Integrating microarray-based spatial transcriptomics and single-cell RNA-sequencing reveals tissue architecture in oesophageal squamous cell carcinoma

Wei Guo
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Bolun Zhou
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Xiang Liu
Echo Biotech Co., Ltd
Qilin Huai
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Lei Guo
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Xuemin Xue
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Fengwei Tan
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Yin Li
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Qi Xue
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Shugeng Gao
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College
Jie He (hejie@cicams.ac.cn)
National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College

Research Article

Keywords: ESCC, scRNA-seq, spatial transcriptomics, tumour microenvironment

Posted Date: April 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1539014/v1

Abstract

Background

The tumour microenvironment (TME) serves as an important factor in tumorigenesis and metastasis. Although distinct cell subsets can be identified via single-cell RNA sequencing (scRNA-seq), the spatial composition of cells within the TME is difficult to characterise.

Methods

Tissue samples were collected from three patients with oesophageal squamous cell carcinoma (ESCC), and scRNA-seq was performed to identify distinct cell subsets. In addition, a microarray-based spatial transcriptomic (ST) method was used to characterise the spatial landscape of expression data via an array of spots. Using multimodal intersection analysis (MIA) to integrate scRNA-seq and ST, the exact cellular components of the tumour and stromal regions were annotated.

Results

The subpopulations of seven stromal cells were identified within the TME of ESCC, and the architecture of scRNA-seq-determined subsets was mapped in cancer and stromal regions. The distribution of various stromal cells and their subpopulations was heterogeneous. Compared with immune cells, non-immune stromal cells were significantly enriched in the TME. Most subsets of epithelial cells were enriched in the cancer regions, whereas inflammatory cancer-associated fibroblasts were correlated with the stromal regions. Furthermore, TME features were different between metastatic and non-metastatic samples and between the primary and metastatic sites of the metastatic sample.

Conclusions

This study revealed the spatial landscape of various cell subsets within the TME and the potential cross-talk among diverse cells, which provides novel insights into cancer intervention.

Introduction

Oesophageal cancer is the eighth main cause of cancer-specific death worldwide, leading to approximately 0.54 million deaths in 2020 (1). Oesophageal squamous cell carcinoma (ESCC) and oesophageal adenocarcinoma (EAC) are the two main subtypes, with ESCC accounting for the majority of all cases (1, 2). The risk factors differ between these two subtypes: alcohol and smoking are associated with ESCC, whereas obesity and reflux are associated with EAC (3–5). Surgery, chemotherapy, chemoradiotherapy, and immunotherapy are the standard treatments for oesophageal cancer (6). Despite the improvement in treatment strategies, some patients have unfavourable clinical outcomes, and the 5-

year survival rate is < 25% (7). In recent years, immunotherapy has emerged as a novel strategy and has been considered a promising therapy for oesophageal cancer, with several clinical trials investigating the safety and efficacy of immunotherapy or combination therapy. Luo et al. showed that anti-PD1 treatment prolonged the survival rate of many patients with advanced or metastatic ESCC; however, some patients did not respond to specific therapy (8). The tumour microenvironment (TME) is heterogeneous in oesophageal cancer, and intratumour heterogeneity may be closely related to different responses to a specific treatment strategy, such as immunotherapy (9, 10). However, few effective strategies have been reported to classify the TME subtypes and analyse the association between the diversity of TME and metastasis in ESCC, which prevents the development of precise medications for patients.

Single-cell RNA sequencing (scRNA-seq) has been considered an unprecedented method to unveil the potentially significant understanding of different cell subtypes, including multiple cell subpopulations in tumours. Various cellular subsets have been classified in cancer, such as subpopulations of tumour-associated stromal cells and tumour-infiltrating immune cells (11, 12). In addition, the discovery of a multicellular ecosystem within the TME has provided novel insights into the intratumoral transcriptional heterogeneity in many cancers (13–15). Furthermore, the subpopulations and proportion of cells may act as potential predictors of clinical outcomes in patients receiving a specific therapy, indicating the promising future of scRNA-seq in cancer treatment (16, 17). However, tissue dissociation before scRNA-seq analysis usually leads to the loss of spatial information, restricting the investigation of cellular cross-talk in the TME.

The spatial transcriptomic (ST) technology, which is complementary to the scRNA-seq technology, can overcome the abovementioned restriction (18). Using spatially barcoded histological microarrays, ST provides an intact two-dimensional landscape of transcripts over an entire tissue section. ST has been widely used to evaluate the spatial heterogeneity of melanoma (19), colorectal cancer (20), pancreatic ductal adenocarcinoma (21), prefrontal cortex (22), the heart (23) and the mouse brain (24). Although ST can provide detailed two-dimensional information of transcripts in different tissues, it has limitations, such as the relatively low cellular resolution. The transcriptomes may differ from spot to spot depending on the number of cells captured by each spot, resulting in heterogeneity of the results (21). To overcome their limitations, we integrated scRNA-seq and ST analyses to comprehensively analyse ESCC tissues.

In this study, we identified the subsets of seven cell types and analysed the ST architecture of three ESCC samples. We characterised the features of TME and analysed the distribution of different cell subpopulations in the cancer and stromal regions of ESCC. In addition, we identified different cell subsets in the cancer region between metastatic and non-metastatic samples and in all regions between the primary and metastatic sites of the metastatic sample. The results revealed the comprehensive landscape and cellular ecosystem of ESCC, which may provide novel insights into its treatment and prevention.

Materials And Methods

Patients and sample collection

Three patients who were diagnosed with pathologically confirmed ESCC and underwent surgery at the Department of Thoracic Surgery of the National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, were enrolled. After samples were resected, tumor tissues were cut into three pieces along the long axis: one was processed for pathological diagnosis, one was processed for scRNA-seq, and the other one was processed for the spatial transcriptomic. One metastatic lymph node was also collected, and was cut into twoe pieces: one was processed for pathological diagnosis, one was processed for scRNA-seq. This study was approved by the National Cancer Center/Cancer Hospital Ethics Committee. Written informed consent was obtained from all participants included in this study.

Cell capture and cDNA synthesis

Using the Single Cell 3' Library and Gel Bead Kit V3 (10x Genomics, 1000075, Pleasanton, CA, USA) and Chromium Single Cell B Chip Kit (10x Genomics, 1000074), a cell suspension (300—600 living cells/μL, determined by Count Star) was loaded onto the Chromium Single Cell Controller (10x Genomics) to generate single-cell gel beads in the emulsion according to the manufacturer's protocol. Briefly, single cells were suspended in PBS containing 0.04% BSA. Approximately 11,000 cells were added to each channel, and approximately 7,000—10,000 cells were estimated to be recovered as target cells. The captured cells were lysed, and the released RNA was barcoded through reverse transcription in individual GEMs. Reverse transcription was performed on an S1000TM Touch Thermal Cycler (Bio-Rad) at 53°C for 45 min, followed by 85°C for 5 min and hold at 4°C. The cDNA was generated, amplified and assessed for its quality on an Agilent 4200 system (performed by CapitalBio Technology, Beijing).

Single-cell RNA-sequencing library preparation

scRNA-seq libraries were constructed using the Single Cell 3´ Library and Gel Bead Kit V3 according to the manufacturer's instructions and subsequently sequenced using an Illumina Novaseq6000 sequencer with a sequencing depth of at least 100,000 reads per cell using a paired-end 150 bp (PE150) reading strategy (performed by CapitalBio Technology, Beijing).

Staining and imaging

Cryosections of 10-μm thickness were cut and mounted onto the GEX arrays. The sections were placed on a thermocycler adaptor with the active surface facing up and incubated for 1 min at 37°C. Subsequently, the sections were fixed for 30 min with methyl alcohol at -20°C and stained with haematoxylin and eosin (H&E) (eosin, Dako CS701, haematoxylin, Dako S3309, bluing buffer CS702). Brightfield images were taken on a Leica DMI8 whole-slide scanner at 10x resolution.

