Check for updates

Single-cell RNA sequencing reveals distinct tumor microenvironment of ground glass nodules and solid nodules in lung adenocarcinoma

Xiaofeng Huang†, Zhimeng Lu†, Xuewei Jiang, Zhe Zhang, Kun Yan and Guiping Yu*

Department of Cardiothoracic Surgery, Jiangyin Clinical College of Xuzhou Medical University, Jiangyin, China

Introduction: Lung adenocarcinoma (LUAD) is the most prevalent lung cancer. LUAD presents as ground glass nodules (GGN) and solid nodules (SN) in imaging studies. GGN is an early type of LUAD with good prognosis. However, SN exhibits a more malignant behavior than GGN, including worse pathological staging and tumor prognosis. The mechanism leading to the different malignancy levels of GGN and SN remains elusive.

Methods: Three patients with GGN and three patients with SN diagnosed with early LUAD were enrolled. The tumor samples were digested to a single-cell suspension and analyzed using 10× Genomic Single-cell ribonucleic acid sequences (scRNA-seq) techniques.

Results: A total of 15,902 cells were obtained and classified into nine major types. The tumor microenvironment (TME) was subsequently described in detail. ScRNA-seq revealed that ribosome-related pathways and cell adhesion played similar but distinct roles in the two groups. SN also had more active cell proliferation, enriched cell cycle regulatory pathways, and severe inflammatory responses.

Conclusion: We observed changes in the cellular composition and transcriptomic profile of GGN and SN. The study improved the understanding of the underlying mechanisms of lung carcinogenesis and contributed to lung cancer prevention and treatment.

KEYWORDS

single-cell RNA sequencing, lung adenocarcinoma, ground glass nodule, solid nodule, tumor microenvironment, macrophage

Introduction

Chest computed tomography (CT) has dramatically increased the detection of early-stage lung adenocarcinoma (LUAD), which have a radiographic appearance of ground glass nodules (GGN) (Travis et al., 2011). Compared to solid nodules (SN), GGNs grow more slowly and have a better prognosis (Zhang et al., 2020). Solid components in pulmonary nodules are closely associated with patient prognosis, with a higher proportion of solids indicating a worse pathological staging and tumor prognosis (Travis et al., 2016). The tumor microenvironment (TME) plays a crucial role in shaping tumor biological and clinical

behaviors in addition to cancer cells (Quail and Joyce, 2013). However, whether the TME in GGN is similar to the TME tissue components or invasive components in SN remains unknown. Moreover, the transition of the TME in GGN to the TME in SN has yet to be established. The developing single-cell ribonucleic acid (RNA) sequencing (scRNA-seq) technology can now provide an unbiased transcriptome analysis of single cells and their genetic heterogeneity (Papalexi and Satija, 2018). ScRNA-seq was performed on samples from three patients with GGN and three with SN diagnosed as LUAD. We characterized the GGN and SN ecosystems, identifying differences in their cell types and expression of molecular signatures, providing insightful biological information for further research.

Materials and methods

Patients and tissue samples

Tissue samples were collected from six patients who underwent lung surgery at the Department of Thoracic Surgery, Jiangyin Clinical College of Xuzhou Medical University, from January 2020 to December 2020. This study included the following inclusion criteria (Travis et al., 2011): the imaging presentation of GGN in three patients and of SN in the other three patients (Zhang et al., 2020), no lesions other than pulmonary nodule (Travis et al., 2016), no preoperative antitumor treatment (Quail and Joyce, 2013), pathological examination suggestive of LUAD. The detailed clinical information is presented in Supplementary Table S1. This study was approved by the Ethics Committee of Jiangyin Clinical College of Xuzhou Medical University [approval No.2019ER (027)]. All participants provided written informed consent.

Preparation of single-cell suspensions

Fresh tissues were stored on ice in the sCelLive™ Tissue Preservation Solution (Singleron) within 30 min after surgery. Hanks Balanced Salt Solution (HBSS) was used three times to wash the specimens, mince them into small pieces, and digest them in 3 mL sCelLive™ Tissue Dissociation Solution (Singleron) by Singleron PythoN™ Tissue Dissociation System at 37°C for 15 min. A 40-micron sterile strainer was used to collect and filter the cell suspension. After adding the GEXSCOPE® red blood cell lysis buffer (RCLB, Singleron), the mixture [Cell: RCLB = 1:2 (volume ratio)] was incubated for 5–8 min at room temperature to remove red blood cells. After centrifuging at $300 \times g$ 4°C for 5 min, the mixture was suspended in PBS softly after centrifugation. Finally, Trypan Blue staining was applied to the samples, and cell viability was assessed microscopically.

RT & amplification & library construction

Single-cell suspensions (2×10^5 cells/mL) with PBS (HyClone) were placed onto a microwell chip using the Singleron Matrix® Single Cell Processing System. Reverse transcription of the mRNA captured by the Barcoding Beads is followed by PCR amplification of the cDNA generated from the Barcoding Bead collection. Sequencing adapters are then ligated to the fragmented cDNA. GEXSCOPE® Single Cell RNA Library Kits (Singleron) were used to construct the scRNA-seq libraries (Dura et al., 2019). Illumina novaseq 6,000 was used to sequence 150 bp paired end reads from individual libraries diluted to 4 nM, pooled, and processed.

Primary analysis of raw read data

Gene expression matrices were generated using CeleScope (https://github.com/singleron-RD/CeleScope) v1.9.0 pipeline from raw scRNA-seq reads. Raw reads were first processed with CeleScope to remove low-quality reads followed by Cutadapt v1. 17 (Martin, 2011) to trim poly-A tails and adapters. Next, cell barcodes and UMI were extracted. Next, we mapped reads to the reference genome GRCh38 (Ensembl version 92 annotation) using STAR v2.6.1a (Dobin et al., 2013). Finally, each cell's UMIs and gene counts have been acquired with featureCounts v2.0.1 (Liao et al., 2014) software and used to generate expression matrix files for subsequent analysis.

Quality control, dimension-reduction, and clustering

Python 3.7 was used to perform quality control, dimension reduction, and clustering using scanpyv1.8.2 (Wolf et al., 2018). The expression matrix for each sample dataset was filtered by the following criteria: 1) cells with a gene count less than 200 or with a top 2% gene count were excluded; 2) cells with a top 2% UMI count were excluded; 3) cells with mitochondrial content >50% were excluded; 4) genes expressed in less than 5 cells were excluded. There were 15,902 cells retained after filtering for downstream analyses, with each cell having on average 1,217 genes and 3,433 UMIs. A normalized data matrix was generated by normalizing the raw count matrix by the total number of counts per cell. The top 2000 variable genes were selected by setting flavor = 'Seurat'. Principle Component Analysis (PCA) was performed on the scaled variable gene matrix, and the top 20 principle components were used to reduce the dimensions and cluster the genes. Cells were split into 21 clusters using the Louvain algorithm with a resolution parameter of 1.2. Cell clusters were visualized by using Uniform Manifold Approximation and Projection (UMAP) {t-Distributed Stochastic Neighbor Embedding (t-SNE)}.

Statistics and repeatability

Two-tailed Wilcoxon rank-sum tests were used to compare cell distributions between the two groups. Student's t-test was used to compare gene expression or gene signatures between two groups of cells. Two-tailed Wilcoxon rank-sum tests were performed to compare cell distributions between paired groups 1 and 2. Statistical analyses and presentations were conducted using R. Statistical tests used in figures were shown in figure legends, and statistical significance was set at $p < 0.05$. The exact value of n was

shown in the figures, and figure legends and what n represents was shown in the figure legends.

Differentially expressed genes (DEGs) analysis

To identify differentially expressed genes (DEGs), we used the Seurat FindMarkers function based on the Wilcox likelihood-ratio test with default parameters. We selected the genes expressed in more than 10% of the cells in a cluster, with an average log (Fold Change) value greater than 0.25 as DEGs. The cell types of each cluster were annotated using the DEGs and knowledge from the literature as canonical markers. The expression of markers in each cell type was displayed using heatmaps/dot plots/violin plots generated with the Seurat DoHeatmap/DotPlot/Vlnplot functions. Doublet cells were identified as expressing markers for different cell types and removed manually.

Cell type annotation

SynEcoSys database was used to identify the cell type identity of each cluster based on canonical markers found in DEGs. In addition, Seurat 3.1.2 was used to generate heatmaps/dot plots/violin plots that illustrated the expression of markers used to identify each cell type.

Pathway enrichment analysis

The "clusterProfiler" R package 3.16.1 was used to analyze Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) data to investigate DEG functions (Yu et al., 2012). Pathways with p_adj value less than 0.05 were considered significantly enriched. Gene Ontology gene sets, including molecular function (MF), biological process (BP), and cellular component (CC) categories, were used as references. GSVA pathway enrichment analysis used the average gene expression of each cell type as input data (Hänzelmann et al., 2013).

UCell gene set scoring

Gene set scoring was performed using the R package UCell v 1.1.0 (Andreatta and Carmona, 2021). Based on the Mann-Whitney U statistic, UCell scores rank query genes according to their expression levels in individual cells. The rank-based scoring method of UCell is suitable for large datasets that contain multiple samples and batches.

scRNA-seq-based CNA detection

The InferCNV package detected the CNAs in malignant cells (cutoff = 0.1, denoise = TRUE, HMM = F, and k_obs_groups = 8) (Kumar et al., 2020). The CNAs of malignant cells were estimated using T cells as baselines. The genes expressed in over 20 cells were sorted according to their loci on each chromosome. Relative expression values were centered at 1 using a 1.5 standard deviation from residual-

normalized expression values. The relative expression of each chromosome was smoothened using a sliding window of 101 genes to remove the effects of gene-specific expressions.

Trajectory analysis

Cell differentiation trajectory was reconstructed using Monocle2 (Qiu et al., 2017) Highly-variable genes (HVGs) was used to sort cells based on their spatial-temporal differentiation. Next, we used DDRTree to perform FindVairableFeatures and dimension-reduction. Finally, the trajectory was visualized using the plot_cell_trajectory function. Next, CytoTRACE (Gulati et al., 2020) (a computational method that predicts the differentiation state of cells from single-cell RNA-sequencing data using gene Counts and Expression) was used to predict the differentiation potential of monocyte clusters.

