Microenvironment Shapes Cell State, Plasticity, and Heterogeneity of Small Cell Lung Cancer

Parth Desai[1,2], Nobuyuki Takahashi[1,3], Rajesh Kumar[1], Samantha Nichols[1], Justin Malin[1], Allison Hunt[4], Christopher Schultz[1], Yingying Cao[5], Desiree Tillo[6], Darryl Nousome[6], Lakshya Chauhan[7], Linda Sciuto[1], Kimberly Jordan[8], Vinodh Rajapakse[1], Delphine Lissa[9], Yang Zhang[1], Suresh Kumar[1], Lorinc Pongor[10], Abhay Singh[7], Brett Schroder[1], Ajit Kumar Sharma[1], Tiangen Chang[5], Rasa Vilimas[1], Danielle Pinkiert[1], Chante Graham[1], Donna Butcher[11], Andrew Warner[11], Robin Sebastian[1], Mimi Mahon[12], Karen Baker[12], Jennifer Cheng[12], Ann Berger[12], Ross Lake[13], Melissa Abel[1], Manan Krishnamurthy[1], George Chrisafis[1], Peter Fitzgerald[6], Micheal Nirula[1], Shubhank Goyal[1], Devon Atkinson[14], Nicholas W. Bateman[15], Tamara Abulez[15], Govind Nair[16], Andrea Apolo[17], Udayan Guha[18], Baktiar Karim[11], Rajaa El Meskini[14], Zoe Weaver Ohler[14], Mohit Kumar Jolly[7], Alejandro Schaffer[5], Eytan Ruppin[5], David Kleiner[19], Markku Miettinen[19], G Tom Brown[19], Stephen Hewitt[19], Thomas Conrads[4], Anish Thomas[1]

1. Developmental Therapeutics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
2. Fox Chase Cancer Center, Temple University Hospital and Lewis Katz School of Medicine, Philadelphia, Pennsylvania.
3. Department of Medical Oncology, National Cancer Center Hospital East, Kashiwa, Japan.
4. Women's Health Integrated Research Center, Inova Health System, Falls Church, Virginia, USA.
5. Cancer Data Science Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
6. CCR Collaborative Bioinformatics, Resource, Office of Science and Technology Resources, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
7. Center for Biosystems Science and Engineering, Indian Institute of Science, Bangalore, India.
8. Department of Immunology and Microbiology, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA.
9. Laboratory of Human Carcinogenesis, Center for Cancer Research National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
10. HCEMM Cancer Genomics and Epigenetics Research Group, Szeged 6728, Hungary
11. Molecular Histopathology Laboratory, Laboratory Animal Sciences Program, Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, Maryland, USA.
12. Pain and Palliative care services, National Institutes of Health Clinical Center, Bethesda, Maryland, USA.
13. Laboratory of Genitourinary cancer Pathogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
14. Center for Advanced Preclinical Research, Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, Maryland, USA.
15. The Henry M. Jackson Foundation for the Advancement of Military Medicine Inc., Bethesda, MD, 20817, USA.
16. National Institute of Neurological Disorders and Stroke, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
17. Genitourinary Malignancies Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
18. Thoracic and GI Malignancies Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
19. Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.

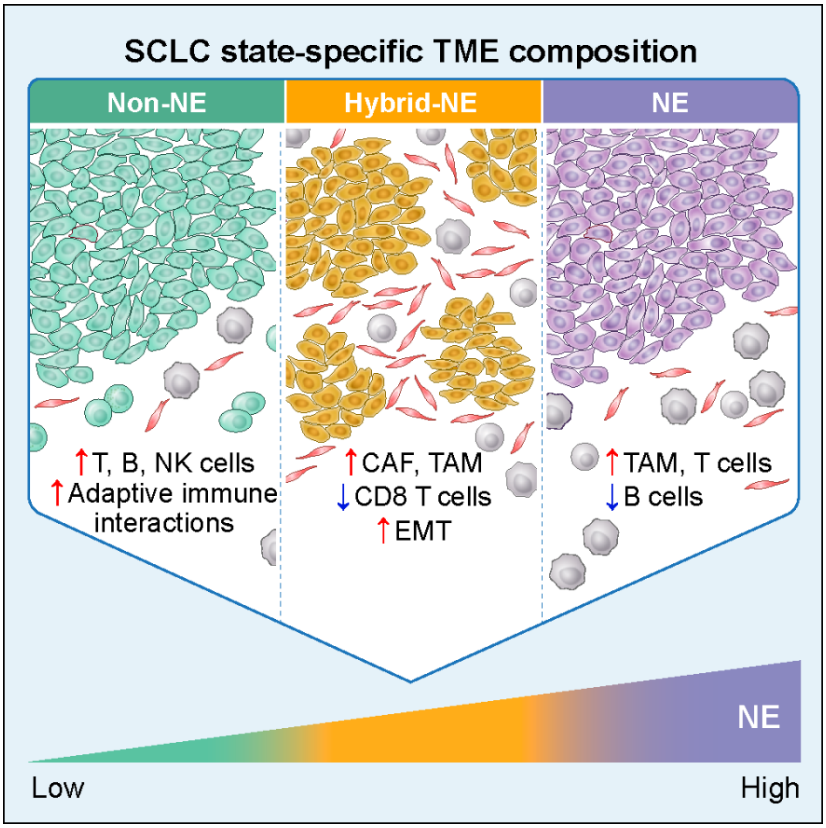
Corresponding author: Anish Thomas, MD, Center for Cancer Research, National Cancer Institute Building 10, Room 4-5330; Bethesda, MD 20892; +1 240 760 7343; anish.thomas@nih.gov

Key words: small cell lung cancer; tumor microenvironment; tumor heterogeneity; spatial transcriptomics; cancer associated fibroblasts; intercellular communication; rapid research autopsy.

Conflict of Interest disclosure: AT received grants paid to NCI from EMD Serono Research & Development, AstraZeneca, Gilead Sciences, and Prolynx during the conduct of the study. No other disclosures were reported.

Abstract

Small-cell lung cancer (SCLC) is the most fatal form of lung cancer. Intra-tumoral heterogeneity, marked by neuroendocrine (NE) and non-neuroendocrine (non-NE) cell states, defines SCLC, but the drivers of SCLC plasticity are poorly understood. To map the landscape of SCLC tumor microenvironment (TME), we applied spatially resolved transcriptomics and quantitative mass spectrometry-based proteomics to metastatic SCLC tumors obtained via rapid autopsy. The phenotype and overall composition of non-malignant cells in the tumor microenvironment (TME) exhibits substantial variability, closely mirroring the tumor phenotype, suggesting TME-driven reprogramming of NE cell states. We identify cancer-associated fibroblasts (CAF) as a crucial element of SCLC TME heterogeneity, contributing to immune exclusion, and predicting an exceptionally poor prognosis. Together, our work provides the first comprehensive map of SCLC tumor and TME ecosystems, emphasizing their pivotal role in SCLCs adaptable nature, opening possibilities for re-programming the intercellular communications that shape SCLC tumor states.



Introduction

Intratumor heterogeneity is a fundamental problem in cancer[1]. A major contributor to intratumor heterogeneity is phenotypic plasticity, which endows tumor cells with the ability to assume distinct cell identities, enabling metastatic capabilities and drug resistance[2,3]. A better understanding of the determinants of phenotypic plasticity is critical and may aid manipulation of cancer cell states and targeting the associated vulnerabilities.

Small-cell lung cancer (SCLC), a high-grade neuroendocrine cancer, represents a paradigm to study tumor heterogeneity and its consequences. As the most metastatic, treatment-resistant, and fatal form of lung cancer[4,5], SCLC exhibits a high degree of intratumoral heterogeneity, harboring cells of neuroendocrine (NE) and non-neuroendocrine (non-NE) states[6-12], further defined by differential expression of lineage-defining transcription factors *ASCL1, NEUROD1*, and *POU2F3*. A fourth subtype has been characterized by *YAP1* expression[9,13-16] or low expression of all three transcription factors accompanied by an inflamed gene expression program[17]. SCLC subtypes, defined by the dominant cell states in each tumor, exhibit distinct therapeutic vulnerabilities[10,16]. Immunogenic plasticity and Notch signaling of non-NE SCLC underlie their favorable responses to immune checkpoint blockade[16,18]. NE SCLC is characterized by replication stress, rendering them susceptible to DNA repair-targeted agents[10]. While heterogeneity and plasticity are important determinants of SCLC clinical responses, the origins and organization of SCLC heterogeneity are poorly understood[12,17,19,20].