Permeabilisation and reverse transcription

Gene expression was analysed using a Visium Spatial Gene Expression slide and Reagent Kit (10x Genomics, PN-1000184). A slide cassette was used to create leakproof wells for adding reagents.

Approximately 70-μL permeabilisation enzyme was added and incubated at 37℃ for 60 min. Each well was washed with 100-μL SSC, and 75-μL reverse transcription Master Mix was added for cDNA Synthesis.

cDNA library preparation for sequencing

At the end of first-strand synthesis, the RT Master Mix was removed from the wells, and 75 μL of 0.08-M KOH was added and incubated for 5 min at room temperature. Subsequently, KOH was removed, and the wells were washed with 100-uL EB buffer. Thereafter, 75-μL Second Strand Mix was added to each well for second-strand synthesis. cDNA amplification was performed on an S1000TM Touch Thermal Cycler (Bio-Rad). Spatial libraries were constructed using the Visium Spatial Library Construction Kit (10x Genomics, PN-1000184) according to the manufacturer's instructions and subsequently sequenced using an Illumina Novaseq6000 sequencer with a sequencing depth of at least 100,000 reads per spot using a pair-end 150 bp (PE150) reading strategy (performed by CapitalBio Technology, Beijing).

scRNA-seq data analysis

The CellRanger software was used for quantitative analysis of the gene expression of scRNA-seq data, and the results were filtered using the R package Seurat. Cells with < 200 or > 2500 expressed genes and mitochondrial gene content of > 10% were removed. The remaining cells were used for downstream analysis. Data of all samples were combined using the merge function of Seurat and standardised using the NormalizeData function of Seurat. The FindVariableGenes function was used to determine 2,000 highly variable genes. The RunPCA and RunUMAP functions were used for dimensionality reduction clustering. The final number of principal components (PCs) is 20, which is determined by the inflection point of the ElbowPlot function. And the P-value of 20 PCs is < 0.05 in the JackStrawPlot result. Based on the clustering results, cell types were annotated using a combination of the R package singleR and previously reported gene markers. The data of each cell type were extracted, and each cell type was clustered using the abovementioned dimensionality reduction clustering method. The R package monocle2 was used to analyse the pseudotime of epithelial cells. Differentially expressed genes (DEGs) were analysed at the single-cell level using the FindAllMarkers function of Seurat (min.pct = 0.25).

Spatial transcriptomic analysis

The gene expression of ST sequencing data was analysed using the CellSpace software. After the gene expression was quantified, a downstream analysis was performed using the R package Seurat. The SCTransform function of Seurat was used to standardise the spatial transcriptomic data of three samples, whereas RunPCA and RunUMAP were used for dimensionality reduction and clustering (PC = 30). The FindAllMarkers function was used to analyse the characteristic genes of each cluster in the spatial transcriptome (min.pct = 0.25, logfc.threshold = 0.25). Combined with the characteristic genes of each cell type in a single cell, multimodal intersection analysis (MIA) was performed to obtain the main cell types distributed in each spatial region. All spatial spots were divided into cancer (regions with abundant epithelial cells) and stromal (regions with abundant non-epithelial cells) regions. Pseudotime analysis of T1 and N1 epithelial cells was performed using monocle2, and the three states of epithelial

cells resulting from the analysis were annotated onto the spatial transcriptome using the FindTransferanchors function of Seurat. Based on the annotation, spatial spots were divided into four regions. The spatial regions annotated by states 1 and 2 with a concentration of metastatic epithelial cells were named ‘Cancer Region 1’ and ‘Stromal Region 1’, respectively, whereas the remaining cancer and stromal regions were named ‘Cancer Region 2’ and ‘Stromal Region 2’, respectively. DEGs between two cancer regions and between two stromal regions were analysed using the FindMarkers function of Seurat.

Functional enrichment and survival analyses

The KEGG pathway of each DEG was annotated using BLAST to align the gene sequence to protein sequences obtained from the KOBAS database. The p-value of enrichment significance was evaluated using Fisher's exact test for each pathway and corrected using the BH method to obtain the q-value. The coxph function of the R package survival was used to perform Cox regression survival analysis on 162 colorectal cancer samples in the TCGA database. Two survival curves were generated using the DEGs of cancer and stromal regions. CIBERSORT was used to score immune cells in the spatial spot of T1 samples, and the ReactomeGSA package was used to score the Hallmark gene set in the spatial spot.

Statistical analysis

The statistical software, threshold and methods of each bioinformatic analysis are described in Results, Methods and Figure legends.

Results

Identification of cell populations via scRNA-seq

Tissue specimens from 3 patients with ESCC were collected (Table 1), and scRNA-seq was performed to evaluate the types and proportion of different cell types in ESCC. A total of 14 clusters were identified based on gene expression (Fig. 1A). A UMAP plot demonstrating the distribution of gene expression in each sample was also created (Fig. 1B). A total of 20,324 cells were classified as epithelial cells, NK and T cells, endothelial cells, fibroblasts, B cells, monocytes and neutrophils, which were identified based on the expression of specific marker genes shown in the UMAP plot (Fig. 1C). For each cell type, three marker genes were selected (Figure S1), and the most significant marker gene was found to be widely expressed and distributed in its cell type (Fig. 1E). The number of endothelial cells, B cells and neutrophils was higher in T1, T2 and T3 samples, respectively (Fig. 1D).

Table 1
Characteristics of the 3 patients included in this study for scRNA-seq analysis and spatial transcriptomics

Patient number	Age	Gender	TNM stage	Smoking history	Drinking history	Family history of ESCC
T 1	76	Male	T2N3M0	Yes	Yes	No
T 2	43	Female	T3N0M0	No	Yes	No
T 3	60	Male	T1N2M0	Yes	No	Yes

ScRNA-seq: single-cell RNA sequencing; ESCC: esophageal squamous cell carcinoma

Spatial transcriptomic regionalisation of ESCC samples

To assess the spatial distribution of different cell types, ST analysis was performed on T1, T2 and T3 samples. ST sections of the three samples were stained with H&E (Fig. 2A–C). According to the ST sequencing data, spots in the spatial sections were divided into 10 regions (Fig. 2D–F), with each spatial region expressing different characteristic genes (Figure S2).

MIA(21), a hypergeometric distribution test method, was used to integrate the results of ST and scRNA-seq analyses. The association between the spatial regions and single-cell clusters was analysed to identify the major cell types distributed in each region (Fig. 2G–I). In addition, the spatial regions with epithelial cells as the dominant component were classified as cancer regions (almost all epithelial cells are cancer cells), whereas other regions were classified as stromal regions (Fig. 2J–L).

scRNA-seq and spatial transcriptomic analyses of epithelial cells, fibroblasts and endothelial cells