Cell-cell interaction analysis (CellPhoneDB)

CellPhoneDB v2.1.0 (Efremova et al., 2020) analyzed cell-cell interactions based on receptor–ligand interactions between 2 cell types/subtypes. First, the average level of ligand-receptor expression in interacting clusters was calculated by permuting all cell labels 1,000 times randomly. Then, the individual ligand or receptor expression was thresholded based on the average log gene expression distribution across all cell types. Finally, the significant cell-cell interactions were defined as p-value <0.05 and average log expression >0.1, visualized with the circle v0.4.10 R package.

Transcription factor regulatory network analysis

The transcription factor network was constructed by pyscenic v0.11.0 (Van de Sande et al., 2020) using the scRNA expression matrix and transcription factors in AnimalTFDB. The GRNBoost2 model predicts regulatory networks by co-expressing regulators and targets. CisTarget was applied to exclude indirect targets and to search transcription factor binding motifs. Following that, AUCell was used to quantify regulon activity for each cell. Finally, Pheatmap in R was used to visualize top TF regulons with high RSS (Regulon Specificity Score).

Results

Single-cell transcriptome analysis of multicellular GGN and SN ecosystems

A microwell-based scRNA sequence analysis was conducted. We collected 15,902 single isolated cells (4,630 from GGN and 10,392 from SN) from six LUAD samples (Supplementary Figure S1A; Supplementary Figure S6; Supplementary Table S1). The cells were classified into nine major cell types using the Uniform Manifold Approximation and Projection (UMAP) clustering analysis (Figure 1A). SynEcoSys database identified nine cell
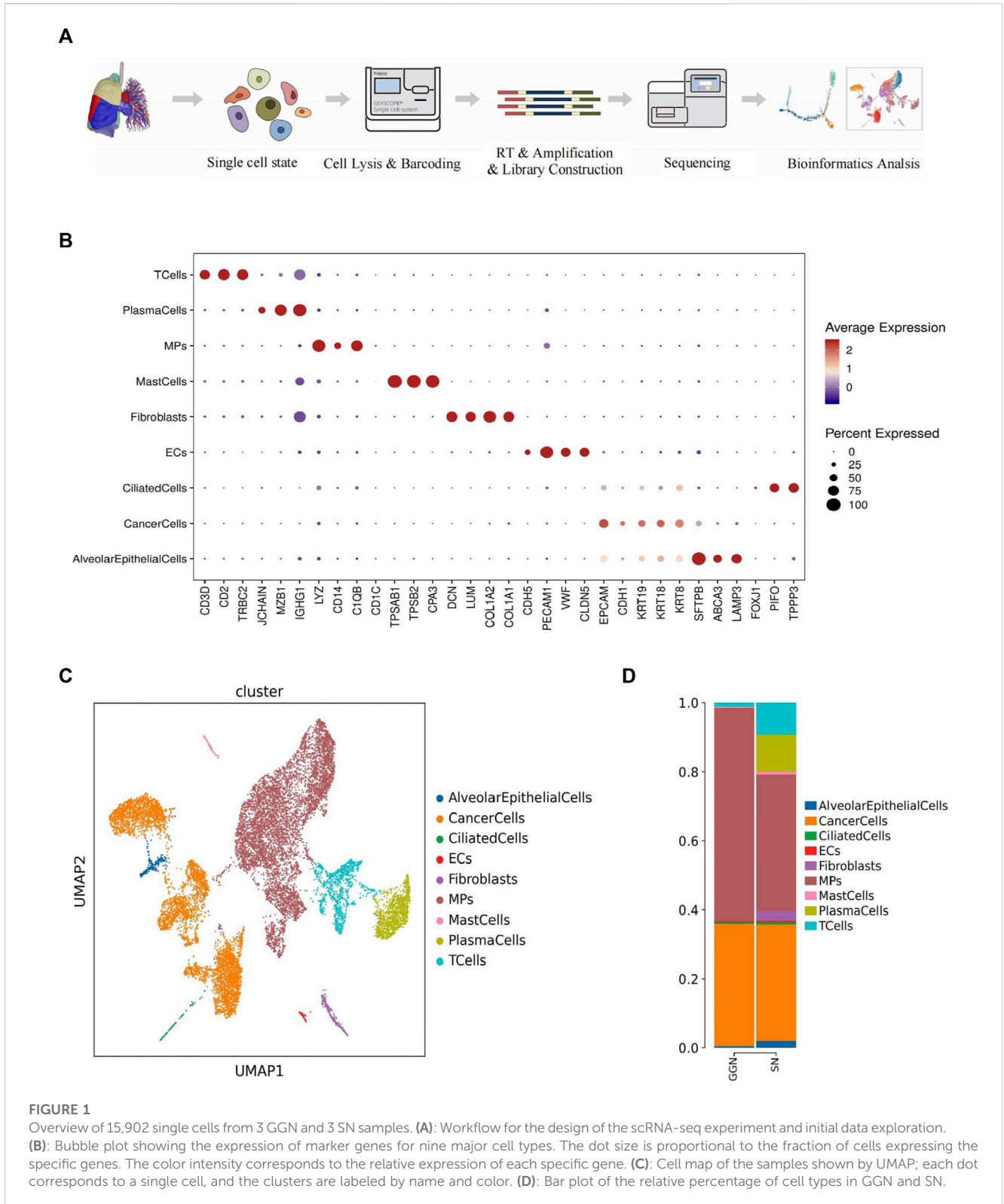
FIGURE 1
Overview of 15,902 single cells from 3 GGN and 3 SN samples. **(A)**: Workflow for the design of the scRNA-seq experiment and initial data exploration. **(B)**: Bubble plot showing the expression of marker genes for nine major cell types. The dot size is proportional to the fraction of cells expressing the specific genes. The color intensity corresponds to the relative expression of each specific gene. **(C)**: Cell map of the samples shown by UMAP; each dot corresponds to a single cell, and the clusters are labeled by name and color. **(D)**: Bar plot of the relative percentage of cell types in GGN and SN.

types annotated with cell type annotation markers, including alveolar epithelial cells (GGN, 22; SN, 227), cancer cells (GGN, 1,644; SN, 3,806), ciliated cells (GGN, 23; SN, 74), endothelial cells (GGN, 10; SN, 43), fibroblasts (GGN, 3; SN, 317), the mononuclear phagocytic system (MPs) (GGN, 2858; SN, 4,468), mast cells (GGN,

14; SN, 90), plasma cells (GGN, 4; SN, 1,205), and T cells (GGN, 52; SN, 1,042) (Figures 1B,C).

The GGN mainly comprised MPs (62%) and cancer cells (36%). However, SN comprised plasma cells (11%) and T cells (9%) in addition to MPs (40%) and cancer cells (34%). Malignant cells were present in
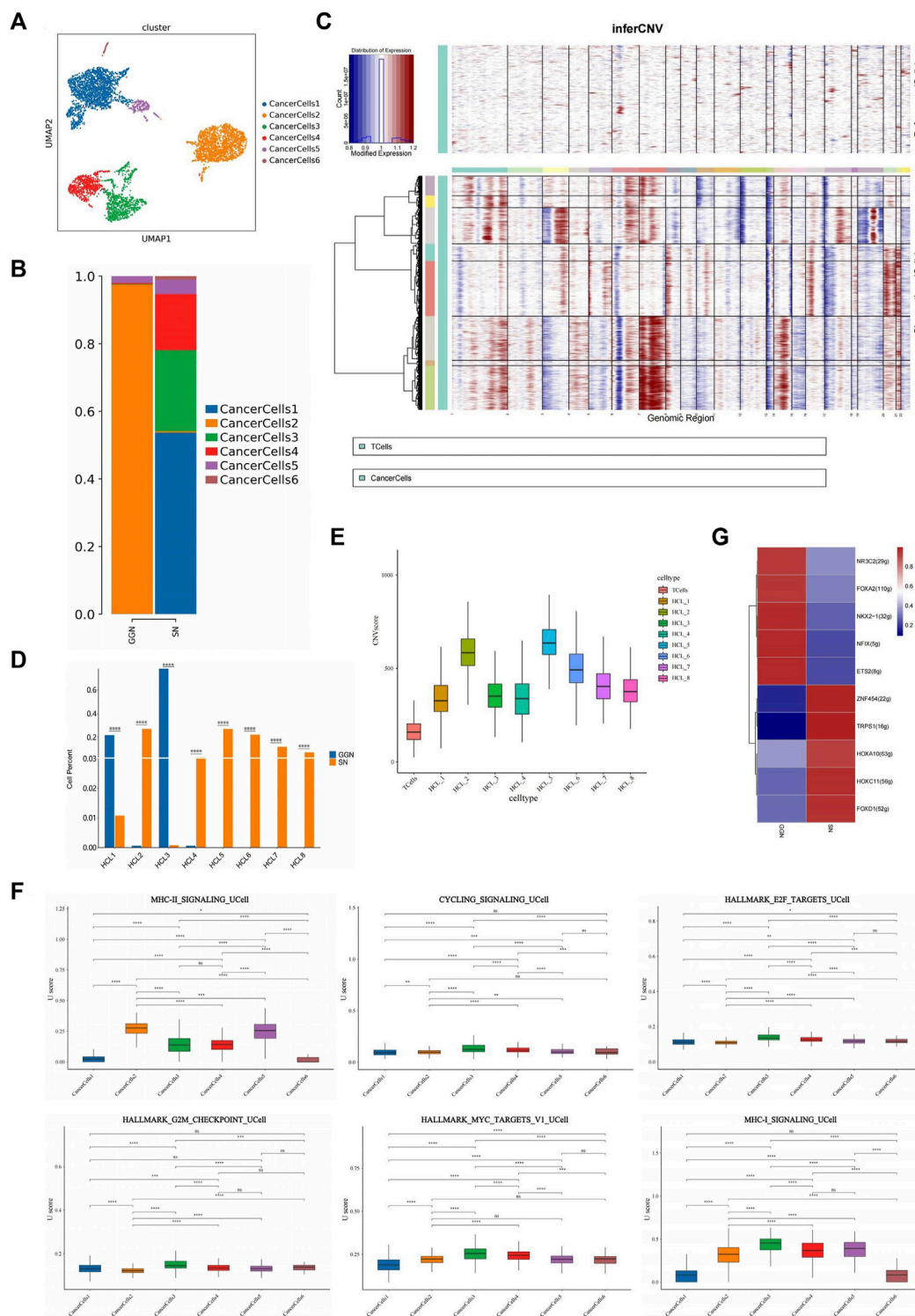
FIGURE 2
Identification and characterization of cancer cells in GGN and SN. **(A)**: UMAP plot showing cancer cells clustered into six subclusters **(B)**: Bar plot of the relative percentage of cancer cell subclusters in GGN and SN. **(C)**: Heat map showing CNVs Were inferred based on single-cell RNA sequencing data of individual cells from samples. Non-cancer cells were treated as references (top), and CNVs were observed in cancer cells (bottom). The color shows the log2 CNV ratio (HCL_1–8). Red: amplifications; blue: deletions. **(D)**: Bar plot of the relative percentage of cancer cell subclusters in GGN and SN. **(E)**: CNV value of cancer cell subclusters. **(F)**: Box plots of UCell score for different signaling pathways in cancer cell subclusters. Two-sided unpaired Wilcoxon rank-sum test was used for analysis; all differences with $p < 0.05$ are indicated, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns non-significance. **(G)**: SCENIC analysis of cancer cells (GGN vs SN). Shown are the top 5 upregulated/downregulated transcription factors, respectively.