Tumor cell state represents the combined influences of both cell-intrinsic (e.g., mutational background, epigenetic state) and cell-extrinsic (e.g., cell-to-cell interactions, microenvironment) factors[21]. In genetically engineered mouse models (GEMMs), the cell-of-origin profoundly influences the SCLC cell states[22,23]. Notch signaling, generally suppressed in NE SCLC, can induce a transition from NE to non-NE cell state[24,25]. MYC, frequently amplified on extrachromosomal DNA[26], can activate Notch signaling to promote the temporal evolution of SCLC sequentially from an ASCL1 to a NEUROD1 to a non-NE state[24,27]. However, model systems that have informed SCLC biology to date harbor minimal to no tumor microenvironment (TME)[23,28], and as such, very little is known of the cell-extrinsic drivers of SCLC tumor state. Importantly, SCLC TME has not been examined in depth in human tumors, especially in a spatial context, owing to the challenges of obtaining tumor samples. SCLC is often diagnosed using fine needle aspirates, and biopsies at relapse are not standard. Research biopsies are difficult to obtain due to rapid cancer progression and patient comorbidities, and when available, they may not portray the true extent of TME heterogeneity[29]. Further, the sequencing approaches that have been applied to human SCLC to date, including bulk and single-cell RNA-sequencing (scRNA-seq), do not inform the spatial interactions between tumor cells and TME.

Thus, while most patients with SCLC are diagnosed with and succumb to metastatic disease[30], the current understanding of SCLC heterogeneity, derived largely from model systems, is limited to cell-intrinsic influences[23]. Our understanding of cell-extrinsic influences on SCLC plasticity is rudimentary, but a key role of these non-genetic mechanisms is suggested by two important observations: (i) transcriptional subtypes of human SCLC are not strongly associated with specific mutational patterns[31] and (ii) divergence of NE gene expression programs between human tumors and patient-derived xenografts which lack human TME[9]. We hypothesized that dynamic interactions between TME and tumor cells shape SCLC tumor states, plasticity, and heterogeneity. Here, we studied metastatic and treatment-resistant SCLC in patients who underwent research autopsies and applied spatially resolved transcriptomics, integrating the data with whole-genome sequencing, bulk, and immunohistochemistry (IHC), multispectral imaging of multiplex immunofluorescence, and mass spectrometry (MS)-based

proteomics seeking to map the cell-extrinsic factors that shape SCLC cell states in their positional context.

Results

Patient and tumor characteristics

SCLC surgical resections are rarely performed since the tumors are almost always widely metastatic by the time of diagnosis[30]. Diagnostic core needle biopsies and cytology samples provide an inadequate representation of the TME heterogeneity[29]. Thus, to characterize human SCLC tumor and TME in its spatial context, we performed rapid research autopsies on ten patients (Clinicaltrials.gov identifier: NCT01851395) and profiled formalin-fixed, paraffin-embedded (FFPE) tumors using NanoString GeoMx Digital Spatial Profiler and GeoMx Whole Transcriptome Atlas (Methods; **Fig. 1A,** Fig.S1A). Patients were mostly male and smokers (8 of 10 each), with a median age of 64 (range: 47-75 years) (Table S1). All patients had received at least two systemic therapies including platinum and etoposide, and immunotherapy in eight of ten cases. The study included a single metastatic site from each patient, selected to represent frequent sites of SCLC metastases[32]; seven derived from liver, two lymph nodes, and one adrenal gland. Expert pathologists reviewed tumor sections to confirm the diagnoses. Nine out of 10 tumors were morphologically classified as SCLC. One tumor had combined morphology, with SCLC and squamous differentiation (patient #7). Patient #6 had a diagnostic biopsy of combined small cell carcinoma with adenocarcinoma differentiation, but only small cell morphology was observed on the autopsy tumor. All tumors underwent whole genome sequencing (WGS) and six of ten tumors were also laser micro-dissected and profiled using mass spectrometry-based proteomics. Mutations or copy number alterations of *TP53* and *RB1* were observed in most (*RB1* in 7 of 10 and *TP53* in 6 of 10) tumors (Fig. S1B). Genomic events that functionally mimic *TP53* and *RB1* inactivation[5], such as *MDM2* or *CCND1* amplifications were found in three tumors. *MYC* family genes were amplified (>4 copy numbers) in six cases. The two never-smoker tumors had no known pathogenic *TP53* alterations detected but had *MYCL* amplification and *RB1* deletion in one case and *CCND1* amplification in the other.

Tissue sections were profiled using fluorescently labeled antibodies targeting epithelial cells (pan-cytokeratin) and immune cells (CD45) to discriminate tumor and TME in selected regions, measuring an average 600-micron diameter (range, 400-650), with an average of three regions (range, 3-6) per tumor section. Tumor and TME were profiled independently based on pan-cytokeratin staining and morphological assessment, even when the two compartments were adjoined or interdigitated, yielding two unique whole transcriptome profiles per region of interest. Four tumor-only areas with minimal visible TME and four histologically normal areas with no visible surrounding tumor were also profiled. Thus, transcriptomes of 72 regions, including 36 tumor, 32 TME, and four normal segments were generated from ten tumors. Barcoded oligonucleotide probes were designed for a total of 18,677 genes representing the whole transcriptome[33]. Following probe hybridization, ultraviolet cleavage, and barcode collection, gene expression was quantified by polymerase chain reaction (PCR) amplification and Illumina sequencing. Two TME segments that did not meet the sequencing quality metrics (Fig. S1C) were excluded from further analyses. While TME and tumor segments had comparable areas of capture, more nuclei were profiled from tumor segments than TME segments (Fig. S1D). Sequencing saturation was high (>90%) for all segments. Sublevel sections of areas profiled for gene expression were additionally examined using IHC, multiplex protein immunofluorescence, whole genome sequencing, and mass spectrometry-based proteomics.

Figure 1

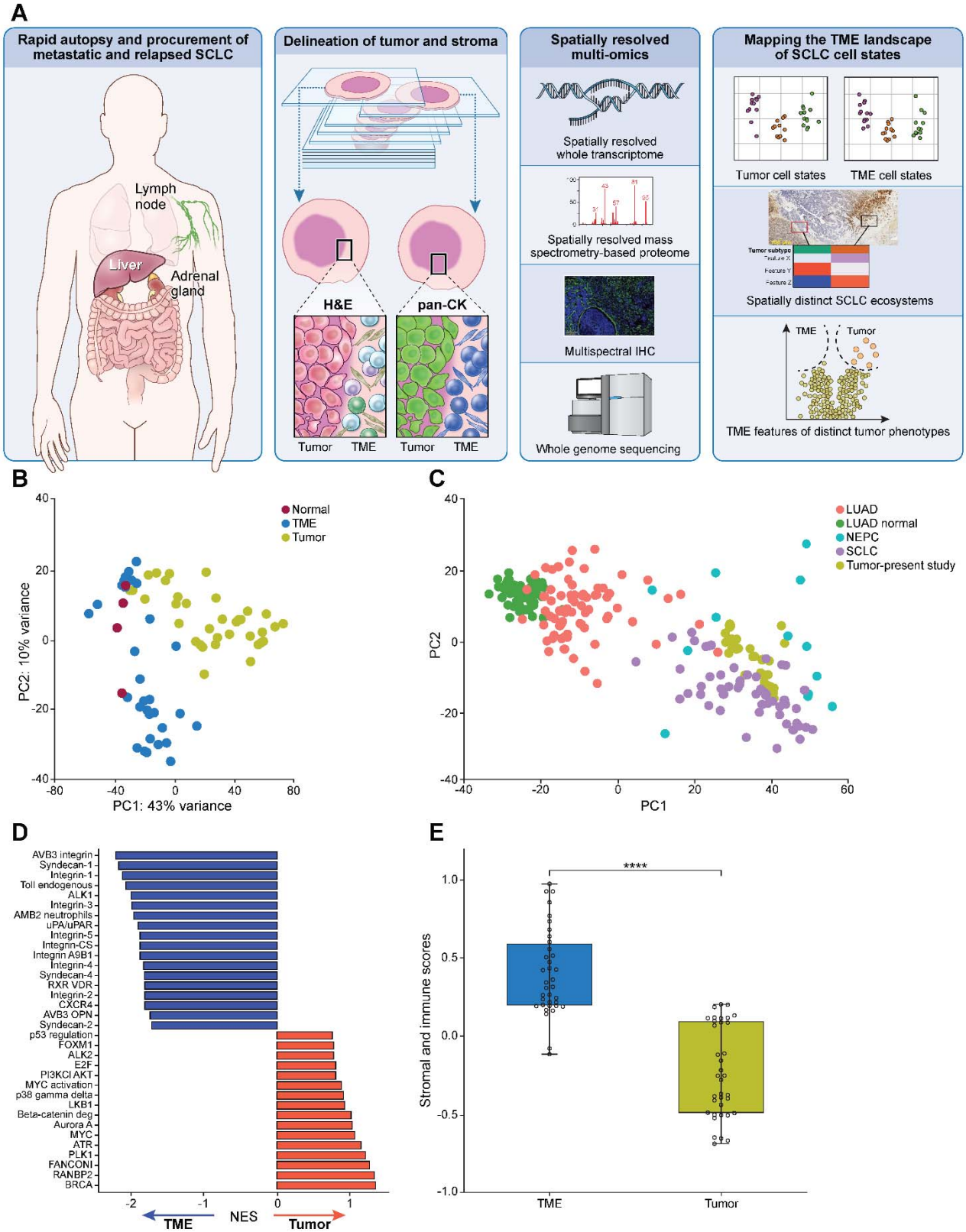

A
Rapid autopsy and procurement of metastatic and relapsed SCLC
Lymph node
Liver
Adrenal gland
Delineation of tumor and stroma
H&E
pan-CK
Tumor
TME
Tumor
TME
Spatially resolved multi-omics
Spatially resolved whole transcriptome
Spatially resolved mass spectrometry-based proteome
Multispectral IHC
Whole genome sequencing
Mapping the TME landscape of SCLC cell states
Tumor cell states
TME cell states
Spatially distinct SCLC ecosystems
TME
Tumor
TME features of distinct tumor phenotypes
B
Normal
TME
Tumor
PC2: 10% variance
PC1: 43% variance
C
LUAD
LUAD normal
NEPC
SCLC
Tumor-present study
PC2
PC1
D
AVB3 integrin
Syndecan-1
Integrin-1
Toll endogenous
ALK1
Integrin-3
AMB2 neutrophils
uPA/uPAR
Integrin-5
Integrin-CS
Integrin A9B1
Integrin-4
Syndecan-4
RXR VDR
Integrin-2
CXCR4
AVB3 OPN
Syndecan-2
p53 regulation
FOXM1
ALK2
E2F
PI3KCI AKT
MYC activation
p38 gamma delta
LKB1
Beta-catenin deg
Aurora A
MYC
ATR
PLK1
FANCONI
RANBP2
BRCA
TME
NES
Tumor
E
****
Stromal and immune scores
TME
Tumor