The spatial distribution of every cell type is different (Fig. 2G–I). The distribution of subsets of each cell type was compared between the cancer and stromal regions. The subsets of stromal cells (epithelial cells, fibroblasts and endothelial cells) were enriched in both regions. A total of 8 subsets (C0–C8) were identified for epithelial cells, with each subset having unique marker genes (Fig. 3A–C). The T1 sample (metastatic sample) was largely represented by C0, C1 and C4 subsets, whereas non-metastatic samples mostly contained other subsets (Fig. 3D). Most epithelial cell subsets, especially C5 and C7, were distributed in the cancer regions, whereas C4 was mostly distributed in the stromal regions, and C6 did not have a preference (Fig. 3E). GSVA revealed that NEIL3-mediated resolution of ICLs was significantly enriched in C5 and C7 subsets, indicating that DNA repair may be associated with tumorigenesis (Fig. 3F, Table S1). Furthermore, two main subtypes of fibroblasts, namely, inflammatory cancer-associated fibroblasts (iCAFs) and myofibroblasts (myCAFs), were identified based on specific marker genes (Fig. 4A, D). In addition, eight subsets were identified, and their distribution in the three patients was demonstrated in the UMAP plot (Fig. 4B-C). Each of the two subtypes had highly expressed genes, such as ACTA2 for myCAFs and CFD for iCAFs (Fig. 4E). The T1 sample was mostly enriched in iCAFs, whereas myCAFs did not have a preference (Fig. 4F, Table S2). iCAFs were significantly enriched in stromal regions, whereas

myCAFs were enriched in both cancer and stromal regions (Fig. 4G). The functions and pathways differed between these two subsets, such as vitamins in myCAFs and hydroxycarboxylic acid-binding receptors in iCAFs (Fig. 4H). Furthermore, the endothelial cell subsets were analysed in the three patients, which revealed three subsets with unique marker genes (Figure S3A-C). The C2 subset was mostly enriched in the T1 sample, whereas the C1 subset was highly enriched in non-metastatic samples (Figure S3D). In addition, the C2 and C0 subsets were mostly found in the cancer and stromal regions, respectively (Figure S3E). Similar pathways were found to be enriched as in the aforementioned cell types, such as NEIL3-mediated resolution of ICLs in the C2 subset and hydroxycarboxylic acid-binding receptors in the C0 and C1 subsets (Figure S3F, Table S3). These findings suggested that the distribution of stromal cells was not similar, and each cell type had its subset that could stratify cancer and stromal regions in ESCC.

scRNA-seq and spatial transcriptomic analyses of NK and T cells, B cells, monocytes and neutrophils

Based on the integration of scRNA-seq and ST, we analysed the distribution of some immune cells. However, the subsets of NK and T cells, B cells, monocytes and neutrophils were not enriched in either the cancer or stromal region. The following five main categories of NK and T cells were identified: CD4 + T cells, CD8 + T cells, regulatory T cells, NK cells and unspecified (Fig. 5A). The UMAP plots revealed 15 classifications and the distribution of these cell types in the 3 patients with ESCC (Fig. 5B-C) and demonstrated marker genes and highly expressed genes in each category (Fig. 5D-E). MIA revealed that exhausted CD4 + cells and cytotoxic CD8 + cells were mostly enriched in non-metastatic samples, whereas other cells were enriched in all three samples (Fig. 5F). Only proliferating CD4 + cells had a relative higher enrichment score in the cancer and stromal regions of all samples, whereas naïve CD4+, exhausted CD8 + and NKT cells were mainly concentrated in the cancer regions of the T1 sample (Fig. 5G). Some pathways were enriched in NK and T cells, such as intracellular oxygen transport in exhausted CD4 + cells (Fig. 5H, Table S4). A total of 6, 9 and 4 subsets were identified for B cells, monocytes and neutrophils, with highly expressed genes in each subset (Figure S4A-B, S5A-B, S6A-B). The C1–3 subsets of B cells and the C2 subset of neutrophils were mainly enriched in the T1 sample (Figure S4C, S6C), whereas most subsets of monocytes were enriched in non-metastatic samples (Figure S5C). In addition, MIA revealed that the characteristic genes of most subsets of these three cell types were mainly enriched in the cancer region of the T1 sample, indicating that these subsets were characterised by sample heterogeneity (Figure S4D, S5D, S6D). GSVA of each subgroup of these three cell types showed no significant difference (Figure S4E, S5E, S6E; Table S5, S6, S7).

Distribution of metastasis-associated epithelial cells in the spatial transcriptome

On analysing the ST data of the 3 patients, T1 was identified as the primary site in a patient with lymph node metastasis, and a metastatic sample (N1) was collected from this patient for scRNA-

seq. Combining the single-cell data of T1 and N1 samples, a total of 12,205 cells were divided into 7 cell types (Fig. 6A). In epithelial cells, 7 subsets were identified via cell subpopulation analysis (Fig. 6B). In both C1 and C5 subsets, some cells of the primary and metastatic origin were present. However, most other subsets belonged to a single sample. For pseudotime analysis (Fig. 6C), epithelial cells were divided into three evolutionary branches; of which, states 1 and 2 mainly included metastatic cells, whereas state 3 mainly included primary cells. State 1 contained a total of 2478 cells; of which, 1757 (70.90%) were of metastatic origin and 721 (29.10%) were of primary origin. State 2 contained a total of 1101 cells; of which, 756 (68.66%) were of metastatic origin and 345 (31.34%) were of primary origin. State 3 contained a total of 1028 cells; of which, 90 (8.75%) were of metastatic origin and 938 (91.25%) were of primary origin (Fig. 6D). These data were used to annotate spots in the spatial transcriptome. States 1 and 3 were annotated in the cancer region, whereas state 2 was annotated in the stromal region (Figs. 2J). Pseudotime analysis on T1 and N1 samples revealed that epithelial cells of N1 (metastasis) were mainly distributed on the right branch (Figure S8).

The spatial regions annotated by states 1 and 2 with concentrated metastatic cells were named ‘Cancer Region 1’ and ‘Stromal Region 1’, respectively. The remaining cancer and stromal regions were named ‘Cancer Region 2’ and ‘Stromal Region 2’, respectively (Fig. 6E). Cancer Region 1 was considered the treatment group, and Cancer Region 2 was considered the control group. The upregulated 119 DEGs in Cancer Region 1 were mainly enriched in pathways related to ‘metabolism of xenobiotics by cytochrome P450’ and ‘chemical carcinogenesis’, whereas the downregulated DEGs were mainly related to ‘ribosomal’ pathways. Among the downregulated genes, some genes were annotated to pathways related to ‘glutathione metabolism’, ‘carbon metabolism’ and ‘glycolysis/gluconeogenesis’ (p-value > 0.05 for all) (Figs. 6F, 6H, Table S8). In two stromal regions, 21 upregulated genes were enriched in the ‘oxidative phosphorylation’ pathway, whereas 59 downregulated genes were mainly enriched in pathways related to ‘ECM–receptor interaction’, ‘focal adhesion’ and ‘PI3K–Akt signalling’ (Figs. 6G, 6I, Table S9). The survival analysis of TCGA data showed that DEGs in both cancer and stromal regions were significantly related to overall survival (Fig. 6J-K), and higher expression of DEGs was related to poorer survival probability in both regions.

In addition, the functional enrichment of each spot was scored, and significant differences were observed in many tumour-related pathways, such as the Notch and TGF-β pathways, between the two cancer regions. The enrichment scores of these pathways were significantly higher in Cancer Region 1 than in Cancer Region 2. The enrichment of other pathways, including epithelial–mesenchymal transition (EMT) and angiogenesis, was significantly different between the two stromal regions, with higher enrichment in Stromal Region 2. (Figure S7A-B). Furthermore, the immune-infiltrating levels of different regions were evaluated using the CIBERSORT algorithm. The proportion of naïve CD4 + T cells, gammadelta T cells, M0 macrophages and eosinophils was higher in Cancer Region 2, whereas that of activated NK cells was higher in Cancer Region 1. As for the stromal regions, most immune cells had higher infiltration in Stromal Region 2, whereas only plasma cells had higher infiltration in Stromal Region 1 (Figure S7C-D).

Discussion

In this study, we identified subpopulations of seven cell types and addressed their spatial map in three heterogeneous ESCC samples. We integrated scRNA-seq and ST analyses using MIA to evaluate the enrichment score of different cell populations across the cancer and stromal regions, revealing the comprehensive map of various cell types and subsets in the TME. In addition, we evaluated the correlation of distinct cell subpopulations with the stromal and cancer regions in ESCC. The findings revealed heterogeneity between the primary and metastatic sites, providing meaningful biological insights into the mechanisms of ESCC metastasis.