both groups in a comparable proportion. The epithelial cells in GGN mainly consisted of cancer cells, whereas SN had a few alveolar epithelial cells (2%) and fibroblasts (3%) (Figure 1D). GGN and SN have complex cellular ecosystems, and their cell proportions differ. Herein, we focused on the functions of important cellular clusters.

Cancer cells from GGN and SN show different transcriptional characteristics

We detected cancer cells using markers typical of LUAD and alveolar epithelium (EPCAM, CDH1, KRT19, KRT18, KRT8) (Figure 1B). A total of 5,450 cancer cells were categorized into six subclusters (GGN, 1,644; SN, 3,806) (Figure 2A). The clustering revealed group specificity, with CancerCells (CaC) 2 predominantly present in GGN, whereas SN primarily comprised CaC1, CaC3, and CaC4. A single subcluster dominated the presence of cancer cells in GGN. However, the cancer cells in SN revealed a mixture of multiple subclusters with no obvious dominant cluster. Patients with SN had greater inter-tumour heterogeneity than those with GGN (Figure 2B).

Cancer cells exhibited a higher copy number variation (CNV) than immune cells. CNV is a general term used to describe a molecular phenomenon of genome sequence duplication, The analysis of CNVs is an important aspect of tumor molecular diagnosis (Pös et al., 2021). Extensive chromosome 13 deletion was observed in cancer cell samples (Figure 2C). Human chromosome 13 (HSA13), which comprises 1,381 genes, 41 novel genes, and 477 pseudogenes, has the lowest proportion of repetitive sequences and is connected with the initiation and progression of various human cancers and disorders (Bailey et al., 2002). The prognosis of patients with LUAD is adversely affected by copy number deletion of chromosome 13 (Han et al., 2019). The CNV values divided cancer cells into eight subclusters (HCL_1–8) (Supplementary Figure S1B). GGN mainly comprised HCL_1 and HCL_3, whereas SN primarily comprised HCL_2, HCL_4, HCL_5, HCL_6, HCL_7, and HCL_8 (Figure 2D). Chromosomes 16, 18, and 20 in the GGN cancer cell subclusters underwent significant amplification. In addition, chromosomes 1 and 7 of the SN cancer cell subclusters underwent significant amplification, whereas chromosomes 10 and 15 underwent deletion (Figure 2C). The CNV values of HCL_2, HCL_5, and HCL_6 were significantly higher than those of other subclusters (Figure 2E), indicating that SN cancer cells have a higher degree of CNV than GGN cancer cells. Studies have proved that CNV variation is related to tumor malignancy and prognosis of patient (Pös et al., 2021), suggesting that SN tumor cells are more malignant.

GGN cancer cells express significantly higher levels of surfactant-related proteins than SN cancer cells, such as surfactant protein (SFTP) A1, SFTPA2, and SFTPB. In addition, cancer cells in SN upregulated MT-RNR2, DST, and MALAT1 (Supplementary Figure S1C). Furthermore, cancer cells in GGN highly expressed alveolar type II (AT2) cell-related marker genes (SFTPA1, SFTPA2, SFTPB, ABCaC3), suggesting that cancer cells in GGN might originate from AT2 cells. However, the alveolar type I

(AT1) cell marker genes PDPN or AGER was not highly expressed in either group (Supplementary Figure S1D) (Jacob et al., 2017; Evans and Lee, 2020), suggesting that AT1-like cells had a low expression level in both samples.

Gene Ontology (GO) pathway enrichment analysis revealed that cancer cells in GGN upregulate ribosome-related pathways, establish protein localization to the endoplasmic reticulum, cotranslational protein targeting to membrane, Major Histocompatibility Complex (MHC) class II protein complex binding, and other pathways. In addition, the Kyoto Encyclopaedia of Genes and Genomes (KEGG) analysis revealed ribosome-related pathway enrichment ((Supplementary Figure S2A). Ribosomal protein (RP) family-related genes (RPS21,RPS2,RPS18,RPLP1,RPL5,RPL13 A) were significantly upregulated in GGN (Supplementary Figure S1E). Thus, ribosome-related pathways play a crucial role in GGN tumourigenesis. However, RNA splicing, cell cycle phase transition regulation, transcription coregulator activity, and other pathways were upregulated in cancer cells in SN. The KEGG enrichment analysis also revealed enrichment of the pathways associated with spliceosome and protein processing in the endoplasmic reticulum (Supplementary Figure S2B). RNA splicing-related pathways were enriched in SN. Aberrant RNA splicing factor expression in oncogenes and cancer suppressor genes can regulate post-transcriptional mechanisms to promote tumor growth (David et al., 2010; Hsu et al., 2015). According to the gene set variation analysis (GSVA) and UCell gene set score analysis, CaC2 scored higher in MHC-II_SIGNALLING; CaC1, CaC3, and CaC4 scored higher in cell cycle regulation pathways such as CYCLING_ SIGNALLING, HALLMARK_E2F_TARGETS, and HALLMARK_ G2M_CHECKPOINT compared with CaC2; CaC3 and CaC4 scored higher in the HALLMARK_MYC_TARGETS_ V1 and MHC-I_SIGNALLING (Figure 2F; Supplementary Figure S2C). Therefore, compared with SN, GGN was enriched in the MHC-II_SIGNALLING pathway. SN was enriched in cell cycle regulation pathways and the MHC-I_SIGNALLING pathway. The above results suggest different functional patterns of the two groups of cancer cells, resulting in differences in proliferation and aggressiveness.

Single-cell regulatory network inference and clustering analysis (SCENIC) showed significant upregulation of E26 transformation-specific transcription factor 2 (ETS2) in GGN and transcriptional repressor erythroid transcription factor binding 1 (TRPS1) in SN (Figure 2G). ETS2, an ETS transcription factor family oncogene, inhibits lung cancer cells' growth, migration, and invasion by suppressing mesenchymal-epithelial transition (MET) expression (Kabbout et al., 2013). TRPS1, a GATA family transcriptional regulatory factor, has been shown to induce tumor angiogenesis, affect VEGFA expression, and promote tumor cell proliferation in tumors (Hu et al., 2014). TRPS1 can also induce MET to increase malignant cell migration and invasion through ZEB2 regulation (Stinson et al., 2011).

An analysis of the trajectory of six cancer cells was performed to determine the developmental relationship between cancer cells in GGN and SN. The dominant subclusters of cancer cells in GGN were concentrated at the tail end of the trajectory. In contrast, cancer cells in SN were scattered in the anterior-middle part of the trajectory. Therefore, cancer cells in SN were more dispersed and differentiated earlier in the trajectory than in GGN, and the cancer cell stemness in the former might be stronger (Supplementary Figure S2D). In addition,
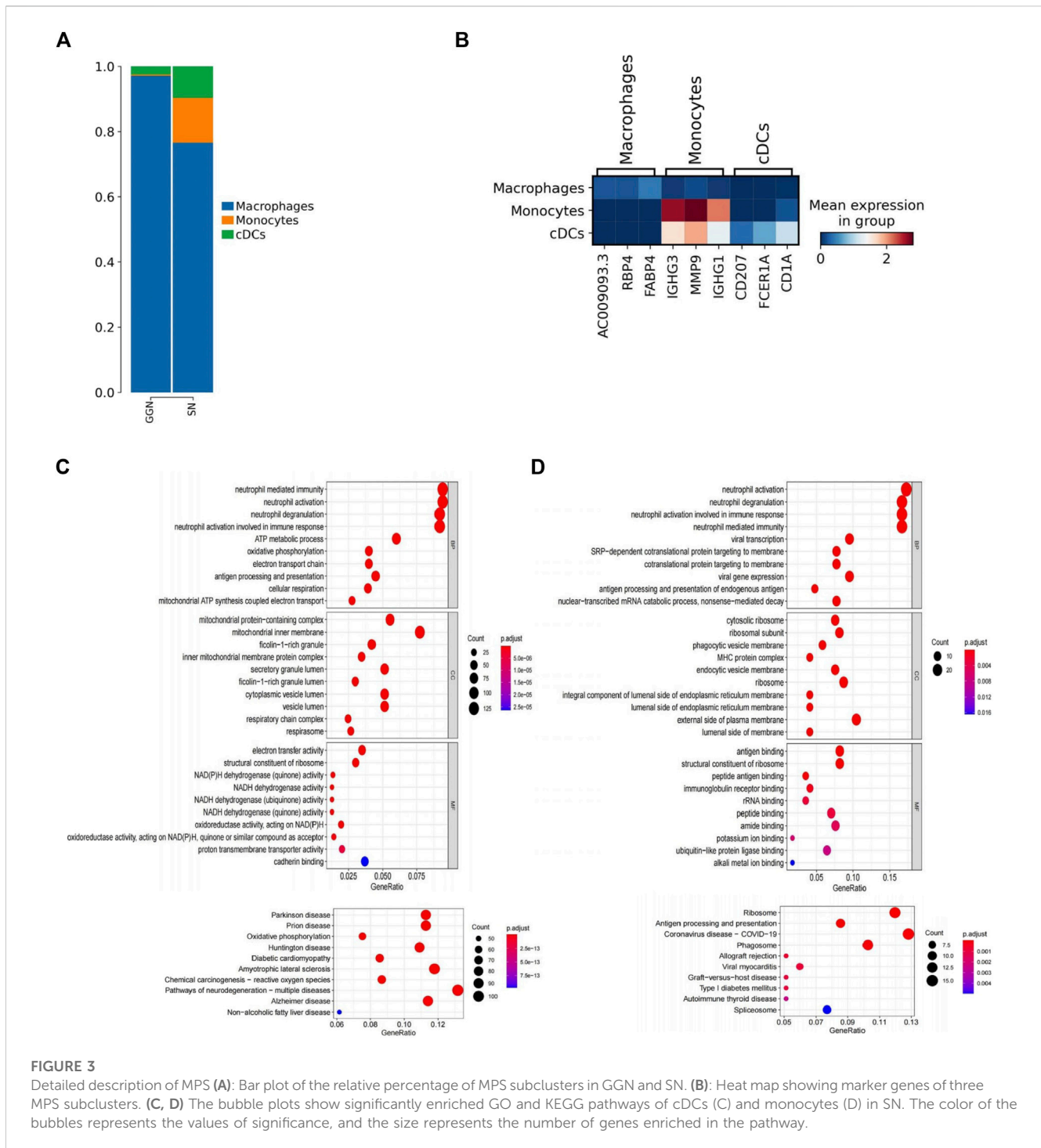
FIGURE 3
Detailed description of MPS **(A)**: Bar plot of the relative percentage of MPS subclusters in GGN and SN. **(B)**: Heat map showing marker genes of three MPS subclusters. **(C, D)** The bubble plots show significantly enriched GO and KEGG pathways of cDCs (C) and monocytes (D) in SN. The color of the bubbles represents the values of significance, and the size represents the number of genes enriched in the pathway.