Figure 1: Dissection of metastatic and relapsed SCLC using spatial transcriptomics.

A. Workflow of SCLC tumor sampling, tissue sectioning, genomics, and spatially resolved transcriptomics and proteomics.

B. PCA of gene expression derived from tumor (n=36), TME (n=30), and normal (n=4) segments, 2500 genes with highest variance.

C. Projection of tumor segments (n=36) to PCA performed on lung adenocarcinoma, NEPC, SCLC, and adjacent normal lung gene expression[34].

D. NES (GSEA) of differentially expressed PID pathways[35] between tumor and TME segments.

E. Stromal and immune score[36] (ssGSEA) computed for TME and tumor segments[#]

Abbreviations: Pan CK, pan-cytokeratin; SCLC, small cell lung cancer; NEPC, neuroendocrine prostate cancer; LUAD, Lung adenocarcinoma; LUAD normal, adjacent normal lung; NES, normalized enrichment score; TME, tumor microenvironment; GSEA, gene set enrichment analysis; ssGSEA, single sample gene set enrichment analysis; PID, pathway interaction database; **** statistical significance at $p<0.0001$; [#] student t-test.

Distinct gene expression profiles of SCLC tumor and TME

To characterize the distinct transcriptomic features of tumor and TME, we evaluated the top 2500 variably expressed genes across all segments (**Fig. 1B**). This analysis revealed separate clustering of tumor and TME. Maximum variance was observed in the principal component (PC)1 axis which captured 43% of differences between tumor and TME. The histologically normal-appearing tissue adjacent to tumor clustered together with TME underscoring similarities between them. However, at a patient level, the normal tissue clustered separately from tumor and TME (Fig. S1E). Tumor gene expression profiles were similar to that of SCLC tumors[5] and neuroendocrine prostate cancer (NEPC)[34] (**Fig. 1C**) and distinct from lung adenocarcinoma and normal lung.

We applied several analytical approaches to better understand the specificity of tumor and TME gene expression profiles. Segments annotated as tumor exhibited higher NE gene expression scores[9,37] than TME and normal segments (Fig. S1F), consistent with the characteristic expression of NE genes in SCLCs. Gene Set Enrichment Analysis (GSEA) showed upregulation of pathways related to DNA repair, replication stress, and MYC in tumors, compared with TME which were enriched for pathways related to extracellular matrix and inflammation (**Fig. 1D**). Single sample Gene Set Enrichment Analysis (ssGSEA)[36] demonstrated significantly higher inferred fraction of stroma and immune cells in TME compared with tumors (**Fig. 1E**). SCLC tumors with lower NE scores[37] had higher stromal and immune scores (Fig. S1G), consistent with intrinsic tumor immunity of non-NE SCLC[16-18]. Estimated tumor purity of most tumor segments were more than 95% and exceeded those of bulk-tumor derived estimates[5,9,38] (Fig. S1H). Thus, despite its spatial proximity, SCLC tumor and TME harbor distinct gene expression states and programs.

Figure 2

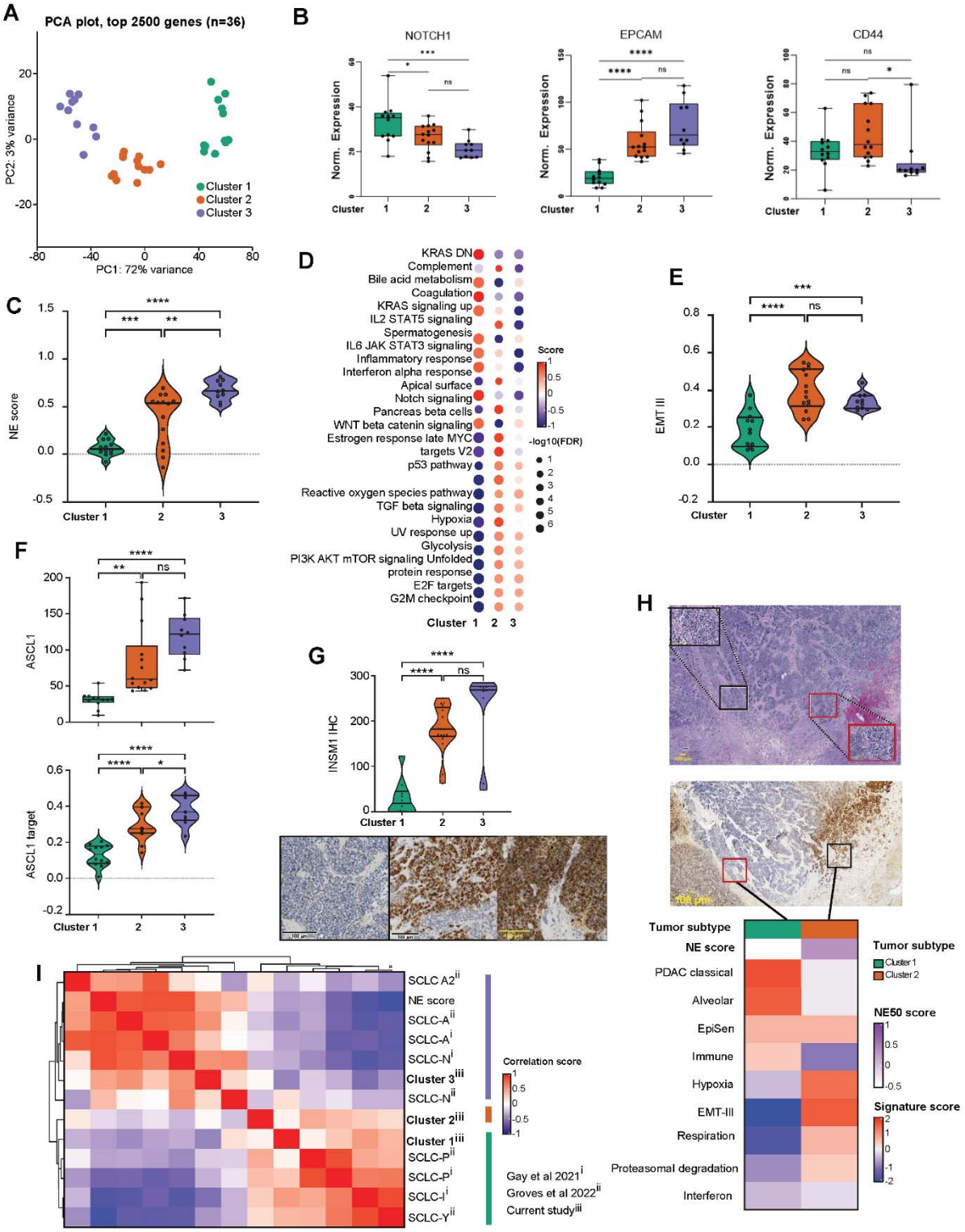


Figure 2: Spatial intra-tumoral heterogeneity of SCLC neuroendocrine differentiation.

A. PCA of 2500 genes with highest variance across tumor segments (n=36).

B. Normalized expression of NOTCH1, EPCAM and CD44 for the three tumor clusters[#].

C. NE scores across the tumor clusters[#].

D. Differentially enriched Hallmark pathways (ssGSEA) across the tumor clusters[$].

E. EMT III[39] (ssGSEA) scores across the tumor clusters[#].

F. *ASCL1* expression (top, Q3 value) and *ASCL1* target scores (bottom, ssGSEA) across the tumor clusters[#].

G. INSM1 protein expression across the tumor clusters. H-score range from 0-300[#]

H. Hybrid-NE (red square) and non-NE (black square) regions in morphologically similar and spatially proximate segments of the same tumor (patient #10). H/E, INSM1 IHC, and differentially upregulated cancer metaprograms[39,40] shown. Medium power (20x), high power (40x) inset. Scale bar at 100 µm.