TME is a complex tissue environment for cancer development and progression, which contains various types of cells, including tumour, immune and stromal cells (25). Compared with tumour cells within the TME, stromal cells have genetic stability and are considered a promising therapeutic biomarker for cancer treatment (26). In addition, immune cells within the TME play a key role in tumorigenesis and metastasis and may be important for immunotherapy effectiveness as evidenced by the immunosuppressive mechanisms possibly associated with Treg–macrophage interactions in ESCC (27). Chen et al. used scRNA-seq to show that the TME of ESCC is heterogeneous, demonstrating substantial differences in stromal cells between tumour and normal tissues, which may contribute to carcinogenesis (28). However, the main restrictions for TME in cancer treatment are temporal and spatial differences, which largely contribute to the heterogeneity of cancer (29). To the best of our knowledge, this study is the first to present a spatial landscape of multiple cell subpopulations in ESCC, indicating its intra- and inter-tumoral heterogeneity, which may provide novel insights into the investigation of efficient therapies. MIA revealed that cell subpopulations were heterogeneous in the tumour and stromal subregions among all samples. The number and variety of cells were much greater in the cancer region of the T1 sample and stromal region of the T3 sample.

Stromal cells were more significantly enriched in the TME of ESCC, suggesting their more active role in tumorigenesis and metastasis. A recent study demonstrated that more epithelial cells and fibroblasts were found in ESCC tumour tissues and the adjacent normal tissues, respectively, which is consistent with the results of this study (30). EMT detaches epithelial cells and promotes metastasis and therapeutic resistance, indicating the potential value of stromal cells of TME in the development of novel therapies (31). The inflammatory and remodelling processes are regulated by the interactions between oesophageal epithelial cells and eosinophils, which may drive various tumorigenic phenotypes in ESCC (32, 33). In this study, most subtypes of epithelial cells were more abundant in cancer regions than in stromal regions. Two subsets of epithelial cells were substantially enriched in cancer regions and unrelated to stromal regions, which may be considered powerful indicators for ESCC. CAFs are important for tumorigenesis and are a heterogeneous component within the TME (34). In this study, two CAF subsets—iCAFs and myCAFs—were identified. iCAFs were mostly clustered in the stromal regions, whereas no difference was found in the distribution of myCAFs between the cancer and stromal regions. Fang et al. found that CXCL1 promotes the formation of iCAFs via the CXCR2–pSTAT3 pathway, contributing to the progression of ESCC (35). In addition, IL-6 regulates the interaction between

fibroblasts and tumour cells within the TME in ESCC (36). We hypothesised that iCAFs interact with tumour cells through specific factors in the stromal region instead of direct interaction in the tumour region. In addition, some specific pathways enriched in iCAF subpopulations may be candidates for future research, such as ‘NTF3-activated NTRK3 signalling’ and ‘COX reactions’. Non-immune stromal and endothelial cells were heterogeneous among the three ESCC samples in this study and did not have a clear preference for tumour or stromal subregions.

Immune cells within the TME play a key role in cancer progression by interacting with tumour cells by secreting different chemokines, cytokines and other signalling molecules (37). In this study, compared with non-immune cells, immune cell subpopulations were heterogeneous in ESCC, and most subsets were enriched in the cancer region of the T1 sample. TME components can promote EMT, invasion and angiogenesis to facilitate metastasis in various cancers, such as oesophageal (10), pancreatic (38) and breast (39) cancers. In this study, T1 sample was obtained from the primary site of metastatic ESCC, whereas the other two samples were obtained from primary ESCC. The types and proportion of immune cells within the TME in metastatic ESCC were different from those in primary cancer, including naïve CD4 + T cells, neutrophils and B cells, and might affect cancer cells in different ways. Gu et al. showed that tumour-educated B cells secrete HSPA4-targeting IgG and subsequently facilitate the metastasis of breast cancer (40). In addition, neutrophils can interact with L-17-producing γδ T cells and promote metastasis in breast cancer (41). We conducted pseudotime analysis using T1 and N1 samples and presented the spatial features of cancer and stromal regions to determine DEGs and different pathways between the primary and metastatic regions. A recent study used scRNA-seq analysis and revealed differences in cells between the primary and metastatic sites of head and neck cancer (42). In this study, combined with scRNA-seq and spatial information, pathways related to metabolism and tumorigenesis were found to be enriched at metastatic sites, showing the dynamic activity of tumour and stromal cells during metastasis. In addition, patients with higher expression of DEGs had substantially unfavourable clinical outcomes, indicating that these metastatic genes were related to a poor prognosis of ESCC. Therefore, cells within the TME play a significant role in cancer metastasis, which may serve as prognostic predictors and provide new insights into the investigation of effective therapies for ESCC.

Although this study is the first to reveal the spatial features of various cell subsets in ESCC and provide a widely applicable method to comprehensively map an entire tissue using MIA, it has some limitations. First, the number of patients enrolled in this study was limited. Because this was an exploratory and the first study on ESCC in this field, it was inappropriate to enrol more patients. Based on the findings, we believe that more patients can be enrolled for further study, and more distinct TME features in ESCC can be identified. Second, the resolution of ST technology remains to be a shortcoming. The size of the ST array may not be sufficient to cover the whole tissue, and ST arrays do not achieve comparable resolution for every spot at the single-cell scale. In addition, the transcriptomic data can be only accessed within the cells of each spot and are lost at intervals between every two spots. With the development of ST technology, higher resolution and shorter interval distance may be achieved in the future.

Conclusions

With the integration of scRNA-seq and ST analyses, this study presents the first comprehensive spatial landscape of different cell subpopulations in ESCC, indicating the intratumour heterogeneity of TME features and tumours. Transcriptomic diversities and distinct spatial patterns were detected between non-metastatic and metastatic samples. The findings show that the precise composition and spatial landscape of tumours may vary from person to person, which may provide novel insights into the discovery of prognostic factors and development of effective therapeutic interventions.

Abbreviations

TME, tumor microenvironment

scRNA-seq, single-cell RNA sequencing

ST, spatial transcriptomics

MIA, multimodal intersection analysis

DEGs, differentially expressed genes

ESCC, Esophageal squamous cell carcinoma

EAC, esophageal adenocarcinoma

H&E, hematoxylin and eosin

EMT, epithelial-mesenchymal transition

CAF, Cancer-associated fibroblasts

Declarations

Acknowledgments

We sincerely thank the staff of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, for their support during the study.

Competing interests

Xiang Liu is employee of Echo Biotech Co., Ltd. All remaining authors report no conflict of interest.

Ethics approval and consent to participate

This study was performed in accordance with the Declaration of Helsinki and was approved by the National Cancer Center/Cancer Hospital Ethics Committee. All patients provided written informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The data that supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Funding

This study was supported by the National Key R&D Program of China (2021YFC2500900), the National Natural Science Foundation of China (82002451), the CAMS Initiative for Innovative Medicine (2021-1-I2M-015,), the Special Research Fund for Central Universities, Peking Union Medical College (3332020024), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2018PT32033), and the Beijing Hope Run Special Fund of Cancer Foundation of China (LC2019B15).

Authors' contributions

JH, SG, and QX supervised the project, design, interpretation, manuscript revision, and final approval of the version to be submitted. WG, BZ, and XL were involved in conceptualization, data acquisition, analysis, and interpretation. WG and BZ wrote the first draft. WG, QH, LG, and XX prepared the figures and tables. FT, BZ, YL and QX reviewed all specimens enrolled in the study. WG, BZ, XX and FT acquired data and provided material support. All authors reviewed and approved the final manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Lin Y, Totsuka Y, He Y, Kikuchi S, Qiao Y, Ueda J, et al. Epidemiology of esophageal cancer in Japan and China. J Epidemiol. 2013;23(4):233–42.