expressions of C16orf89, CD74, CD55, and CCND1 were significantly upregulated at the end of the trajectory as the trajectory developed (Supplementary Figure S2E). CD74 is a type II transmembrane protein with various biological functions that promote the angiogenesis of lung cancer cells when co-expressed with macrophage migration inhibitory factor (MIF) (McClelland et al., 2009). In addition, CD74 can form oncogenic fusion genes with NRG1 to promote the progression of lung cancer cell (Murayama et al., 2016). CD55 and CCND1 are reportedly associated with cancer progression, as demonstrated in liver cancer (Meng et al., 2017; Nie et al., 2020).

Varying distribution of MPs subtypes between GGN and SN

MPs were identified by marker genes (LYZ, CD14, C1QB, CD1C) (Figure 1B). A total of 6,726 MPs (2,858 in GGN and 4,468 in SN) were divided into three subclusters: macrophages, monocytes, and conventional dendritic cells (cDCs). Among them, macrophages accounted for most of the subclusters. In GGN, macrophages were almost dominant, and a small number of cDCs(3%) were also included. However, in addition to most
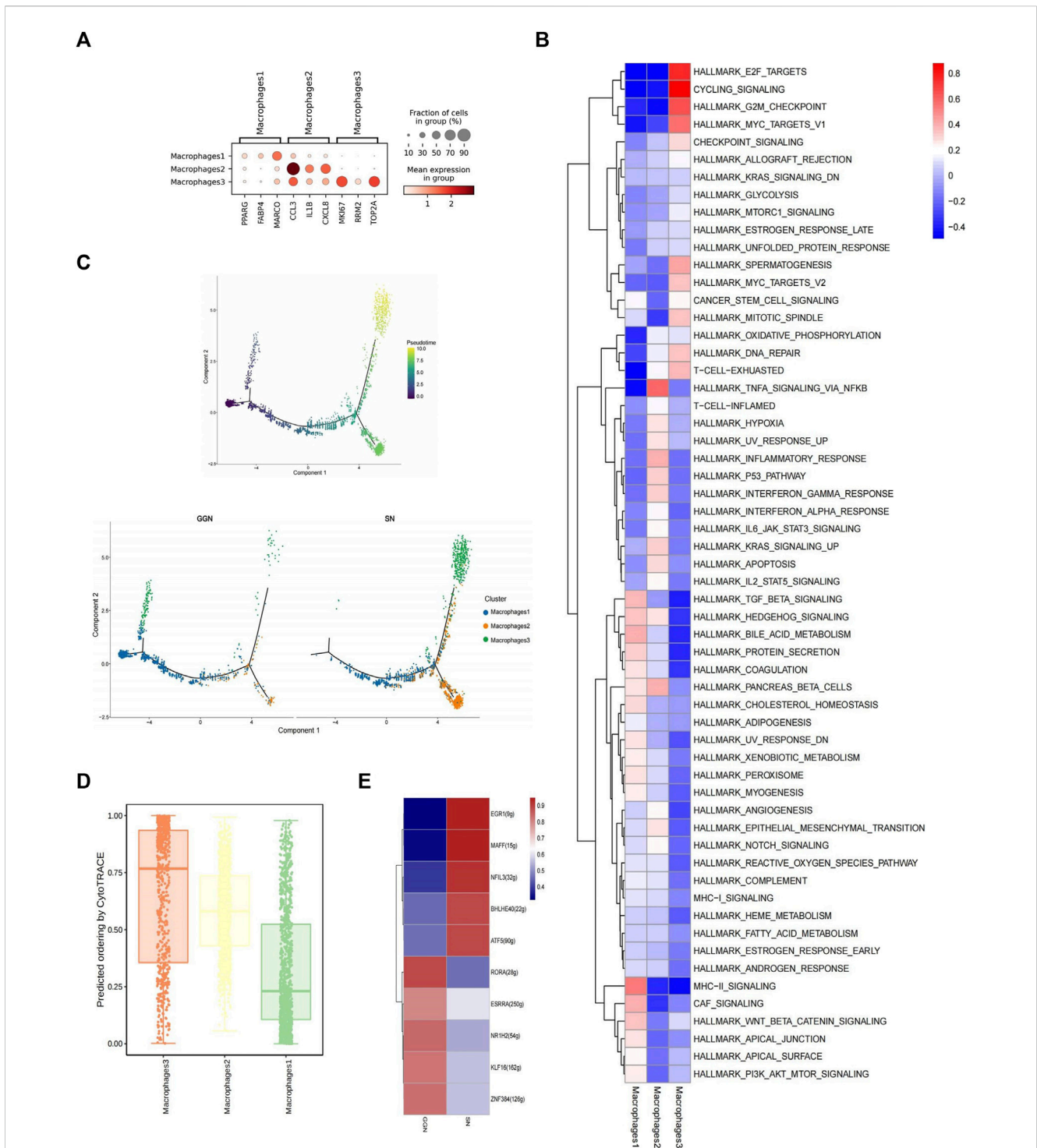
FIGURE 4
Different composition and characteristics of macrophages between GGN and SN. **(A)**: Bubble plot showing marker genes of macrophage subcluster. The color of the bubbles represents the average expression level of the gene, and the size represents a fraction of cells. **(B)**: Heat map showing the difference in metabolic pathways scored by GSVA of macrophage subclusters. **(C)**: Differentiation trajectory of cancer cells was predicted by monocle 2. **(D)**: Box plot showing cytotrace scores of macrophage subclusters, ranked from top to bottom by the median value. **(E)**: SCENIC analysis of macrophages (GGN vs SN). Shown are the top 5 upregulated/downregulated transcription factors, respectively.

macrophages, SN comprised a certain percentage of monocytes (14%) and cDCs (10%) (Figure 3A). CDCs (CD207+, FCER1A+, CD1A+) (Figure 3B) were more enriched in SN. The GO and KEGG enrichment pathway analyses revealed that the upregulated

enrichment pathways of cDCs in SN included neutrophil-related immune response and oxidative phosphorylation pathways (Figure 3C). In SN, monocytes (IGHG3+, MMP9+, IGHG1) (Figure 3B) accounted for approximately 14% of the MPs. In

contrast with GGN, monocytes in SN are recruited extensively into the tumor tissue. They are differentiated into macrophages and dendritic cells, altering the TME and promoting immune evasion, angiogenesis, and metastatic growth (Olingy et al., 2019). GO and KEGG pathway analysis findings indicate that monocytes in the SN are up-regulating neutrophil-related immune response, ribosome, and antigen processing and presentation pathways (Figure 3D). SN comprises abundant monocytes and cDCs and demonstrates, Moreover, both cells were enriched in neutrophil-related immune response pathways. SN showed stronger immune activity.

Macrophages in SN exhibit stronger proliferative capacity and inflammatory response

Macrophages, one of the important components of the innate immune system, are crucial for normal homeostasis and disease development and are closely associated with tumor development (Murray and Wynn, 2011). MPs comprise abundant macrophages (Figure 3A). We identified three macrophage subclusters, namely, alveolar-resident macrophage 1 (Mac1) (PPARG+, FABP4+, MARCO+), inflammatory cytokine-enriched macrophage 2 (Mac2) (CCL3+, IL1b+, CXCL8+), and proliferative macrophage 3 (Mac3) (MKI67+, RRM2+, TOP2A+) (Ma et al., 2022) (Figure 4A). GGN mainly comprised Mac1 (89%), as well as smaller quantities of Mac2 (5%) and Mac3 (6%). However, SN primarily comprised Mac1 (41%) and Mac2 (46%) and a smaller quantity of Mac3 (13%) (Supplementary Figure S3A).

According to the combined analysis of GSVA analysis and UCell gene set score, Mac1 scored higher in the HALLMARK_BILE_ACID_METABOLISM pathway; Mac3 scored higher in the cell cycle regulatory pathways such as HALLMARK_G2M_CHECKPOINT, CYCLING_SIGNALLING, HALLMARK_E2F_TARGETS, and HALLMARK_MYC_TARGETS_V1 (Figure 4B; Supplementary Figure S3A); Mac2 scored high in the HALLMARK_TNFA_SIGNALLING_VIA_NFKB (Figure 4B; Supplementary Figure S3B). In addition, Mac2 highly expressed inflammatory factors such as IL1B, CCL3, CCL4, CXCL2, and CXCL8 (Supplementary Figure S3C), suggesting a more severe inflammatory response in SN, which may act through the TNFA_SIGNALLING_VIA_NFKB pathways. However, Mac3 highly expressed proliferation-related genes, such as MKI67, RRM2, and TOP2A (Supplementary Figure S3C). It might promote tumor-associated macrophage proliferation and tumor growth by regulating the cell cycle through the above pathways (Murray and Wynn, 2011).