I. Pairwise correlation of tumor-cluster signatures from current study and previously published SCLC gene signatures[7,17,41] computed on SCLC tumor transcriptomes[5].

Abbreviations: SCLC, small cell lung carcinoma; PCA, principal component analysis; ssGSEA, single sample gene set enrichment analysis; FDR, false discovery rate; EMT, Epithelial mesenchymal transformation; TF, transcription factors; ASCL1, Achaete-scute complex homologue 1; NEUROD1, neuronal differentiation 1; YAP1, yes-associated protein 1; POU2F3, POU class 2 homeobox 3; Q3, 3rd quantile value; UMAP, uniform manifold approximation and projection; NE, neuroendocrine; INSM1, insulinoma-associated protein 1; NES, normalized enrichment score; IHC, Immunohistochemistry; KRAS DN, KRAS downregulation; ns, statistically non-significant; H/E- hematoxylin & Eosin; *statistical significance at $p<0.05$; **statistical significance at $p<0.001$; ***statistical significance at $p<0.001$; ****statistical significance at $p<0.0001$; [#]Tukey's multiple comparison test; [$]FDR correction using Benjamini and Hochberg (BH) method.

Spatial Intra-tumoral heterogeneity of SCLC neuroendocrine differentiation

Seeking to classify SCLC tumors in an unbiased manner, we identified the optimal number of clusters for the 36 tumor segments as k=3 (see methods; Fig. S2A). The three clusters partitioned clearly at 2500 highly variant genes (**Fig. 2A**). In four of ten tumors profiled (Patients #3, 5, 7, 10), spatially proximate segments of the same tumor were separated into different clusters (Fig. S2B). Cluster 1 showed relatively high expression of *Notch* genes and *REST*, a repressor of neural gene expression and direct target of *Notch1* (**Fig. 2B**, Fig. S2C). Consistent with the negative regulation of NE differentiation by Notch, cluster 1 showed reduced expression of several key NE genes (*INSM1*, *BEX1*, *NCAM1*). In contrast, Cluster 3 showed upregulation of NE genes and Notch inhibitory ligands *DLL1* and *DLL3*. Cluster 2 exhibited features of both clusters 1 and 3, including simultaneous expression of both NE and non-NE genes[7]. Additionally, they expressed multiple EMT genes (*FN1, SNAI2*), in contrast to clusters 1 and 3 which showed variable expression of EMT genes. Cluster 2 also showed co-expression of mesenchymal and epithelial genes *CD44* and *EPCAM* that are characteristically expressed in non-NE and NE cells respectively[42]. *MYC* was amplified in bulk genome sequences of cluster 1-containing tumors, consistent with the known role of MYC in driving a non-NE phenotype[14,24,27]. However, the tumor clusters were not defined by genomic alterations including mutations and or copy number alterations of *TP53* or *RB1* (Fig. S2D).

In line with these observations at the gene expression level, cluster 1 exhibited significantly lower NE gene signature score[37] (mean=0.06; range, -0.07 to 0.21) than cluster 3 which showed the highest NE score (mean=0.67, range=0.51 to 0.81), and cluster 2 (mean=0.38; range, -0.13 to 0.69) which had intermediate NE scores (**Fig. 2C**). Consistently, cluster 3 was enriched for NE SCLC hallmarks including DNA repair and replication stress[43], E2F targets, and G2M checkpoints[10] (**Fig. 2D,** Fig. S2E). Cluster 1 was enriched for non-NE SCLC hallmarks such as inflammation and immunity[9,16,44]. Cluster 2 shared features of both clusters 1 and 3, including replication stress, E2F targets, G2M checkpoint, and immune pathways. Additionally, cluster 2 exhibited selective upregulation of hypoxia and epithelial-mesenchymal transition (EMT) III[39] (**Fig. 2E**). EMT-III is a cancer metaprogram of coordinately upregulated mesenchymal and epithelial markers consistent with a hybrid cellular state, previously described in neuroendocrine cancers[39].

Transcription factors function as molecular switches to regulate the expression of cell-type or lineage-specific target genes. Clusters 2 and 3 exhibited significantly higher expression of NE lineage-defining transcription factor *ASCL1* and its downstream target genes (**Fig. 2F**) than *NEUROD1* or non-NE lineage defining *POU2F3* (Fig. S2F). Cluster 1 exhibited low and comparable expression of all four transcription factors and their downstream target genes, reminiscent of SCLC-I (inflamed subtype) (Fig. S2G)[17]. IHC of sublevel sections (Table S2) confirmed these observations. The highest ASCL1 protein expression was observed in cluster 3, followed by cluster 2, and minimal expression in cluster 1. NEUROD1 and POU2F3 were rarely expressed across the three clusters (Fig. S2H). INSM1, a transcription factor that regulates global neuroendocrine gene programs[45] was also significantly highly expressed in clusters 3 and 2 compared with cluster 1 (**Fig. 2G**). YAP1 protein was expressed in cluster 2, but importantly these tumors did not express *YAP1* transcriptional target signature (Fig. S2I) or a pan-cancer *YAP1/HIPPO* signature[46] (Fig. S2J). In contrast, YAP1 protein was highly expressed in TME segments of cluster 2, and to a lesser extent cluster 3, but with concomitant upregulation of *YAP1* target genes in the TME. The functional role of stromal *YAP1* signaling in cluster 2 and 3 remains to be explored in future studies.

Morphologically all three SCLC clusters exhibited similar nuclear and cytoplasmic features. However, Cluster 2 tumors were more likely to localize at the invasive margin of tumors, forming nests and buds surrounded by desmoplastic tissue[47] (Fig. S2K). In contrast, Cluster 1 and Cluster 3 tumors formed large sheets or closely packed interconnected ribbons. GSEA of cancer hallmark capabilities[48,49] revealed striking enrichment of nearly every cancer hallmark in Cluster 2 tumors (Fig. S2L) compared with the other two subtypes indicating distinctly aggressive features of this phenotype. Even within the same tumor specimen, despite spatial proximity and morphological similarity between the unique segments, Cluster 2 tumor segments exhibited (**Fig.2H,** S2M, Table S3) specific upregulation of EMT-III[39] and hypoxia pathways**.** Given the technical constraints of our approach, which does not achieve single-cell resolution, we explored the potential for Cluster 2 tumors to represent potential contamination of tumor and stromal cells. However, the tumor and TME components of hybrid-NE segments distinctly separated along PC1 versus PC2 (Fig. S2N), and the hybrid-NE tumor segments demonstrated significantly lower stromal and immune scores compared to the corresponding hybrid-NE TME segments (Fig. S2O). Although, we cannot fully exclude the possibility that hybrid NE segments identified in our spatial profiling represent a mixture of NE or non-NE cells, the observation of this phenotype is reminiscent of transitional states identified in single cell studies[50] and in-vitro identification of cells with intermediate activation of Notch pathway[25].

To benchmark cluster 2 and to clarify potential inter-relationships, we assessed how features of cluster 2 overlapped with previously described SCLC phenotypes[7,17,41]. We created cluster-specific gene signatures from the top contributors to the first and second PCs (Methods, Table S4). Pairwise correlations of these signatures revealed strong correlation of cluster 1 with the non-NE SCLC and cluster 3 with NE SCLC

signatures. However, cluster 2 was unique and not well described by prior signatures (**Fig. 2I,** S2P). Tumor clusters 1, 2 and 3 are henceforth referred to as non-NE, hybrid-NE, and NE tumor subtypes respectively. This classification is supported by the relative expression of canonical transcription factors, their transcriptional targets, gene expression programs, and matched protein expression.