3. Arnold M, Laversanne M, Brown LM, Devesa SS, Bray F. Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030. Am J Gastroenterol. 2017;112(8):1247–55.
4. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet. 2017;390(10110):2383–96.
5. Lagergren J, Lagergren P. Recent developments in esophageal adenocarcinoma. CA Cancer J Clin. 2013;63(4):232–48.
6. Shah MA, Hofstetter WL, Kennedy EB. Immunotherapy in Patients With Locally Advanced Esophageal Carcinoma: ASCO Treatment of Locally Advanced Esophageal Carcinoma Guideline Rapid Recommendation Update. J Clin Oncol. 2021;39(28):3182–4.
7. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. Lancet. 2013;381(9864):400–12.
8. Luo H, Lu J, Bai Y, Mao T, Wang J, Fan Q, et al. Effect of Camrelizumab vs Placebo Added to Chemotherapy on Survival and Progression-Free Survival in Patients With Advanced or Metastatic Esophageal Squamous Cell Carcinoma: The ESCORT-1st Randomized Clinical Trial. Jama. 2021;326(10):916–25.
9. Yan T, Cui H, Zhou Y, Yang B, Kong P, Zhang Y, et al. Multi-region sequencing unveils novel actionable targets and spatial heterogeneity in esophageal squamous cell carcinoma. Nat Commun. 2019;10(1):1670.
10. Lin EW, Karakasheva TA, Hicks PD, Bass AJ, Rustgi AK. The tumor microenvironment in esophageal cancer. Oncogene. 2016;35(41):5337–49.
11. Gohil SH, Iorgulescu JB, Braun DA, Keskin DB, Livak KJ. Applying high-dimensional single-cell technologies to the analysis of cancer immunotherapy. Nat Rev Clin Oncol. 2021;18(4):244–56.
12. Rodda LB, Lu E, Bennett ML, Sokol CL, Wang X, Luther SA, et al. Single-Cell RNA Sequencing of Lymph Node Stromal Cells Reveals Niche-Associated Heterogeneity. Immunity. 2018;48(5):1014-28.e6.
13. Aoki T, Chong LC, Takata K, Milne K, Hav M, Colombo A, et al. Single-Cell Transcriptome Analysis Reveals Disease-Defining T-cell Subsets in the Tumor Microenvironment of Classic Hodgkin Lymphoma. Cancer Discov. 2020;10(3):406–21.
14. Friebel E, Kapolou K, Unger S, Núñez NG, Utz S, Rushing EJ, et al. Single-Cell Mapping of Human Brain Cancer Reveals Tumor-Specific Instruction of Tissue-Invading Leukocytes. Cell. 2020;181(7):1626-42.e20.
15. Obradovic A, Chowdhury N, Haake SM, Ager C, Wang V, Vlahos L, et al. Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. Cell. 2021;184(11):2988–3005.e16.
16. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. Cell. 2020;181(2):442 – 59.e29.

17. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. Cell. 2020;182(5):1232-51.e22.
18. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol. 2015;33(5):495–502.
19. Thrane K, Eriksson H, Maaskola J, Hansson J, Lundeberg J. Spatially Resolved Transcriptomics Enables Dissection of Genetic Heterogeneity in Stage III Cutaneous Malignant Melanoma. Cancer Res. 2018;78(20):5970–9.
20. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level. Cancer Discov. 2022;12(1):134–53.
21. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat Biotechnol. 2020;38(3):333–42.
22. Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat Neurosci. 2021;24(3):425–36.
23. Asp M, Giacomello S, Larsson L, Wu C, Fürth D, Qian X, et al. A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. Cell. 2019;179(7):1647-60.e19.
24. La Manno G, Siletti K, Furlan A, Gyllborg D, Vinsland E, Mossi Albiach A, et al. Molecular architecture of the developing mouse brain. Nature. 2021;596(7870):92–6.
25. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med. 2013;19(11):1423–37.
26. Bejarano L, Jordão MJC, Joyce JA. Therapeutic Targeting of the Tumor Microenvironment. Cancer Discov. 2021;11(4):933–59.
27. Zheng Y, Chen Z, Han Y, Han L, Zou X, Zhou B, et al. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. Nat Commun. 2020;11(1):6268.
28. Chen Z, Zhao M, Liang J, Hu Z, Huang Y, Li M, et al. Dissecting the single-cell transcriptome network underlying esophagus non-malignant tissues and esophageal squamous cell carcinoma. EBioMedicine. 2021;69:103459.
29. Prasetyanti PR, Medema JP. Intra-tumor heterogeneity from a cancer stem cell perspective. Mol Cancer. 2017;16(1):41.
30. Zhang X, Peng L, Luo Y, Zhang S, Pu Y, Chen Y, et al. Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. Nat Commun. 2021;12(1):5291.
31. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. Nat Rev Mol Cell Biol. 2019;20(2):69–84.
32. Dunn JLM, Caldwell JM, Ballaban A, Ben-Baruch Morgenstern N, Rochman M, Rothenberg ME. Bidirectional crosstalk between eosinophils and esophageal epithelial cells regulates inflammatory and remodeling processes. Mucosal Immunol. 2021;14(5):1133–43.

33. Monkkonen T, Debnath J. Inflammatory signaling cascades and autophagy in cancer. Autophagy. 2018;14(2):190–8.
34. Biffi G, Tuveson DA. Diversity and Biology of Cancer-Associated Fibroblasts. Physiol Rev. 2021;101(1):147–76.
35. Fang L, Che Y, Zhang C, Huang J, Lei Y, Lu Z, et al. LAMC1 upregulation via TGFβ induces inflammatory cancer-associated fibroblasts in esophageal squamous cell carcinoma via NF-κB-CXCL1-STAT3. Mol Oncol. 2021;15(11):3125–46.
36. Karakasheva TA, Lin EW, Tang Q, Qiao E, Waldron TJ, Soni M, et al. IL-6 Mediates Cross-Talk between Tumor Cells and Activated Fibroblasts in the Tumor Microenvironment. Cancer Res. 2018;78(17):4957–70.
37. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. Cancer Res. 2019;79(18):4557–66.
38. Ren B, Cui M, Yang G, Wang H, Feng M, You L, et al. Tumor microenvironment participates in metastasis of pancreatic cancer. Mol Cancer. 2018;17(1):108.
39. Deepak KGK, Vempati R, Nagaraju GP, Dasari VR, S N, Rao DN, et al. Tumor microenvironment: Challenges and opportunities in targeting metastasis of triple negative breast cancer. Pharmacol Res. 2020;153:104683.
40. Gu Y, Liu Y, Fu L, Zhai L, Zhu J, Han Y, et al. Tumor-educated B cells selectively promote breast cancer lymph node metastasis by HSPA4-targeting IgG. Nat Med. 2019;25(2):312–22.
41. Coffelt SB, Kersten K, Doornebal CW, Weiden J, Vrijland K, Hau CS, et al. IL-17-producing γδ T cells and neutrophils conspire to promote breast cancer metastasis. Nature. 2015;522(7556):345–8.
42. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell. 2017;171(7):1611-24.e24.