According to the GO and KEGG enrichment analyses in both groups, pathways such as neutrophil-related immune response, ribosome, RNA splicing, RNA catabolic process, and mRNA catabolic process were upregulated in SN (Supplementary Figure S4A), pathways such as upregulation of neutrophil-related immune response, focal adhesion, cell-substrate junction were upregulated in GGN. No significant KEGG pathway enrichment was observed in GGN (Supplementary Figure S4B). Increased pro-inflammatory cytokines and chemokines, such as IL1B, CCL3, CCL4, CXCL2, and CXCL8, were secreted by macrophages in SN (Supplementary

Figure S4C). These cytokines can promote the occurrence and development of inflammatory response, so it could be inferred that TME in SN has a more severe inflammatory response.

The Monocle2 package performed unsupervised trajectory analysis of macrophage transcriptional changes. It showed that Mac1 might differentiate into Mac2 and Mac3. Since Mac1 were mainly expressed in GGN while Mac2,3 were mainly expressed in SN, More Mac1 might be induced conversing to Mac2 and Mac3 in SN (Figure 4C). The greatest differentiation potential of proliferating Mac3 was suggested by CytoTrace analysis (Figure 4D). In addition, the tumor-associated gene Activating Transcription Factor 3 (ATF3) was upregulated, and the oncogene Aldehyde Dehydrogenase 2 (ALDH2) was downregulated as the trajectory progressed (Supplementary Figure S4D).

SCENIC analysis revealed significant Early growth response factor 1 (EGR1) upregulation in SN and Nuclear receptor subfamily 1 Group H Member 2(NR1H2) in GGN (Figure 4E). EGR1 regulates cell proliferation, apoptosis, immune cell activation, and stromal degradation and is closely associated with tumorigenesis (Li et al., 2019). NR1H2 suppresses inflammatory genes in macrophages (A-González NCastrillo, 2011), activates liver X receptors, and inhibits cancer cell growth, including colorectal cancer (Liang et al., 2019).

In summary, macrophages in SN exhibited greater proliferative capacity and inflammatory response than those in GGN.

Significant immune cell enrichment in SN

1,094 T cells were identified based on labeled marker gene expression (CD3D, CD2, TRBC2) (Figure 1B). SN exhibited increased T cell enrichment, with decreased enrichment exhibited by GGN (Figure 1D). T cells were subdivided into four clusters, which were as follows: CD8 +T effector cell (CD8Teff), native T cell (NaiveT), proliferating T cell, and regulatory T cell (Treg). According to the GSVA analysis and UCell gene set score, proliferating T cells scored higher in HALLMARK_MITOTIC_SPINDLE, MHC-II_SIGNALING, and cell cycle regulation pathways, such as CYCLING_SIGNALLING, HALLMARK_E2F_TARGETS, HALLMARK_G2M_CHECKPOINT, HALLMARK_MYC_TARGET_v1 (Figures 5A,B). However, CD8Teff scored higher in the T_CELL_INFLAMED pathway (Figures 5A,B). According to the pseudotime analysis, proliferating T cells appeared first in the early stage of the trajectory, and the remaining three clusters were distributed in the middle and late stages (Figure 5C). In addition, CytoTrace analysis revealed the greatest differentiation potential of proliferating T cells (Figure 5D), suggesting that proliferating T cells promote the development of inflammation and the formation of TME by regulating the cell cycle.

Based on differential gene expression, the 1,209 plasma cells identified by the marker genes (JCHAIN, MZB1, IGHG1) (Figure 1B) were divided into four clusters: PlasmaCells1 (PC1) (SCGB3A1+, ENAM+, JCHAIN+), PlasmaCells2 (PC2) (CD3G+, CD52+, CST+), PlasmaCells3 (PC3) (Z93241.1+, NR4A1+, HSP90AA1+), and PlasmaCells4 (PC4) (EGR1+, IER2+, FOS+) (Supplementary Figure S5A). Plasma cells were abundantly
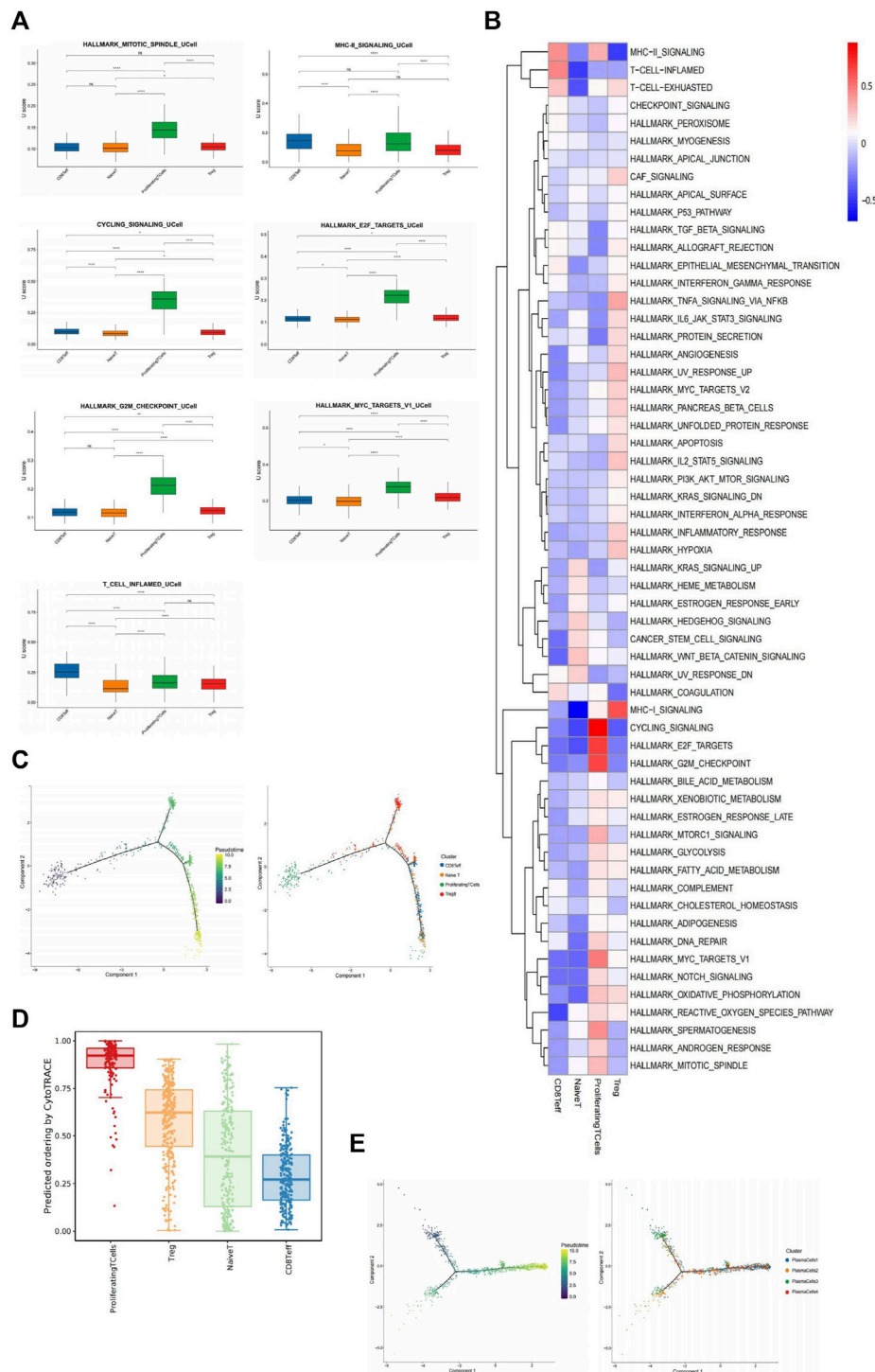
FIGURE 5

Immune cells play a greater role in SN. **(A)**: Box plots of UCell score for different signaling pathways in T cell subclusters. Two-sided unpaired Wilcoxon rank-sum test was used for analysis; all differences with $p < 0.05$ are indicated, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns non-significance. **(B)**: Heat map showing the difference in metabolic pathways scored by GSVA of T cell subclusters. **(C)**: Differentiation trajectory of T cells was predicted by monocle 2. **(D)**: Box plot showing cytotrace scores of macrophage subclusters, ranked from top to bottom by the median value. **(E)**: Differentiation trajectory of four plasma cell subclusters was predicted by monocle 2.