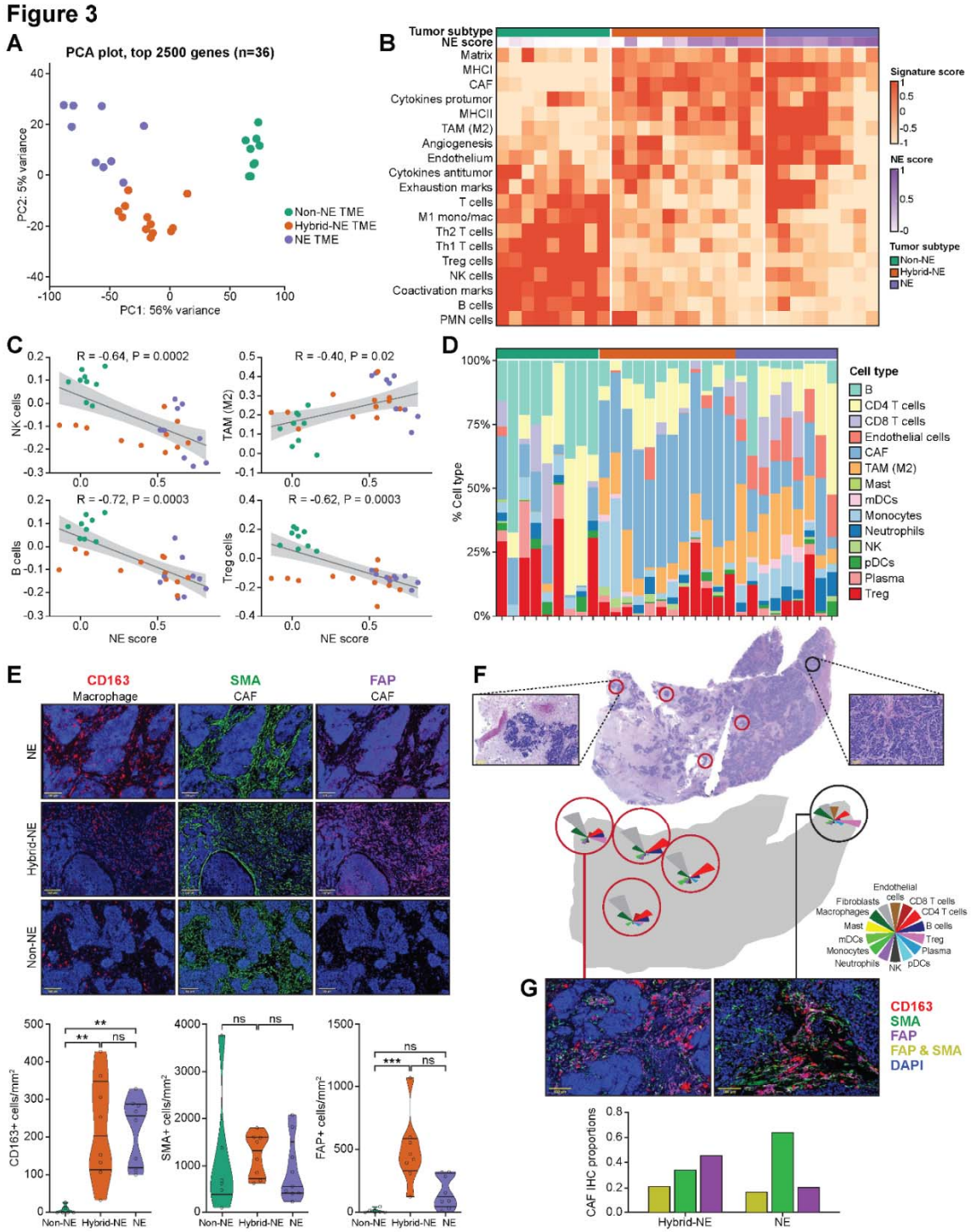


Figure 3: Tumor heterogeneity-linked reprogramming of SCLC TME

A. PCA of 5000 genes with highest variance across the TME segments (n=30).

B. Pan-cancer TME signatures enriched across the TME subtypes[51].

C. Correlation between tumor segment NE scores and immune cell (NK-cell, TAM, B-cell, T-reg) signatures of the spatially proximate TME segment. NE subtypes are indicated in colors.

D. Relative proportion of TME non-malignant cell types sorted by NE subtype of the spatially proximate tumor segments, estimated using CIBERSORT[52].

E. Representative single component multiplex immunofluorescence images showing CD163 (left), SMA (center) and FAP (right) expression across TME subtypes. Quantification shown below cells/mm^{2}[#]. DAPI filter (nuclear stain) is applied on all the images. Scale bar 100 µm.

F. Heterogeneity across spatially proximate hybrid-NE (red circles) and NE (black circle) TME within tumor from patient #5. Top panel shows bird's eye view (4x magnification) H&E image of the tumor section with insets highlighting hybrid-NE (left) and NE (right) regions. Bottom panel shows CIBERSORT-derived relative TME cell type abundance[53]. Inset scale bar 100 µm.

G. Enrichment of FAP+ cells in hybrid-NE TME. Representative (40x magnification) multispectral mIF images of hybrid-NE (left) and NE (right) TME from patient #5 in Fig 3F. Bar plot (below) demonstrating proportion of FAP+ cells, SMA + cells, and combined FAP and SMA+ cells in hybrid-NE (left) and NE (right) TME regions. DAPI (nuclear) filter is on in all the images.

Abbreviations: PCA, principal component analysis; TME, tumor microenvironment; NE, neuroendocrine; mIF, multiplex immunofluorescence; ns, non-significant; DAPI, 4,6-diamidino-2-phenylindole; NE, neuroendocrine; FAP, fibroblast activation protein; SMA, smooth muscle actin; TAM, tumor associated macrophages; T-reg, regulatory T- cells; pDC; plasmacytoid dendritic cells; mDC, myeloid dendritic cells; H&E, Hematoxylin and eosin; *statistical significance at $p<0.05$; **statistical significance at $p<0.01$; ***statistical significance at $p<0.001$; R= spearman's correlation co-efficient; [#]Tukey's multiple comparison test.

Tumor heterogeneity-linked reprogramming of SCLC TME

We next examined the TME in relation to the spatially proximate tumor subtypes. PCA of 5000 most variable genes of the 30 TME segments revealed clear separation to three TME clusters (**Fig. 3A**). NE and hybrid-NE TME were more like each other than non-NE TME.

Applying pan-cancer TME signatures[51], non-NE TME were marked by immune infiltration including natural killer (NK) cells, B cells, and M1 tumor associated macrophages (TAM-M1), and upregulation of immune co-activators (*CD40LG, CD80*) (**Fig. 3B**) compared with hybrid-NE and NE TME[9,44]. Non-NE TME were also enriched for regulatory T-cells (T-regs), immune checkpoints (*PDCD1, LAG3, TIGIT*), and neutrophils. In contrast, hybrid-NE and NE TME were characterized by distinct signatures including cancer associated fibroblasts (CAF) and TAM-M2, matrix remodeling, and pro-tumoral cytokines. The composition of individual immune cell types was correlated with NE differentiation (TAM-M2, r= 0.40, p=0.02; NK cells, r= -0.64, p=0.0002; B cells, r= -0.72, p=0.0003; T-regs, r= -0.52, p=0.0003) (**Fig. 3C**), suggesting reprogramming of the TME and individual immune cell types linked to tumor NE differentiation. Consistent observations were noted when the intratumoral proportion of immune cells was estimated based on CIBERSORT deconvolution[53] (**Fig. 3D**), but additionally revealed a striking enrichment of immunosuppressive CAFs in hybrid-NE TME. Consistent with CAFs constituting the dominant cell-type in hybrid-NE TME, we found significantly lower Shannon index scores[54] in these TMEs

(Fig. S3A), especially compared to NE TME (p=0.0004). Additionally, hybrid-NE TME showed conspicuous absence of CD8+ T cell signatures (Fig. S3B).

To further validate these observations, we performed IHC and multiplex immunofluorescence on sublevel tumor sections to characterize B cells (CD20), T-cells (CD3), fibroblasts (SMA, FAP), macrophages (CD163, CD115, CD11b), and HLA-DR positive cells. Non-NE TME showed higher CD20+ B cell infiltrates compared with NE and hybrid-NE TME, consistent with transcriptomic data (Fig. S3C). Confirming CIBERSORT observations, non-NE and NE TME showed higher CD3+ T cell infiltrates compared with hybrid-NE TME which lacked CD3+ T cells corroborating the earlier observation of reduced intrinsic immune activation in hybrid-NE cells (Fig. 2I). NE and hybrid-NE TME were enriched for TAMs (CD163+ cells) (**Fig. 3E,** S3D) and CD115+ macrophages representing tumor infiltrating mono/macrophages compared to HLA-DR rich regulatory macrophages (Fig. S3E)[55]. All three TME subtypes showed similar proportions of SMA+ fibroblasts. However, hybrid-NE exhibited distinctly increased FAP+ fibroblast signals, while they were absent in non-NE TME.

Intra-tumoral heterogeneity between spatially separated TME regions also correlated with tumor NE subtypes (Fig. S3F). Intra-tumor TME heterogeneity was most evident in tumors that harbored at least one hybrid-NE region (patient #3, #10 and #5). For example, the TME compositions of the NE and hybrid-NE regions from liver metastasis of patient #5 were remarkably distinct (**Fig. 3F**). The hybrid-NE tumors showed budding and nesting, reduced CD8+T cells signatures and enrichment of CAF signatures. Multiplex immunofluorescence confirmed significant enrichment of FAP+ and FAP/SMA+ co-expressing cells in hybrid-NE TME (**Fig 3G,** S3G). Taken together, our observations suggest that NE differentiation is a major determinant of SCLC TME heterogeneity, favoring a highly active immune milieu in non-NE TME. The NE and hybrid-NE TME harbored more primitive immune cell profiles[56], consistent with the known evolutionary trajectory of SCLC from NE to non-NE cell states. Notably, there was remarkable heterogeneity of CAF states across the tumor subtypes, with hybrid-NE subtype specifically enriched for FAP+ CAFs.