Figures

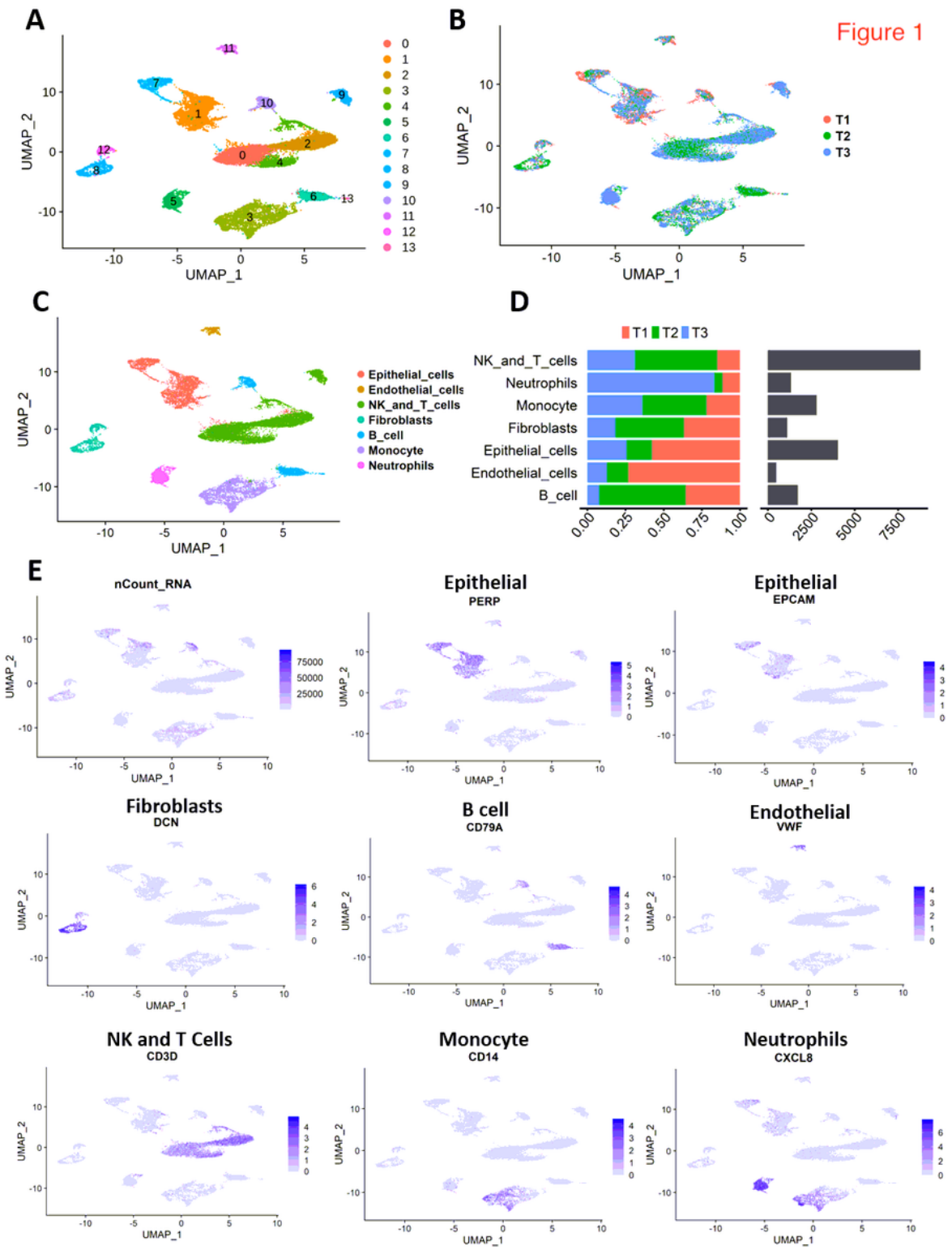

Figure 1

The single cell atlas of ESCC patients (n=3). A. Single cell clustering results of three primary samples based on the expression levels of genes. **B.** The UMAP presentation of the distribution of the three samples. **C.** The UMAP presentation of major annotated cell types according to the expression of selected marker genes. The clusters are annotated with various colors based on specific identities of different cell

types. **D.** Number of cells and sample proportion of each cell type. **E.** The expression of marker genes in different cell types.

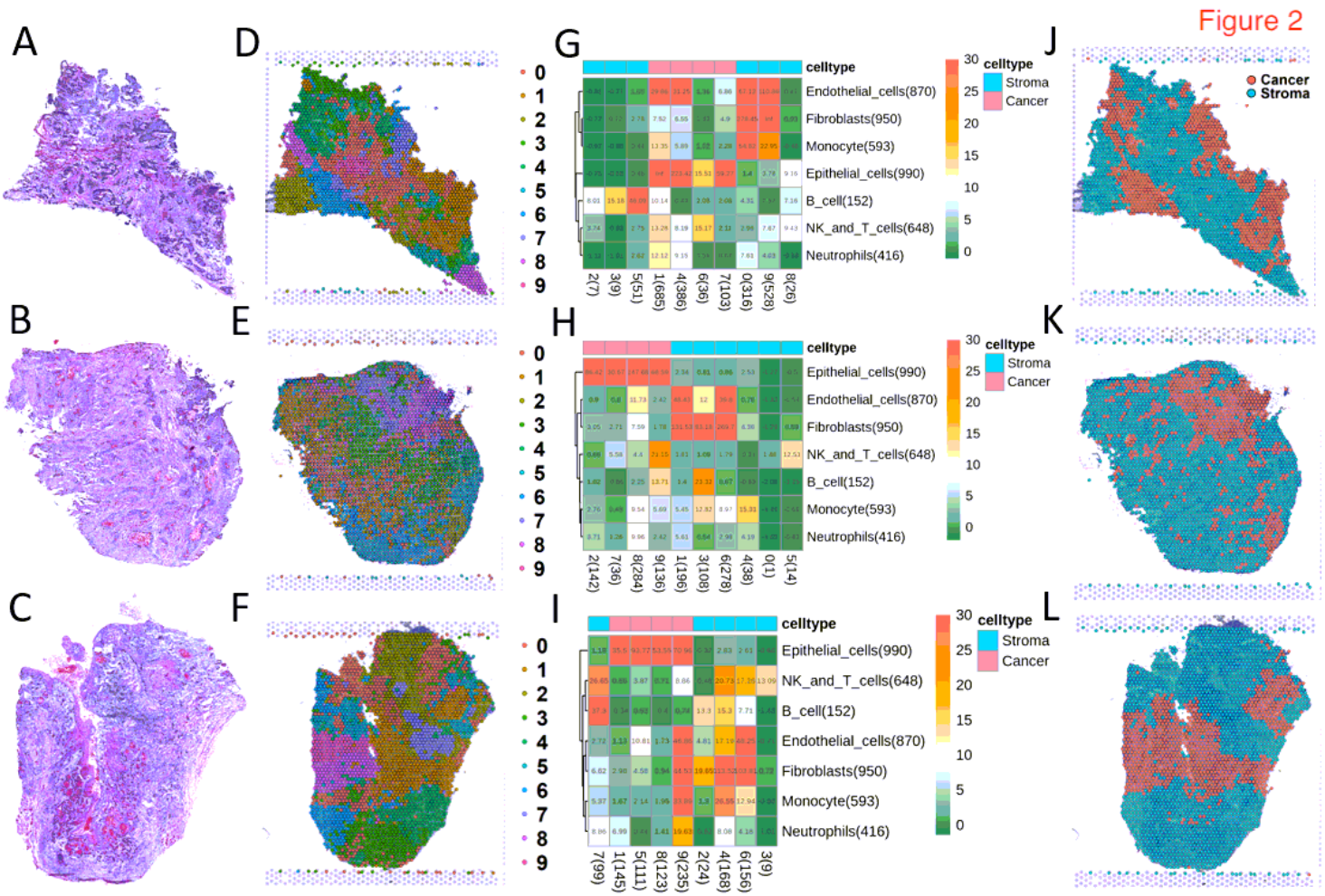

Figure 2

Spatial transcriptome analysis of ESCC and regionalization of cancer and stroma. A-C. H&E staining of transcriptome sections of three samples. (T1, T2 and T3, in turn). **D-F.** The results of spots clustering in three spatial transcriptome data. The numbers of spots in three samples are 2353 (T1), 2849 (T2) and 2896 (T3). **G-I.** Results of correlation significance analysis of spatial transcriptome regions and cell subsets of single-cell data using MIA method. The higher the value, the greater the proportion of highly expressed genes shared by spatial transcriptome regions and cell subsets. For each spatial region, it was classified as the cell type with the highest value. Finally, all spatial regions were divided into cancer regions and stroma regions. **J-L.** According to the analysis results of MIA, all spots was divided into cancer region and stroma region. The red is the area of cancer, and the blue is the area of stroma.