enriched mainly in SN but were minimally expressed in GGN (Figure 1D). According to the combined analysis of GSVA and UCell gene set scores, PC2 scored highest in the T_CELL_INFLAMED and HALLMARK_MITOTIC_SPINDLE pathways (Supplementary Figures S5B,C). According to the pseudotime analysis, PC2 and PC4 clusters were distributed across all time segments, PC1 was primarily distributed in the posterior segment, and PC3 was primarily distributed in the anterior segment
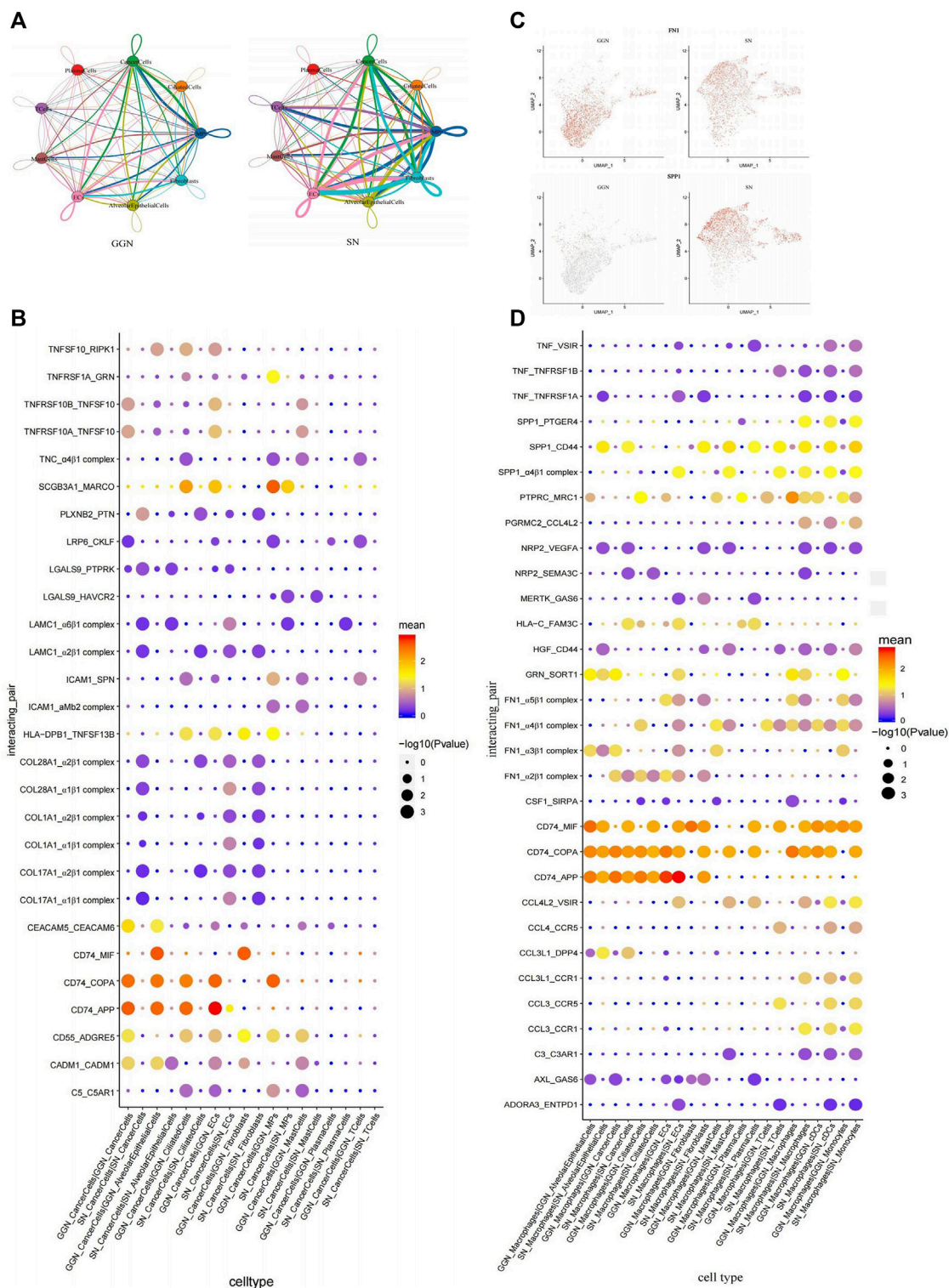
FIGURE 6
Intercellular interactions between GGN and SN. **(A)**: Interaction network diagram of cell types between GGN (left) and SN (right), the network node represents the cell type, the thickness of the network edge represents the total number of ligand and receptor pairs, and the line color represents the ligand cell type. **(B)**: Bubble plot showing ligand-receptor pairs of cancer cells in GGN and SN. Bubbles are labeled by −log10(P) (size) and average expression level of ligand−receptor pairs (color). *X*-axis: cell types of the two groups, *Y*-axis: ligand-receptor pairs. **(C)**: The expression of FN1 and SPP1 of macrophages in GGN and SN. **(D)**: Bubble plot showing ligand-receptor pairs of macrophages in GGN and SN. Bubbles are labeled by −log10(P) (size) and average expression level of ligand−receptor pairs (color). *X*-axis: cell types of the two groups, *Y*-axis: ligand-receptor pairs.

(Figure 5E), suggesting that the other plasma cell subclusters may be formed by PC1 differentiation, and PC3 may be in the mature terminal differentiation stage.

In summary, SN significantly enriched immune cells and formed a complex immune microenvironment.

Cancer cell-associated interactions suggest different patterns of tumor development between GGN and SN

The significant differences in the gene expression patterns of cancer cells between GGN and SN might result in different interactions between cancer cells and TME. Therefore, CellphoneDB was used to study cancer cell-specific interaction pairs in different fractions. According to the cellular interaction pair network diagram, the number of intercellular pairs was higher in SN than in GGN, indicating that cancer cells in SN interact more with other cells (Figure 6A).

The adhesion-related molecules ADGRE5 and ICaCM1 were significantly upregulated in GGN (Supplementary Figure S5D). They interacted with the corresponding receptor CD55 and SPN (Figure 6B) and promoted tumor invasion and metastasis through the adhesion of cancer cells to normal cells (Ksiazek et al., 2010; Meng et al., 2017). Carcinoembryonic antigen cell adhesion molecule 5 (CEACAM5)-CEACAM6 interaction was significant in GGN (Figure 6B). In addition, CEACAM6 might also be involved in key cellular events such as tumor migration and invasion in GGN through adhesion (Blumenthal et al., 2005; Cameron et al., 2012). TNFRSF10 and its receptor have obvious interaction in GGN (Figure 6B), suggesting inflammation and apoptosis between cancer cells, cancer cells, and endothelial cells, and between plasma cells in GGN (Allen and El-Deiry, 2012; Cucolo et al., 2022). CD74 was significantly upregulated in cancer cells in GGN (Supplementary Figure S5D) and was strongly associated with APP and COPA in alveolar epithelial cells, ciliated cells, and endothelial cells (Figure 6B). CD74_APP and CD74_COPA interactions were less commonly reported. However, their high enrichment suggests their crucial role in GGN development and warrants further research.

SN is enriched in integrin-associated complex interactions, which form complexes with collagen family genes (COL28A1, COL1A1, COL17A1) and human laminin γ1 (LAMC1) (Figure 6B). Integrins, the typical adhesion molecules in cancer cells, bind to collagen to promote the occurrence and development of tumors. Different collagen types can bind to various integrins in multiple signaling pathways in cancer cells (Xu et al., 2019). For example, LKB1 is involved in the HIF/LOX pathway through collagen type IV and β1 integrins resulting in enhanced lung cancer cell proliferation and invasiveness (Gao et al., 2010). Integrin α2β1 mediates the adhesion of lung cancer cells to type IV collagen, enhances cancer cell proliferation, and promotes the metastasis of lung cancer cells to the liver (Burnier et al., 2011). In addition, synergistic effects of integrins with LAMC1 promote cell migration and invasion, regulate cell adhesion and motility to promote tumor progression, and are associated with poor prognosis (Zacapala-Gómez et al., 2020). In a word, cell adhesion plays an important role in SN.

In summary, our data suggested that complex interactions existed between cancer cells and cells of various compositions. And there were significant differences between GGN and SN.

The interaction characteristics of macrophages between two groups reflect changes in the immune environment

Macrophage subclusters differed in their functional distribution and the cellular interactions between SN and GGN. Fibronectin 1 (FN1), which is highly expressed by GGN macrophages (Figure 6C), could form complexes with integrin subunits of other cells (FN1_α2β1 complex, FN1_α3β1 complex, FN1_α4β1 complex, FN1_α5β1 complex) (Figure 6D), which promoted cancer cell migration through adhesion (Huaman and Ogunwobi, 2020). Secreted phosphoprotein 1 (SPP1), highly expressed in SN, interacted significantly with CD44, PTGER4, and α4β1 (Figures 6C,D). SPP1, a multifunctional secretory acidic glycoprotein highly expressed in lung cancer, mediates macrophage polarisation and tumor immune evasion (Hu et al., 2005; Zhang et al., 2017). A tight association between Neuropilin 2 (NRP2) and VEGFA and SEMA3C was observed in SN. NRP2 regulates lymphatic vessel growth and promotes tumor metastasis to the lymphatic pathways (Raimondi and Ruhrberg, 2013). SN comprises extensive cell adhesion-related interactions like HGF_CD44 (Xu et al., 2021). In contrast with GGN, CCL4 and CCL3 interact significantly with their receptors in SN (CCL4L2_VSIR, CCL4_CCR5, CCL3L1_DPP4, CCL3L1_CCR1, CCL3_CCR5, CCL3_CCR1) (Figure 6D). The CC chemokines, CCL4 and CCL3, act alone or together by regulating the tumor immune microenvironment to promote tumor progression (Mukaida et al., 2020; Ntanasis-Stathopoulos et al., 2020). TNF-related interactions such as TNF_VSIR (Shapouri-Moghaddam et al., 2018) and CD74_MIF interactions are enriched in SN (Su et al., 2017), exhibiting more inflammation-related tumor features. Complex interactions exist between macrophages and cells of various compositions, significantly different between GGN and SN.

Discussion

GGN is an inert pathological subtype of LUAD with slow progression. GGN is less malignant than SN. Therefore, patients with GGN have a better prognosis than those with SN. Additionally, the clinical treatment strategies differ (Mao et al., 2019; Zhang et al., 2020). The present study comprehensively characterizes cellular profiles and transcriptomics in GGN and SN samples. It is revealed that the distinct cellular features and TMEs of the cell subtypes differ when the dynamic changes in their composition, expression of the molecular signatures, and intercellular interactions are compared. These findings help us understand lung cancer onset and progression mechanisms.

Cancer cells have the greatest heterogeneity compared with other cell types. Nevertheless, the proportion of cancer cells was similar in both groups. However, the higher CNV values of cancer cells in SN compared with cancer cells in GGN indicated a higher variability, which is consistent with the findings reported by Lu et al. (2020). This might be owing to the higher tumor mutational load in SN than in GGN (Chen

et al., 2021). In GGN, the cancer cells comprise one major subcluster. However, SN comprises multiple tumor subclusters. In addition, GGN cancer cells expressed AT2-associated genes, suggesting they may be derived from AT2 cells.