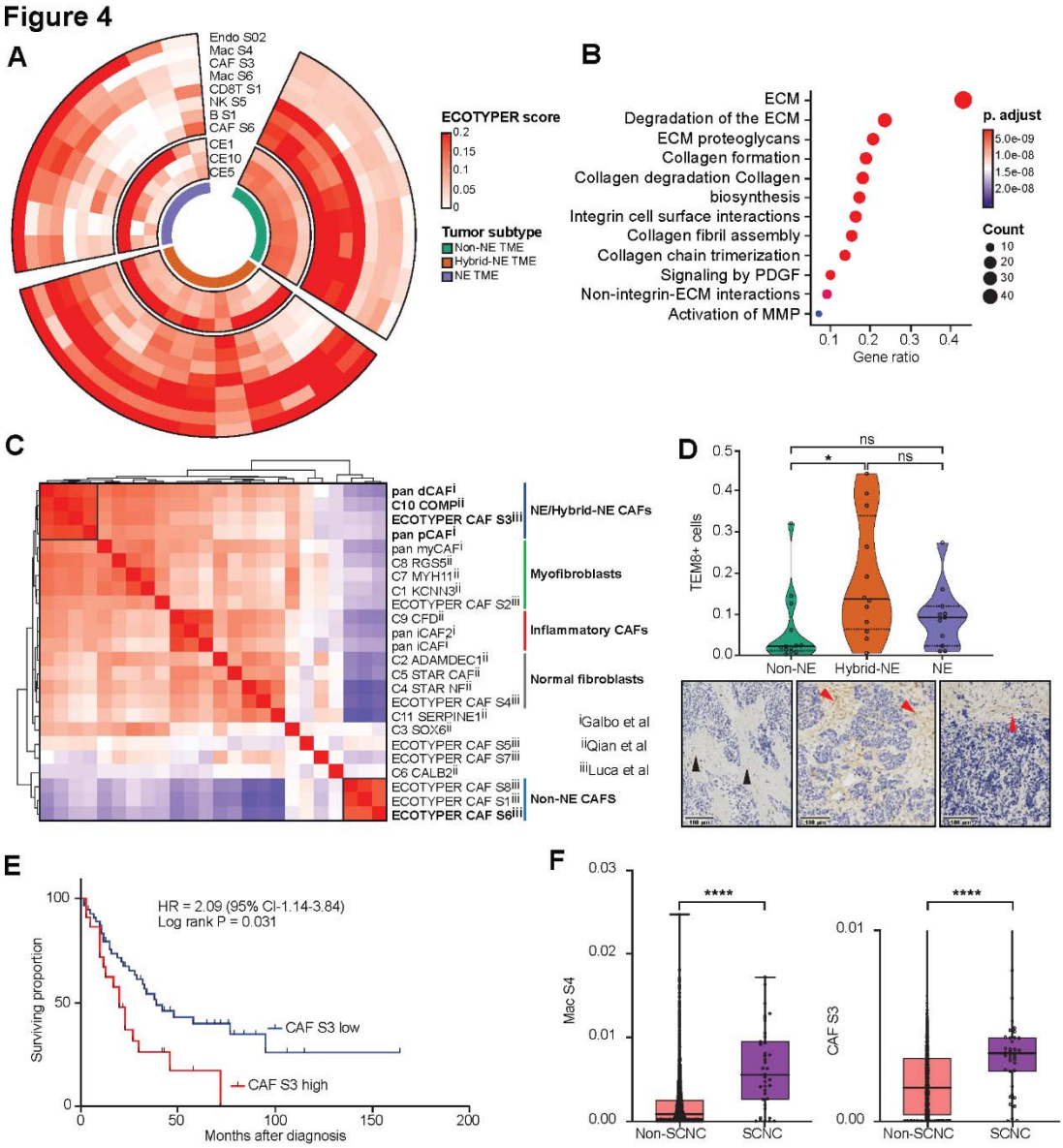


Figure 4. Immunosuppressive CAF cell state enriched in hybrid-NE SCLC

A. Differentially enriched cancer ecotypes and cell states across TME subtypes[52].

B. Expression programs enriched in CAF S3.

C. SCLC mesenchymal cells characterized using CAF signatures[52,57,58] overlayed on sc-RNAseq data of mesenchymal cells extracted from the HTAN dataset[50]. Pair-wise correlation heatmap of ssGSEA-derived enrichment scores are shown.

D. TEM8 protein expression by IHC (% TEM8 expressing cells) across TME subtypes on sublevel sections of tumors profiled using spatially resolved transcriptomics. [%] Representative images showing absence of TEM8 expression in non-NE TME (black arrows) and membranous and

cytoplasmic expression in hybrid-NE and NE TME (red arrows). Color codes as for Fig. 6A. Scale bar set at 100µm.

E. Kaplan Meier curves showing survival (months) of patients with SCLC[5] with high and low CAF S3 expression (cut-off at 75% percentile).

F. Enrichment of CAF S3 (right) and Mac S4 (left) in SCNC pan-cancer[34].[#]

Abbreviations: SCLC, small cell lung cancer; TME, tumor microenvironment; ST, spatial transcriptomics; Endo, endothelial; CAF, cancer associated fibroblasts; Mac, Monocytes/macrophages; ssGSEA, single sample gene set enrichment; SCNC, small cell neuroendocrine carcinoma; OS, overall survival; HR, hazards ratio; CI, confidence interval; IHC, immunohistochemistry; TEM8, tumor endothelial marker 8; ECM, extra cellular matrix; MMP, matrix-metalloproteinases; PDGF- platelet derived growth factor; Diff, Difference; *statistical significance at p<0.05, **** statistical significance at p<0.0001; [%] Tukey's multiple comparison test; [#] student t-test.

Immunosuppressive CAF cell state enriched in hybrid-NE SCLC

We profiled cell states and multicellular communities that organize as functional units[52], referred to as ecotypes, in the spatially resolved gene expression data (**Fig. 4A**, Table S5). NE and hybrid-NE subtypes were enriched for carcinoma ecotype 1 (CE1), associated with lymphocyte deficiency, EMT, and characteristic of cancers with the poorest prognosis[52]. NE and hybrid-NE subtypes additionally showed enrichment of CAF S3 cell state, marked by expression of extracellular matrix, and collagen organization and degradation-related pathways (**Fig. 4B**) and macrophage S4 (Mac S4) (Fig. S4A). In contrast, non-NE TME were enriched for CAF S6, characterized by pathways related to neuronal system, G-coupled protein mediated receptor ligand binding, and peptide ligand binding receptor activity (Fig. S4B).

To contextualize the properties of CAF S3, we performed pairwise correlation of previously published CAF signatures, applying them to human SCLC mesenchymal cells[50] (**Fig. 4C)** and spatially resolved TME segments of current study (Fig. S4C). This analysis revealed a striking similarity of CAF S3 with other aggressive CAF phenotypes such as pan-proliferative CAFs, pan-desmoplastic CAFs[57] and C10 COMP CAFs[58], which in previous studies have been associated with glycolysis, hypoxia, EMT, and metalloproteinase expression. While distinct, these CAF signatures shared expression of several genes (*TEM8, FN1, INHBA, POSTN* and *THY1*; Fig. S4D, Table S6) that have been individually associated with cancer stemness and tumor aggressiveness[59,60]. Tumor endothelial marker 8 (TEM8), also known as anthrax toxin receptor 1 (ANTXR1) is a highly conserved transmembrane receptor broadly overexpressed on CAF, endothelium, and pericytes. TEM8 is also a receptor for Seneca Valley virus, an oncolytic picornavirus, previously described to have selective tropism non-NE SCLC[61]. We profiled sublevel sections of the spatially profiled tumors using TEM8 IHC and confirmed the strong expression of TEM8+ cells in hybrid-NE TME (**Fig. 4D**). We also found heterogeneity of TEM8 expression in spatially separated areas within the same tumor, with expression limited to hybrid-NE TME (Fig. S4E; 2H,2I).

Importantly, SCLC patients[5] whose tumors contained high proportions of CAF S3 had significantly worse outcomes [HR=2.09 (CI 95%- 1.14-3.84, log-rank p=0.031)] compared to those with lower number of CAFs (**Fig. 4E**). CAF S3 remained an independent risk factor for death after controlling for clinical variables including age, sex, and stage (Fig. S4F). CAF S3 and Mac S4 (**Fig. 4F)** cell states were also enriched in pan-cancer small cell neuroendocrine tumors[34], suggesting that these states may in part underlie the aggressiveness of small cell cancers regardless of tissue of origin.

Taken together, these analyses reveal CAF heterogeneity in SCLC TME, with hybrid NE subtype showing remarkable enrichment of CAF S3, marked by high expression of TEM8, with immunosuppressive and metastases aiding capabilities and portending poor prognosis.