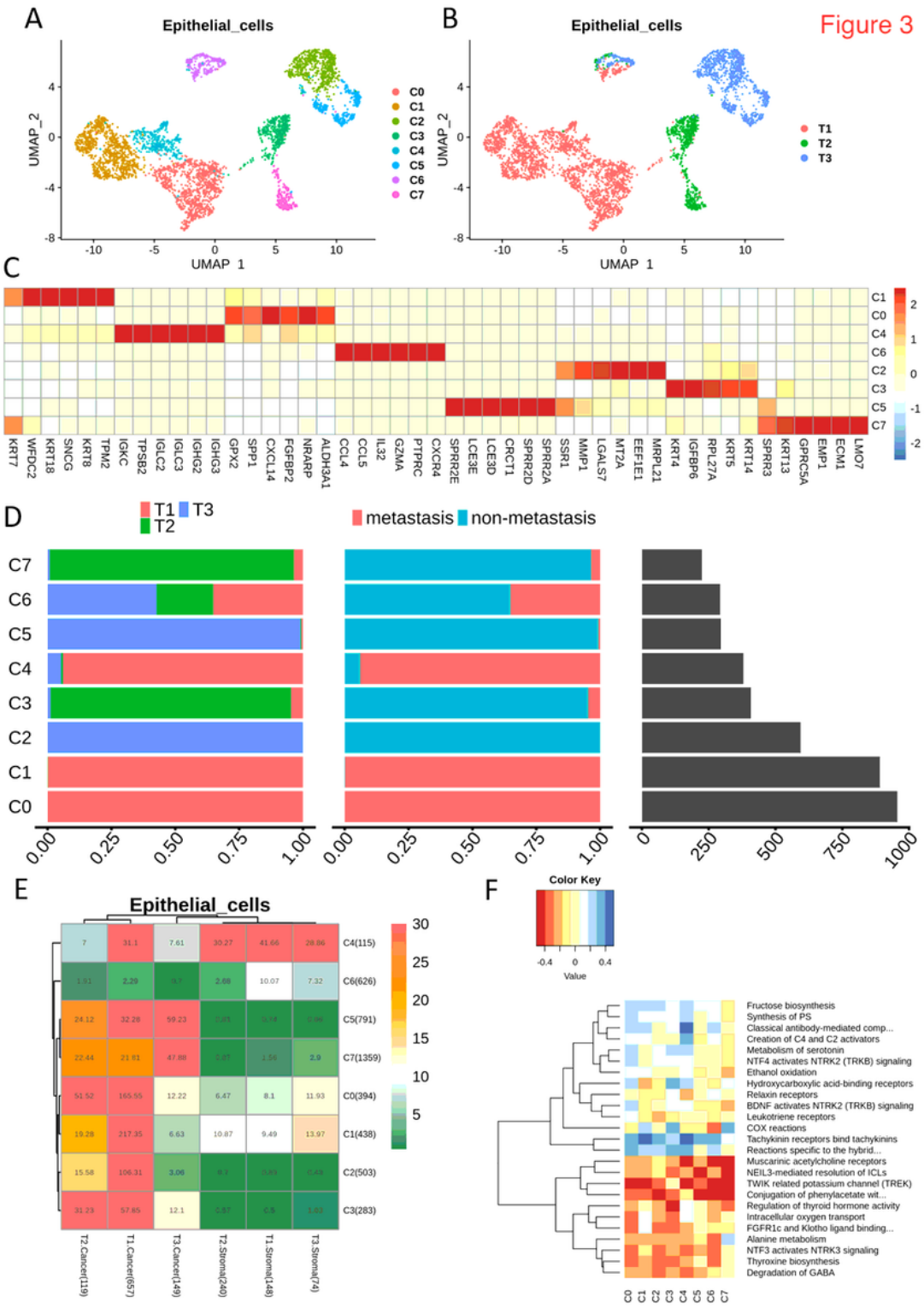

Figure 3

Subsets analysis and MIA mapping of epithelial cell. A. Identification of subgroups of epithelial cells in ESCC. **B.** The distribution of epithelial cells in three ESCC patients. **C**. Characteristic gene expression heatmap of each subgroup of epithelial cell. **D**. Cell number and sample proportion of each cell subsets. **E.** MIA results of epithelial cells in cancer and stroma of spatial transcriptome. **F**. GSVA results of different cell subsets.