Compared with GGN, SN upregulated genes promote cell proliferation and migration like EGR1 and TRPS1, and downregulated oncogenes like NR1H2 and ETS2. In cancer cell analysis, ribosome-related pathways were significantly enriched in GGN. Ribosomes are important organelles for protein synthesis. In addition to performing protein translation, ribosomes co-regulate cancer cell growth and proliferation through the RAS/RAF/MEK/ERK, MYC, and PI3K/AKT/mTOR pathways (Woods et al., 2015). However, in macrophage analysis, ribosome-related pathways were more enriched in SN than in GGN, which might be associated with inflammation in lung cancer. Several studies have reported that tumors, including lung cancer, are strongly associated with inflammation, and upregulation of the ribosome biogenesis rate might be involved in tumor transformation in tissues affected by chronic inflammation, upregulation of rRNA transcription results in increased MDM2-mediated degradation of P53 proteasome, which reduces P53 expression and promotes tumorigenesis (Donati et al., 2011). Cell cycle regulation-related pathways in SN are significantly enriched in cancer cells, macrophages, and T cells. The G2M_ checkpoint pathway plays a crucial role in various human tumors (Evan and Vousden, 2001) and regulates cell growth and apoptosis in ovarian cancer (Zaffaroni et al., 2001) and colon cancer (Yang et al., 2018). E2F_targets (De Meyer et al., 2009) and MYC_targets_V1 (Stine et al., 2015) also involve tumor therapy, drug resistance, immune evasion, and progression. Tumor-associated macrophages are involved in the cell cycle regulation of cancer cells through M2 polarisation (Yoshikawa et al., 2012). The above pathways might be involved in SN development. Considering that immune cells, such as T and plasma cells, are mainly enriched in SN, gene expression of cancer cells and macrophages and pathway analysis suggested a more severe inflammatory response in SN than GGN. Therefore, it is speculated that the inflammatory response is a differentiating factor between GGN and SN.

The present study confirmed region-specific cellular interactions. Significant intercellular interactions between the two groups were screened for analysis. Integrins exerted adhesion interactions in both groups. However, certain differences were observed. In GGN, macrophages interacted with other cells via FN1 and integrin subunit pairing. However, in SN, cancer cells interact with other cells through a complex formed by collagen family genes, LAMC1, and integrin subunits. A clear association of CD44, PTGER4, and α4β1 in SN macrophages, plasma cells, cancer cells, and T cells was observed with upregulated SPP1 in macrophages. The activated interaction pairs were associated with macrophage polarisation and tumor immune evasion (Zhang et al., 2017). Inflammation-associated cellular interaction pairs were observed in both groups. However, the interactions were more pronounced in SN. Thus, SN and GGN exhibit similar but distinct specific interactions.

Our study has certain limitations. We could only partially describe all cell types, subtypes, and phenotypes in one report. Therefore, only certain key observations were reported herein, and further studies at a more comprehensive and deeper level need to be conducted.

Conclusion

In summary, the complex cellular and ecosystem processes in patients with lung cancer were characterized by the scRNA-seq technique, and GGN and SN exhibited similar but different specific TME. Macrophages might also contribute to the differences between SN and GGN in addition to cancer cells. In general, this study contributes to the understanding of LUAD and the discovery of new therapeutic targets.

Data availability statement

The data presented in the study are deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, accession number (GSA-Human: HRA005270).

Ethics statement

The studies involving humans were approved by Ethics Committee of Jiangyin Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

XH, GY and ZL conceived and designed the study. KY and ZZ collected the specimen and prepared single-cell suspension for sequencing. XJ and ZZ performed the experiments. XH and ZL performed the bioinformatic analyses. XH wrote the manuscript. GY edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2023.1198338/full#supplementary-material

References

A-González, N., and Castrillo, A. (2011). Liver X receptors as regulators of macrophage inflammatory and metabolic pathways. *Biochimica Biophysica Acta (BBA) - Mol. Basis Dis.* 1812 (8), 982–994. doi:10.1016/j.bbadis.2010.12.015

Allen, J. E., and El-Deiry, W. S. (2012). Regulation of the human TRAIL gene. *Cancer Biol. Ther.* 13 (12), 1143–1151. doi:10.4161/cbt.21354

Andreatta, M., and Carmona, S. J. (2021). UCell: robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* 19, 3796–3798. doi:10.1016/j.csbj.2021.06.043

Bailey, J. A., Gu, Z., Clark, R. A., Reinert, K., Samonte, R. V., Schwartz, S., et al. (2002). Recent segmental duplications in the human genome. *Sci. (New York, NY)* 297 (5583), 1003–1007. doi:10.1126/science.1072047

Blumenthal, R. D., Hansen, H. J., and Goldenberg, D. M. (2005). Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen). *Cancer Res.* 65 (19), 8809–8817. doi:10.1158/0008-5472.CAN-05-0420

Burnier, J. V., Wang, N., Michel, R. P., Hassanain, M., Li, S., Lu, Y., et al. (2011). Type IV collagen-initiated signals provide survival and growth cues required for liver metastasis. *Oncogene* 30 (35), 3766–3783. doi:10.1038/onc.2011.89

Cameron, S., de Long, L. M., Hazar-Rethinam, M., Topkas, E., Endo-Munoz, L., Cumming, A., et al. (2012). Focal overexpression of CEACAM6 contributes to enhanced tumourigenesis in head and neck cancer via suppression of apoptosis. *Mol. cancer* 11, 74. doi:10.1186/1476-4598-11-74

Chen, K., Bai, J., Reuben, A., Zhao, H., Kang, G., Zhang, C., et al. (2021). Multiomics analysis reveals distinct immunogenomic features of lung cancer with ground-glass opacity. *Am. J. Respir. Crit. care Med.* 204 (10), 1180–1192. doi:10.1164/rccm.202101-0119OC

Cucolo, L., Chen, Q., Qiu, J., Yu, Y., Klapholz, M., Budinich, K. A., et al. (2022). The interferon-stimulated gene RIPK1 regulates cancer cell intrinsic and extrinsic resistance to immune checkpoint blockade. *Immunity* 55 (4), 671–685.e10. doi:10.1016/j.immuni.2022.03.007

David, C. J., Chen, M., Assanah, M., Canoll, P., and Manley, J. L. (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* 463 (7279), 364–368. doi:10.1038/nature08697

De Meyer, T., Bijsmans, I. T., Van de Vijver, K. K., Bekaert, S., Oosting, J., Van Criekinge, W., et al. (2009). E2Fs mediate a fundamental cell-cycle deregulation in high-grade serous ovarian carcinomas. *J. pathology* 217 (1), 14–20. doi:10.1002/path.2452

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). Star: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* 29 (1), 15–21. doi:10.1093/bioinformatics/bts635

Donati, G., Bertoni, S., Brighenti, E., Vici, M., Trere, D., Volarevic, S., et al. (2011). The balance between rRNA and ribosomal protein synthesis up- and downregulates the tumour suppressor p53 in mammalian cells. *Oncogene* 30 (29), 3274–3288. doi:10.1038/onc.2011.48

Dura, B., Choi, J. Y., Zhang, K., Damsky, W., Thakral, D., Bosenberg, M., et al. (2019). scFTD-seq: freeze-thaw lysis based, portable approach toward highly distributed single-cell 3' mRNA profiling. *Nucleic acids Res.* 47 (3), e16. doi:10.1093/nar/gky1173

Efremova, M., Vento-Tormo, M., Teichmann, S. A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* 15 (4), 1484–1506. doi:10.1038/s41596-020-0292-x

Evan, G. I., and Vousden, K. H. (2001). Proliferation, cell cycle and apoptosis in cancer. *Nature* 411 (6835), 342–348. doi:10.1038/35077213

Evans, K. V., and Lee, J. H. (2020). Alveolar wars: the rise of *in vitro* models to understand human lung alveolar maintenance, regeneration, and disease. *Stem cells Transl. Med.* 9 (8), 867–881. doi:10.1002/sctm.19-0433

Gao, Y., Xiao, Q., Ma, H., Li, L., Liu, J., Feng, Y., et al. (2010). LKB1 inhibits lung cancer progression through lysyl oxidase and extracellular matrix remodeling. *Proc. Natl. Acad. Sci. U. S. A.* 107 (44), 18892–18897. doi:10.1073/pnas.1004952107

Gulati, G. S., Sikandar, S. S., Wesche, D. J., Manjunath, A., Bharadwaj, A., Berger, M. J., et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Sci. (New York, NY)* 367 (6476), 405–411. doi:10.1126/science.aax0249

Han, X., Tan, Q., Yang, S., Li, J., Xu, J., Hao, X., et al. (2019). Comprehensive profiling of gene copy number alterations predicts patient prognosis in resected stages I-iii lung adenocarcinoma. *Front. Oncol.* 9, 556. doi:10.3389/fonc.2019.00556

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Hsu, T. Y., Simon, L. M., Neill, N. J., Marcotte, R., Sayad, A., Bland, C. S., et al. (2015). The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* 525 (7569), 384–388. doi:10.1038/nature14985

Hu, J., Su, P., Jia, M., Wu, X., Zhang, H., Li, W., et al. (2014). TRPS1 expression promotes angiogenesis and affects VEGFA expression in breast cancer. *Exp. Biol. Med. (Maywood, NJ)* 239 (4), 423–429. doi:10.1177/1535370214523904

Hu, Z., Lin, D., Yuan, J., Xiao, T., Zhang, H., Sun, W., et al. (2005). Overexpression of osteopontin is associated with more aggressive phenotypes in human non-small cell lung cancer. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 11 (13), 4646–4652. doi:10.1158/1078-0432.CCR-04-2013

Huaman, J., and Ogunwobi, O. O. (2020). Circulating tumor cell migration requires fibronectin acting through integrin B1 or SLUG. *Cells* 9 (7), 1594. doi:10.3390/cells9071594

Jacob, A., Morley, M., Hawkins, F., McCauley, K. B., Jean, J. C., Heins, H., et al. (2017). Differentiation of human pluripotent stem cells into functional lung alveolar epithelial cells. *Cell. stem Cell.* 21 (4), 472–488. doi:10.1016/j.stem.2017.08.014

Kabbout, M., Garcia, M. M., Fujimoto, J., Liu, D. D., Woods, D., Chow, C. W., et al. (2013). ETS2 mediated tumor suppressive function and MET oncogene inhibition in human non-small cell lung cancer. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 19 (13), 3383–3395. doi:10.1158/1078-0432.CCR-13-0341