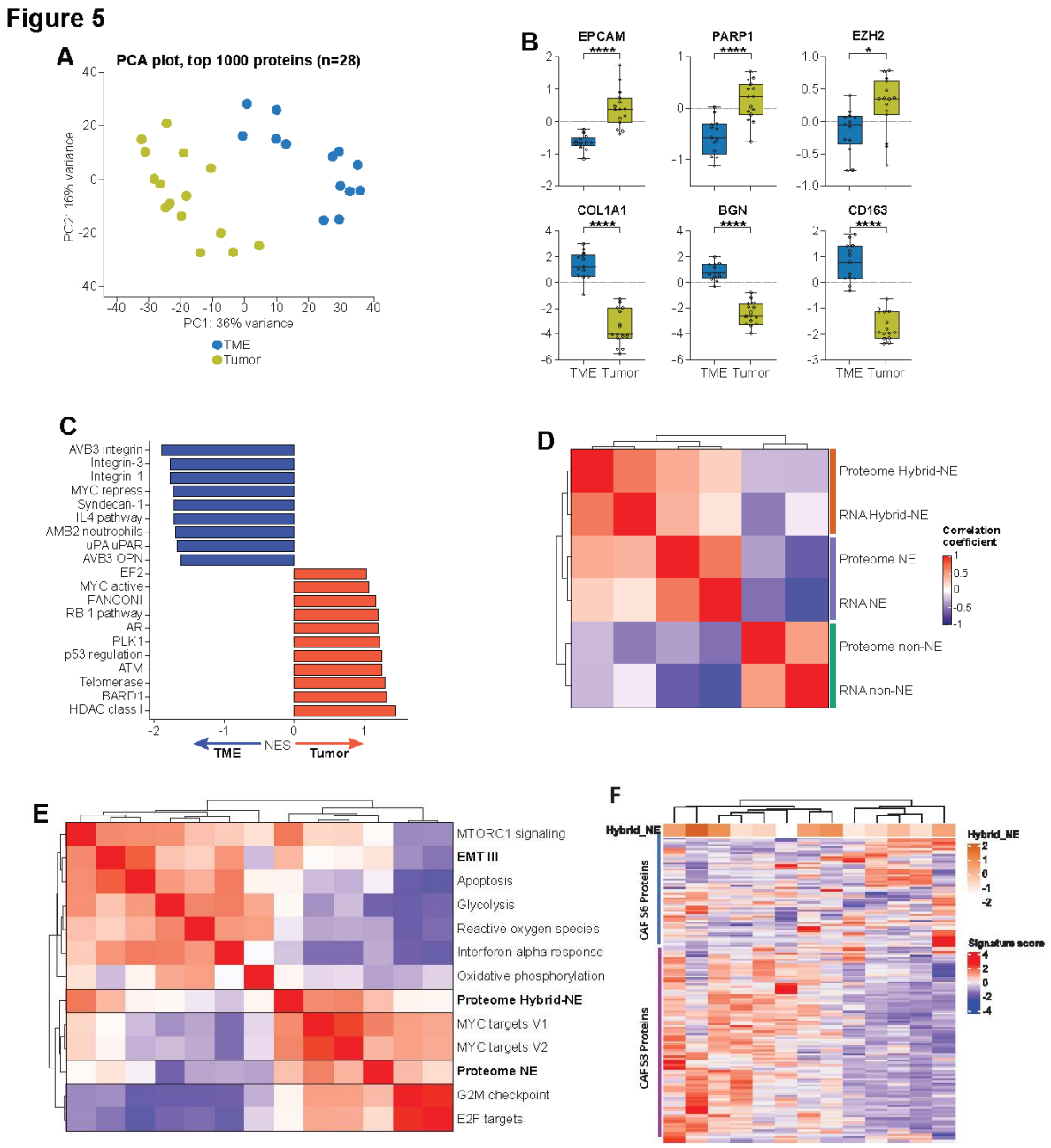


Figure 5: Proteomic characterization of tumor heterogeneity-linked reprogramming of SCLC TME

A. PCA of 1000 proteins with highest variance between tumor (n=15) and TME (n=13).

B. Distribution of tumor (above) and TME (below)-associated proteins. [#]

C. NES (GSEA) of differentially expressed pathways between tumor and TME.

D. Tumor transcript-protein correlation for NE and hybrid NE subtypes. Pairwise correlation of matched tumor proteome and bulk RNA-seq (n=12) derived NE, non-NE, and hybrid-NE signature scores.

E. Pairwise correlation of tumor proteome-derived NE and hybrid-NE signatures with selected hallmark pathways (ssGSEA).

F. Heatmap showing CAF S3 and CAF S6 proteins enrichment in TME proteome and their association with hybrid NE signatures of corresponding tumor proteome.

Abbreviations: SCLC, small cell lung cancer; BGN, biglycan; EZH2, enhancer of zeste homolog 2; MS, mass spectrometry; PCA, principal component analysis; GSEA, gene set enrichment analysis; TME, tumor microenvironment; ssGSEA, single sample GSEA; NE, neuroendocrine; ns, non-significant at $p<0.05$; *statistical significance $p<0.05$; **** statistical significance at $p<0.0001$; [#] student t-test.

Proteomic characterization of tumor heterogeneity-linked reprogramming of SCLC TME

In parallel with spatially resolved transcriptomics, we used laser capture microdissection to separately enrich tumor and TME with the goal of proteomic characterization of SCLC heterogeneity (Fig. S5A). Fifteen tumor and 13 TME segments from rapid autopsy derived tumors of 11 patients were examined (see Methods, Table S7). Twelve tumors had matching bulk RNA-seq data, and six tumors had matching spatially resolved transcriptomic data described earlier(Fig. S5B). The histology-resolved tumor and TME were analyzed by quantitative MS-based proteomics[62], resulting in 7418 and 6655 total proteins quantified in tumor and TME, respectively, and 6155 common proteins across both compartments (Fig. S5C). Enrichment of TME and tumor compartments was effective; tumor and TME samples separated widely on PCA for 1000 proteins with highest variance (**Fig. 5A**). TME segments were significantly enriched for stromal proteins (e.g., COL1A1, BGN, CD163) (**Fig. 5B**) and showed upregulation of stromal pathways (e.g., integrins, syndecan, and interleukin 4) (**Fig. 5C**). Tumor segments were significantly enriched for tumor-specific proteins (e.g., EPCAM, PARP1, EZH2), had higher proteomics-derived NE scores (Fig. S5D), and showed upregulation of tumor-related pathways e.g., DNA replication and repair (ATM, Fanconi, BARD1), tumor suppressors (RB1, E2F targets, P53 regulation) and MYC. As expected, tumor proteomes from different metastatic sites of the same patient were more like each other than tumors from other patients (Fig. S5E).

We then integrated proteomic data with patient-matched spatial transcriptomics data. A significant but modest correlation was observed between the overall tumor protein abundance and gene expression (Fig. S5F, median spearman correlation co-efficient 0.41; range, 0.11-0.47), in line with previously reported transcript-protein abundance correlations[63]. The correlation was more modest for TME segments (median 0.13; range, 0.05-0.26), possibly due to greater cellular heterogeneity of the TME compartment or reduced cellular density. In contrast to the modest correlation across all genes, we found strong correlation between proteomic and RNA-seq derived NE and hybrid-NE signatures (**Fig. 5D,** Fig. S5G). Proteome-based hallmark pathway analyses showed enrichment of DNA repair and replication stress pathways, E2F targets, and G2M checkpoints[10] in NE SCLC. Hybrid-NE proteome was unique, reminiscent of gene expression patterns described earlier (Fig. 2H), showing features of EMT (**Fig.5E**). Due to the low amount of available starting material and resultant low coverage of proteomes, only selected hallmark pathways with at least 50% coverage were included in these analyses (Table S8). TME proteome of hybrid-NE tumors showed enrichment of multiple CAF S3 compared with CAF S6 proteins

(**Fig. 5F**). Additionally, TME proteome analyses supported earlier observations from spatially resolved transcriptomics including TAM-M2 enrichment and B/plasma cell de-enrichment with increasing NE differentiation (Fig. S5H) and CD8+ T cell exclusion in hybrid-NE enriched tumors (Fig. S5I). CD8+ T cell exclusion associated with hybrid-NE enrichment (Fig. S5J) was further confirmed in bulk RNA-seq datasets[9,17].

Thus, histology-resolved proteomics confirmed key observations from spatially resolved transcriptomics, including the strong enrichment of hybrid-NE state with EMT-III, the aggressive CAF S3 subtype, and exclusion of CD8+ T cells. The high correlation between RNA and the corresponding protein levels suggests that RNA serves as a valuable indicator of protein expression for genes that contribute to SCLC NE heterogeneity.

