liu DB, Zhao Q, Chen G, Liu XM, Wang YN, et al. The genomic landscape of small cell carcinoma of the esophagus. Cell Res. 2018;28(7):771–4.
9. Zhao Q, Chen YX, Wu QN, Zhang C, Liu M, Wang YN, et al. Systematic analysis of the transcriptome in small-cell carcinoma of the oesophagus reveals its immune microenvironment. Clin Transl Immunol. 2020;9(10): e1173.
10. Li R, Yang Z, Shao F, Cheng H, Wen Y, Sun S, et al. Multi-omics profiling of primary small cell carcinoma of the esophagus reveals RB1 disruption and additional molecular subtypes. Nat Commun. 2021;12(1):3785.
11. Liu B, Zhang Y, Wang D, Hu X, Zhang Z. Single-cell meta-analyses reveal responses of tumor-reactive CXCL13(+) T cells to immune-checkpoint blockade. Nat Cancer. 2022;3(9):1123–36.
12. Yang Y, Chen X, Pan J, Ning H, Zhang Y, Bo Y, et al. Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes. Cell. 2024;187(17):4790-811 e22.
13. Bresnick AR, Weber DJ, Zimmer DB. S100 proteins in cancer. Nat Rev Cancer. 2015;15(2):96–109.
14. Etzerodt A, Moulin M, Doktor TK, Delfini M, Mossadegh-Keller N, Bajenoff M, et al. Tissue-resident macrophages in omentum promote metastatic spread of ovarian cancer. J Exp Med. 2020;217(4):e20191869.
15. Zheng X, Wang X, Cheng X, Liu Z, Yin Y, Li X, et al. Single-cell analyses implicate ascites in remodeling the ecosystems of primary and metastatic tumors in ovarian cancer. Nat Cancer. 2023;4(8):1138–56.
16. Rocha SF, Schiller M, Jing D, Li H, Butz S, Vestweber D, et al. Esm1 modulates endothelial tip cell behavior and vascular permeability by enhancing VEGF bioavailability. Circ Res. 2014;115(6):581–90.
17. Grigoriu BD, Depontieu F, Scherpereel A, Gourcerol D, Devos P, Ouatas T, et al. Endocan expression and relationship with survival in human non-small cell lung cancer. Clin Cancer Res. 2006;12(15):4575–82.
18. Maurage CA, Adam E, Mineo JF, Sarrazin S, Debunne M, Siminski RM, et al. Endocan expression and localization in human glioblastomas. J Neuropathol Exp Neurol. 2009;68(6):633–41.
19. Zuo L, Zhang SM, Hu RL, Zhu HQ, Zhou Q, Gui SY, et al. Correlation between expression and differentiation of endocan in colorectal cancer. World J Gastroenterol. 2008;14(28):4562–8.
20. De Falco S. The discovery of placenta growth factor and its biological activity. Exp Mol Med. 2012;44(1):1–9.
21. Shah FH, Lee HW. Endothelial and macrophage interactions in the angiogenic niche. Cytokine Growth Factor Rev. 2024;78:64–76.
22. Bergers G, Song S. The role of pericytes in blood-vessel formation and maintenance. Neuro Oncol. 2005;7(4):452–64.
23. Gungabeesoon J, Gort-Freitas NA, Kiss M, Bolli E, Messemaker M, Siwicki M, et al. A neutrophil response linked to tumor control in immunotherapy. Cell. 2023;186(7):1448-64 e20.
24. Gay CM, Stewart CA, Park EM, Diao L, Groves SM, Heeke S, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. Cancer Cell. 2021;39(3):346-60 e7.
25. Wang Z, Liu C, Zheng S, Yao Y, Wang S, Wang X, et al. Molecular subtypes of neuroendocrine carcinomas: a cross-tissue classification framework based on five transcriptional regulators. Cancer Cell. 2024;42(6):1106-25 e8.
26. Spigel DR, Townley PM, Waterhouse DM, Fang L, Adiguzel I, Huang JE, et al. Randomized phase II study of bevacizumab in combination with chemotherapy in previously untreated extensive-stage small-cell lung cancer: results from the SALUTE trial. J Clin Oncol. 2011;29(16):2215–22.
27. Shi J, Cheng Y, Wang Q, Li K, Wu L, Han B, et al. Effect of anlotinib in advanced small cell lung cancer (SCLC) patients relapsed within three months after second-line treatment: a subgroup analysis from a randomized, double-blind phase II trial (ALTER 1202). J Clin Oncol. 2020;38(15_suppl):9063-.
28. Silvestre-Roig C, Kalafati L, Chavakis T. Neutrophils are shaped by the tumor microenvironment: novel possibilities for targeting neutrophils in cancer. Signal Transduct Target Ther. 2024;9(1):77.
29. Wu Y, Ma J, Yang X, Nan F, Zhang T, Ji S, et al. Neutrophil profiling illuminates anti-tumor antigen-presenting potency. Cell. 2024;187(6):1422-39 e24.
30. Gavish A, Tyler M, Greenwald AC, Hoefflin R, Simkin D, Tschernichovsky R, et al. Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. Nature. 2023;618(7965):598–606.
31. Dinh HQ, Pan F, Wang G, Huang QF, Olingy CE, Wu ZY, et al. Integrated single-cell transcriptome analysis reveals heterogeneity of esophageal squamous cell carcinoma microenvironment. Nat Commun. 2021;12(1):7335.
32. Nowicki-Osuch K, Zhuang L, Jammula S, Bleaney CW, Mahbubani KT, Devonshire G, et al. Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition. Science. 2021;373(6556):760–7.
33. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024;42(2):293–304.
34. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8(4):329-37 e4.
35. Luecken MD, Buttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods. 2022;19(1):41–50.
36. Borcherding N, Bormann N. scRepertoire: an R-based toolkit for single-cell immune receptor analysis [version 1; peer review: 2 approved with reservations]. F1000Research. 2020;9:47.
37. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature. 2018;564(7735):268–72.
38. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018;19(1):477.
39. Xu S, Hu E, Cai Y, et al. Using clusterProfiler to characterize multiomics data. Nat Protoc. 2024;19:3292–320. https://doi.org/10.1038/s41596-024-01020-z.
40. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020;367(6476):405–11.
41. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. Nat Genet. 2021;53(9):1334–47.
42. Wang R, Song S, Qin J, Yoshimura K, Peng F, Chu Y, et al. Evolution of immune and stromal cell states and ecotypes during gastric adenocarcinoma progression. Cancer Cell. 2023;41(8):1407-26 e9.
43. Goveia J, Rohlenova K, Taverna F, Treps L, Conradi LC, Pircher A, et al. An integrated gene expression landscape profiling approach to identify lung tumor endothelial cell heterogeneity and angiogenic candidates. Cancer Cell. 2020;37(1):21-36 e13.
44. Xie X, Shi Q, Wu P, Zhang X, Kambara H, Su J, et al. Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. Nat Immunol. 2020;21(9):1119–33.
45. Gu Z. Complex heatmap visualization. Imeta. 2022;1(3):e43.

46. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using Cell Chat. Nat Commun. 2021;12(1):1088.
47. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010;11: 367.
48. Barkley D, Moncada R, Pour M, Liberman DA, Dryg I, Werba G, et al. Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. Nat Genet. 2022;54(8):1192–201.

Publisher's Note