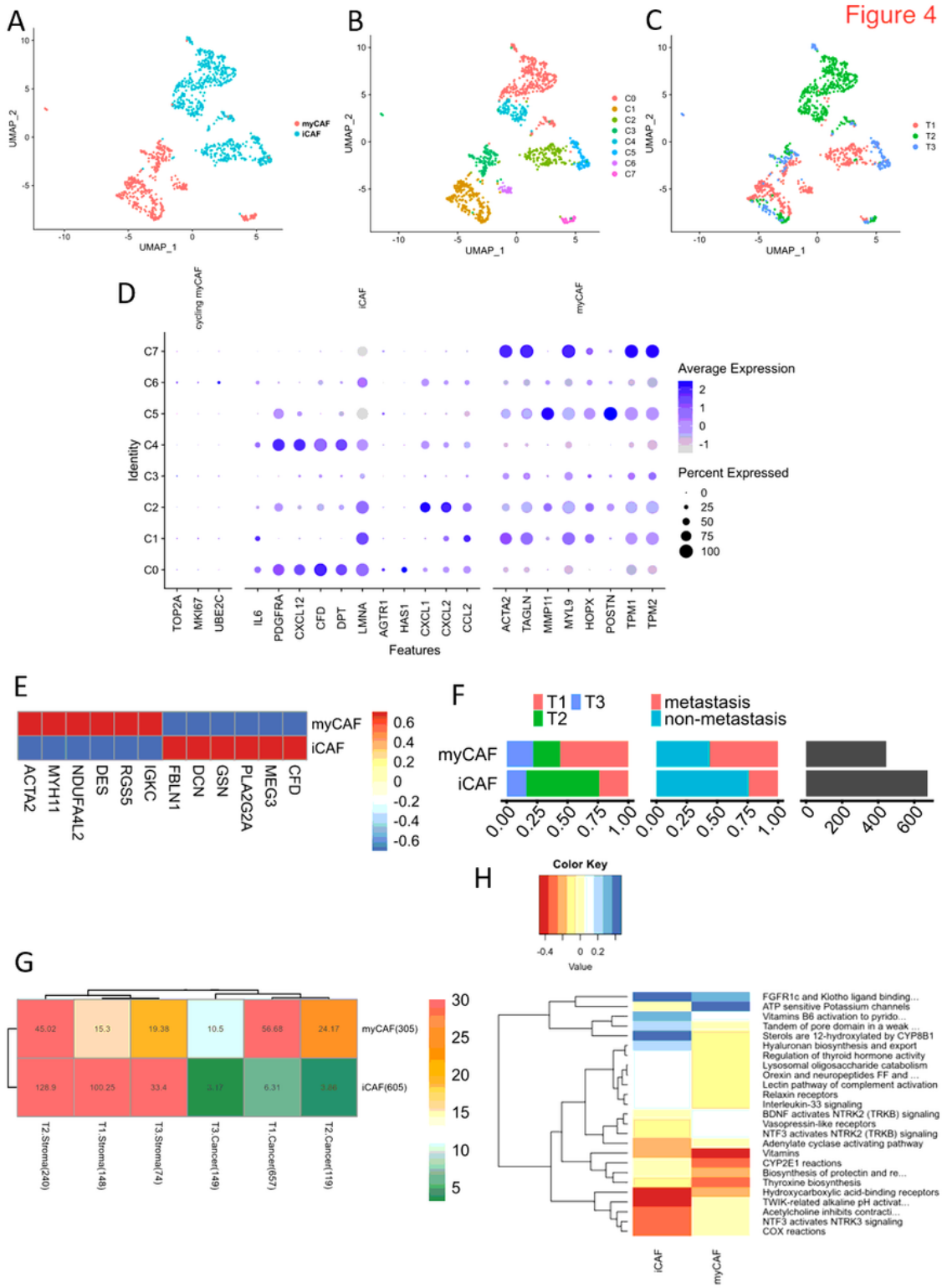

Figure 4

Subsets analysis and MIA mapping of fibroblast. A. Identification of two main subgroups of fibroblast in ESCC. iCAF: inflammatory fibroblast; myCAF, myofibroblast. B. Single cell clustering results of fibroblast subsets. C. The distribution of fibroblast in three ESCC patients. D. Dot plot of marker gene expression in fibroblast of different subsets. E. The highly expressed characteristic gene heatmap of two main fibroblast subgroup. F. Number of cells in fibroblast subsets and three samples proportion. G. MIA results

of the fibroblast subsets distribution in the spatial transcriptome stroma and cancer regions. H. GSVA of fibroblast subsets.

Figure 5

Subsets analysis and MIA mapping of NK and T cell.

A. Identification of subgroups of NK and T cells in ESCC. Five main categories are divided: CD4 + T cell, CD8 + T cell, regulatory T cell, NK cell and unspecified. **B.** Single cell clustering results of NK and T cell subsets. **C.** The distribution of NK and T cells in three ESCC patients. **D.** Dot plot of marker gene expression in NK and T cells of different subsets. **E.** The highly expressed characteristic gene heatmap of NK and T cell subsets. **F.** Number of cells in NK and T cell subsets and three samples proportion. **G.** MIA results of the NK and T cell subsets distribution in the spatial transcriptome stroma and cancer regions. **H.** GSVA results of NK and T cell subsets.

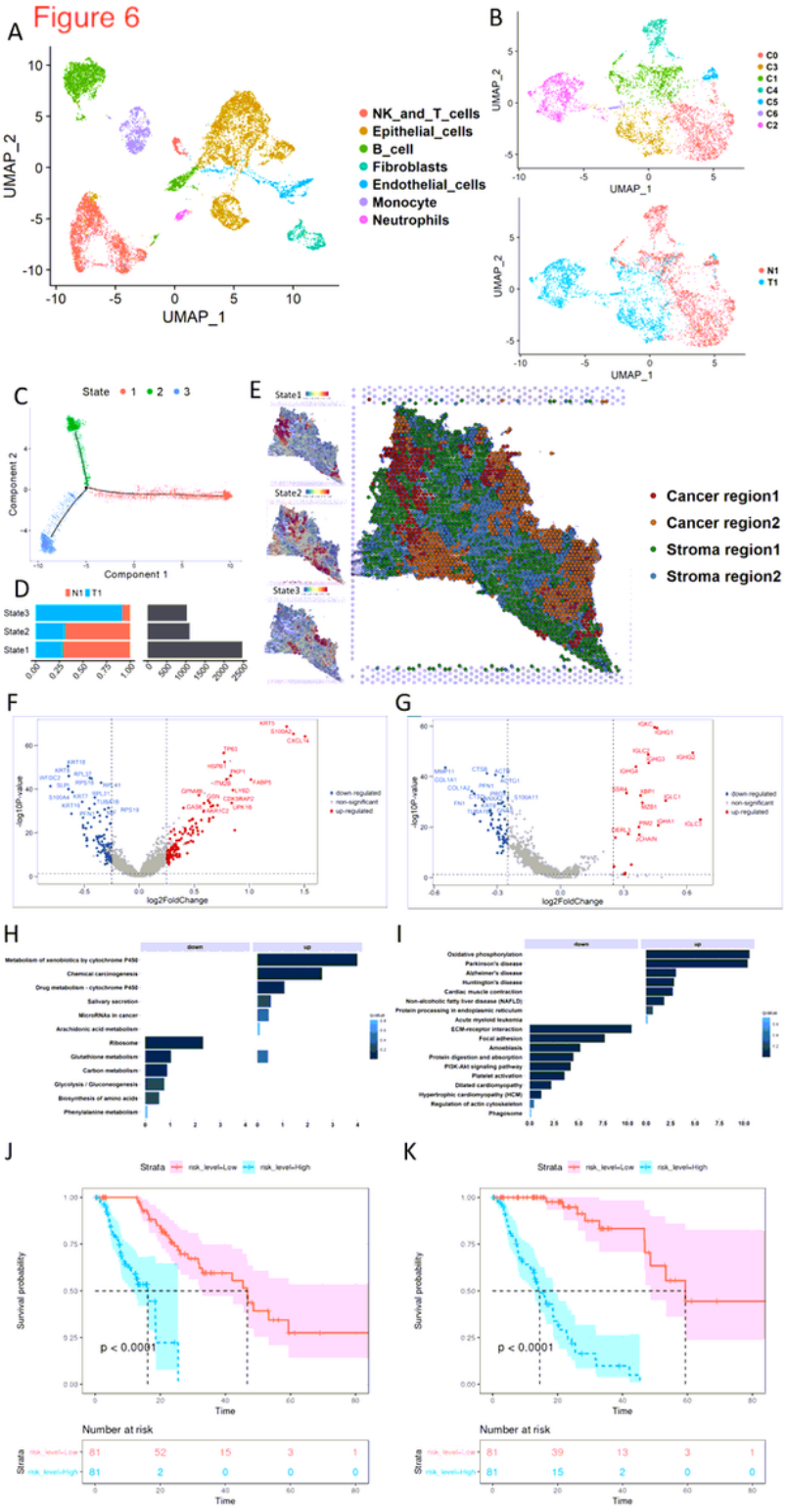

Figure 6

Metastasis related epithelial cells distribution in spatial transcriptome. A. Single-cell clustering and cell type annotated results of T1 (primary) and N1 (lymph node metastases) samples. **B.** Clustering results of epithelial cell subsets in T1 and N1 samples. **C.** Pseudotime analysis of epithelial cells showed that epithelial cells were divided into three evolutionary branches. **D.** Proportion of primary and lymph node metastatic cells in 3 states of epithelial cells. **E.** Three states of epithelial cells were used to annotate the

spatial transcriptome spots, which was spatially divided into four regions. Cancer region 1 represents the region where epithelial cells state 1 is concentrated in. Stroma region 1 is the region where epithelial cells state 2 is concentrated in. The remaining cancer and stroma spots were subdivided into cancer region 2 and stroma region 2. **F.** Volcano plot of differentially expressed genes (DEGs) in cancer region 1 versus cancer region 2. **G.** Volcano plot of DEGs in stroma region 1 versus stroma region 2. **H.** Function enrichment of DEGs in cancer region 1 versus cancer region 2. Top 6 enrichment pathways of up and down regulated genes were shown, respectively. **I.** Function enrichment of DEGs in stroma region 1 versus stroma region 2. **J**. Survival analysis results of DEGs in cancer region 1 and cancer region 2 in TCGA data. **K.** Survival analysis results of DEGs in matrix region 1 and matrix region 2 in TCGA data.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFigures.docx
- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- TableS8.xlsx
- TableS9.xlsx