Ksiazek, K., Mikuła-Pietrasik, J., Catar, R., Dworacki, G., Winckiewicz, M., Frydrychowicz, M., et al. (2010). Oxidative stress-dependent increase in ICAM-1 expression promotes adhesion of colorectal and pancreatic cancers to the senescent peritoneal mesothelium. *Int. J. cancer* 127 (2), 293–303. doi:10.1002/ijc.25036

Kumar, M., Bowers, R. R., and Delaney, J. R. (2020). Single-cell analysis of copy-number alterations in serous ovarian cancer reveals substantial heterogeneity in both low- and high-grade tumors. *Cell. cycleGeorget. Tex)* 19 (22), 3154–3166. doi:10.1080/15384101.2020.1836439

Li, L., Ameri, A. H., Wang, S., Jansson, K. H., Casey, O. M., Yang, Q., et al. (2019). EGR1 regulates angiogenic and osteoclastogenic factors in prostate cancer and promotes metastasis. *Oncogene* 38 (35), 6241–6255. doi:10.1038/s41388-019-0873-8

Liang, X., Cao, Y., Xiang, S., and Xiang, Z. (2019). LXRα-mediated downregulation of EGFR suppress colorectal cancer cell proliferation. *J. Cell. Biochem.* 120 (10), 17391–17404. doi:10.1002/jcb.29003

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* 30 (7), 923–930. doi:10.1093/bioinformatics/btt656

Lu, T., Yang, X., Shi, Y., Zhao, M., Bi, G., Liang, J., et al. (2020). Single-cell transcriptome atlas of lung adenocarcinoma featured with ground glass nodules. *Cell. Discov.* 6, 69. doi:10.1038/s41421-020-00200-x

Ma, R. Y., Black, A., and Qian, B. Z. (2022). Macrophage diversity in cancer revisited in the era of single-cell omics. *Trends Immunol.* 43 (7), 546–563. doi:10.1016/j.it.2022.04.008

Mao, R., She, Y., Zhu, E., Chen, D., Dai, C., Wu, C., et al. (2019). A proposal for restaging of invasive lung adenocarcinoma manifesting as pure ground glass opacity. *Ann. Thorac. Surg.* 107 (5), 1523–1531. doi:10.1016/j.athoracsur.2018.11.039

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17 (1), 10. doi:10.14806/ej.17.1.200

McClelland, M., Zhao, L., Carskadon, S., and Arenberg, D. (2009). Expression of CD74, the receptor for macrophage migration inhibitory factor, in non-small cell lung cancer. *Am. J. pathology* 174 (2), 638–646. doi:10.2353/ajpath.2009.080463

Meng, Z. W., Liu, M. C., Hong, H. J., Du, Q., and Chen, Y. L. (2017). Expression and prognostic value of soluble CD97 and its ligand CD55 in intrahepatic cholangiocarcinoma. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 39 (3), 1010428317694319. doi:10.1177/1010428317694319

Mukaida, N., Sasaki, S. I., and Baba, T. (2020). CCL4 signaling in the tumor microenvironment. *Adv. Exp. Med. Biol.* 1231, 23–32. doi:10.1007/978-3-030-36667-4_3

Murayama, T., Nakaoku, T., Enari, M., Nishimura, T., Tominaga, K., Nakata, A., et al. (2016). Oncogenic fusion gene CD74-NRG1 confers cancer stem cell-like properties in lung cancer through a IGF2 autocrine/paracrine circuit. *Cancer Res.* 76 (4), 974–983. doi:10.1158/0008-5472.CAN-15-2135

Murray, P. J., and Wynn, T. A. (2011). Protective and pathogenic functions of macrophage subsets. *Nat. Rev. Immunol.* 11 (11), 723–737. doi:10.1038/nri3073

Nie, M., Wang, Y., Yu, Z., Li, X., Deng, Y., Wang, Y., et al. (2020). AURKB promotes gastric cancer progression via activation of CCND1 expression. *Aging* 12 (2), 1304–1321. doi:10.18632/aging.102684

Ntanasis-Stathopoulos, I., Fotiou, D., and Terpos, E. (2020). CCL3 signaling in the tumor microenvironment. *Adv. Exp. Med. Biol.* 1231, 13–21. doi:10.1007/978-3-030-36667-4_2

Olingy, C. E., Dinh, H. Q., and Hedrick, C. C. (2019). Monocyte heterogeneity and functions in cancer. *J. Leukoc. Biol.* 106 (2), 309–322. doi:10.1002/JLB.4RI0818-311R

Papalexi, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* 18 (1), 35–45. doi:10.1038/nri.2017.76

Pös, O., Radvanszky, J., Buglyó, G., Pös, Z., Rusnakova, D., Nagy, B., et al. (2021). DNA copy number variation: main characteristics, evolutionary significance, and pathological aspects. *Biomed. J.* 44 (5), 548–559. doi:10.1016/j.bj.2021.02.003

Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y. A., and Trapnell, C. (2017). Single-cell mRNA quantification and differential analysis with Census. *Nat. methods* 14 (3), 309–315. doi:10.1038/nmeth.4150

Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19 (11), 1423–1437. doi:10.1038/nm.3394

Raimondi, C., and Ruhrberg, C. (2013). Neuropilin signalling in vessels, neurons and tumours. *Seminars Cell. & Dev. Biol.* 24 (3), 172–178. doi:10.1016/j.semcdb.2013.01.001

Shapouri-Moghaddam, A., Mohammadian, S., Vazini, H., Taghadosi, M., Esmaeili, S. A., Mardani, F., et al. (2018). Macrophage plasticity, polarization, and function in health and disease. *J. Cell. physiology* 233 (9), 6425–6440. doi:10.1002/jcp.26429

Stine, Z. E., Walton, Z. E., Altman, B. J., Hsieh, A. L., and Dang, C. V. (2015). MYC, metabolism, and cancer. *Cancer Discov.* 5 (10), 1024–1039. doi:10.1158/2159-8290.CD-15-0507

Stinson, S., Lackner, M. R., Adai, A. T., Yu, N., Kim, H. J., O'Brien, C., et al. (2011). TRPS1 targeting by miR-221/222 promotes the epithelial-to-mesenchymal transition in breast cancer. *Sci. Signal.* 4 (177), ra41. doi:10.1126/scisignal.2001538

Su, H., Na, N., Zhang, X., and Zhao, Y. (2017). The biological function and significance of CD74 in immune diseases. *Inflamm. Res. official J. Eur. Histamine Res. Soc. [et al]* 66 (3), 209–216. doi:10.1007/s00011-016-0995-1

Travis, W. D., Asamura, H., Bankier, A. A., Beasley, M. B., Detterbeck, F., Flieder, D. B., et al. (2016). The IASLC lung cancer staging project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. *Journal of Thoracic Oncology* 11 (8), 1204–1223. doi:10.1016/j.jtho.2016.03.025

Travis, W. D., Brambilla, E., Noguchi, M., Nicholson, A. G., Geisinger, K. R., Yatabe, Y., et al. (2011). International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol.* 6 (2), 244–285. doi:10.1097/JTO.0b013e318206a221

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nature protocols* 15 (7), 2247–2276. doi:10.1038/s41596-020-0336-2

Wolf, F. A., Angerer, P., and Theis, F. J. (2018). Scanpy: Large-scale single-cell gene expression data analysis. *Genome biology* 19 (1), 15. doi:10.1186/s13059-017-1382-0

Woods, S. J., Hannan, K. M., Pearson, R. B., and Hannan, R. D. (2015). The nucleolus as a fundamental regulator of the p53 response and a new target for cancer therapy. *Biochimica et biophysica acta* 1849 (7), 821–829. doi:10.1016/j.bbagrm.2014.10.007

Xu, J., Zhang, M. Y., Jiao, W., Hu, C. Q., Wu, D. B., Yu, J. H., et al. (2021). Identification of candidate genes related to synovial macrophages in rheumatoid arthritis by bioinformatics analysis. *International journal of general medicine* 14, 7687–7697. doi:10.2147/IJGM.S333512

Xu, S., Xu, H., Wang, W., Li, S., Li, H., Li, T., et al. (2019). The role of collagen in cancer: from bench to bedside. *Journal of translational medicine* 17 (1), 309. doi:10.1186/s12967-019-2058-1

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a journal of integrative biology* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yang, D., Zhang, X., Zhang, W., and Rengarajan, T. (2018). Vicenin-2 inhibits Wnt/β-catenin signaling and induces apoptosis in HT-29 human colon cancer cell line. *Drug design, development and therapy* 12, 1303–1310. doi:10.2147/DDDT.S149307

Yoshikawa, K., Mitsunaga, S., Kinoshita, T., Konishi, M., Takahashi, S., Gotohda, N., et al. (2012). Impact of tumor-associated macrophages on invasive ductal carcinoma of the pancreas head. *Cancer science* 103 (11), 2012–2020. doi:10.1111/j.1349-7006.2012.02411.x

Zacapala-Gómez, A. E., Ld, C. A-R., Mendoza-Catalán, M. A., Salmerón-Bárcenas, E. G., Zubillaga-Guerrero, M. I., Torres-Rojas, F. I., et al. (2020). Integrin subunit β1 and laminin γ1 chain expression: a potential prognostic biomarker in cervical cancer. *Biomarkers in medicine* 14 (15), 1461–1471. doi:10.2217/bmm-2020-0140

Zaffaroni, N., De Marco, C., Villa, R., Riboldi, S., Daidone, M. G., and Double, J. A. (2001). Cell growth inhibition, G2M cell cycle arrest and apoptosis induced by the imidazoacridinone C1311 in human tumour cell lines. *European journal of cancer* 37 (15), 1953–1962. doi:10.1016/s0959-8049(01)00227-1

Zhang, Y., Du, W., Chen, Z., and Xiang, C. (2017). Upregulation of PD-L1 by SPP1 mediates macrophage polarization and facilitates immune escape in lung adenocarcinoma. *Experimental Cell. research* 359 (2), 449–457. doi:10.1016/j.yexcr.2017.08.028

Zhang, Y., Fu, F., and Chen, H. (2020). Management of ground-glass opacities in the lung cancer spectrum. *The Annals of thoracic surgery* 110 (6), 1796–1804. doi:10.1016/j.athoracsur.2020.04.094