Figure 6

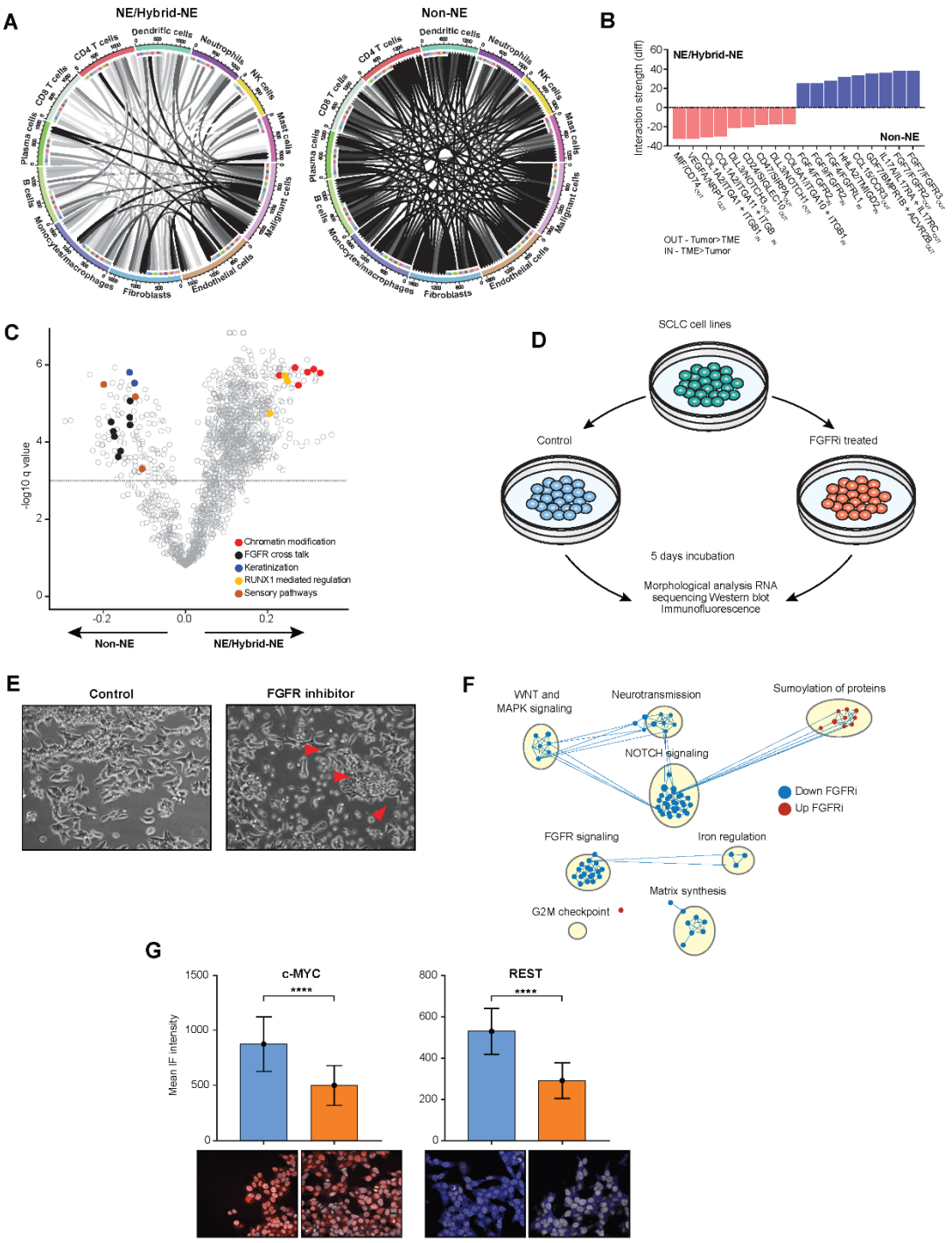


Figure 6. SCLC cell states modulated by FGFR inhibition.

A. Intercellular tumor-TME interactions[52,64] of NE/hybrid-NE (left) and non-NE tumor-TME (right) ecosystems.

B. Differentially over-represented ligand-receptor interactions in NE/hybrid-NE and non-NE tumor ecosystems. Directionality of interaction specified for each pairs as in (TME-> tumor) and out (Tumor->TME).

C. Differentially enriched expression programs (ssGSEA) between non-NE and NE/hybrid-NE TMEs. Recurrent and SCLC-relevant programs are highlighted.

D. Workflow of FGFR inhibition in SCLC cell lines using pan-FGFR inhibitor erdafitinib.

E. DMS-273 cell line treated with FGFR inhibitor (erdafitinib, 33.33 nM). Representative microscopic images (40x magnification) of untreated (left) and treated cells (right). Red arrows showing cells in suspension state.

F. Pathways altered following treatment of DMS-273 with FGFR inhibitor (erdafitinib, 33.33 nM). Network enrichment plot of GSEA shown. Blue dots indicate pathways downregulated and red dots indicate upregulated pathways compared with control cells.

G. c-MYC (left) and REST (right) mean IF intensity following treatment of DMS-273 with FGFR inhibitor (erdafitinib, 33.33 nM). [%]

Abbreviations: FGF8, fibroblast growth factor 8; RNA ISH, ribonucleic acid in situ hybridization; FGFR, fibroblast growth factor receptor ssGSEA; single sample gene set enrichment analysis, NE, neuroendocrine; nM/L, nanomoles/liter; IF, immunofluorescence; **p<0.01; ****p<0.0001; ns, not significant; R, Spearman's correlation coefficient; [#]Tukey's multiple comparison test; [%] student t-test

Tumor-TME crosstalk and tumor state modulation by FGFR inhibition

To determine spatially proximate communications that shape SCLC tumor cell states, we reconstructed cell-cell interactions based on coordinated expression of receptor-ligand pairs in TME and tumor[64]. We combined NE and hybrid-NE subtypes since TME of these two subtypes shared functional similarities (Fig. 3B, 4A).

Non-NE subtype exhibited strikingly higher number and diversity of interactions with TME compared with NE and hybrid-NE subtypes (**Fig. 6A**, S6A)[64,65]. Interactions significantly enriched in non-NE included immune checkpoint-receptor (e.g., PDCD1/PDCD1LG2), cytokine-receptor pairs (e.g., CCL15/CCR3)[9,44], and multiple pathways of the FGF/FGFR signaling system especially FGF8 with multiple FGFR partners (**Fig. 6B,** Fig. S6B**,** Table S9). NE and hybrid-NE subtypes had far fewer interactions, but top interactions involved fibroblasts, macrophages, and endothelial cells further supporting the key role of these cell types in the NE and hybrid-NE ecosystems. These included MIF/CD74[66], CD24/SIGLEC10[67] and CD47/SIRPA[68] signaling pathways. Unsupervised examination of differentially enriched programs between non-NE and NE/hybrid-NE TME corroborated with above observation, revealing multiple FGF/FGFR signaling pathways upregulated in non-NE TME (**Fig. 6C**, Table S10). To further validate our findings, we performed *FGF8* RNA in-situ hybridization on sublevel tumor sections (n=4). Highest *FGF8* signals were observed in non-NE TME (Fig. S6C). Consistently, we also found evidence of high FGFR activity in non-NE SCLC tumor segments supporting unusually high TME to tumor FGFR signaling (Fig.S6D).

Considering that FGFR signaling plays a pivotal role in the growth and differentiation of normal tissues and during embryogenesis[69], we postulated that TME-derived FGF signaling might be crucial for maintaining the non-NE SCLC cell fate. It is important to note that while cell lines lack TME, cell culture media typically contain variable amounts of FGF in fetal bovine serum[70]. We treated non-NE SCLC cell lines NCI H211 and DMS-273 with varying concentrations of FDA approved pan-FGFR inhibitor erdafitinib

for 5 days (**Fig. 6D,** Methods). Treatment with FGFR inhibitor led both cell lines to transition from predominantly adherent to suspension states (**Fig. 6E**), a growth pattern typically associated with NE differentiation[7,8].

Network analysis of transcriptomic data revealed robust downregulation of various FGFR signaling pathways (**Fig.6F**), and well as non-NE signaling pathways such as MAPK[71] and NOTCH signaling[25], along with EMT. Correspondingly, western blot and immunofluorescence showed reduced expression of c-MYC which drives the non-NE cell fate[24] and REST, a repressor of neural gene expression (**Fig. 6G,** Fig. S6E), indicating the dependence of SCLC non-NE cell state on extrinsic FGF signaling. Notably, there was no observable upregulation of NE markers at the protein level. This finding is reminiscent of irreversible fate switches induced by Notch[72], a recurrent mechanism implicated in lineage switching in SCLC[25]. While Notch activation has been demonstrated to promote an NE to non-NE transition in SCLC, studies suggest that Notch blockade cannot completely reverse cells back to an NE- state[24,25]. Caspase activation and cell viability assays (Fig. S6F, S6G) did not reveal a significant increase in apoptosis with erdafitinib, indicating that the reduction in non-NE features is not attributable to apoptotic cell death.

Discussion

SCLCs epitomize cancers with exceptional chemoresistance and metastatic capabilities, driven by significant intratumoral heterogeneity. Tumor-intrinsic factors that heighten SCLC plasticity include Notch pathway activation[25] and MYC hyperactivation[24] through amplification of extra chromosomal elements[26]. However, the cell-extrinsic determinants governing SCLC heterogeneity remain inadequately understood. Our comprehensive approach, integrating histopathology with spatially resolved transcriptomics and MS-based proteomics, whole-genome sequencing and multiplex immunofluorescence, maps SCLC tumor states within their spatial context. Our work highlights the following key findings: (i) recognition of CAFs as a crucial element contributing to SCLC TME heterogeneity, with aggressive CAF subpopulations enriched in the hybrid-NE TME, predicting an exceptionally poor prognosis; (ii) substantial variation in the phenotype and overall composition of non-malignant cells in the metastatic TME, closely mirroring the tumor NE state, indicating TME-driven reprogramming of NE cell states. Notably, we find that TME-derived FGF signaling directs the tumor cells towards a non-NE cell state. Collectively, our work provides crucial insights into SCLC heterogeneity and the pivotal role of the TME in facilitating SCLC plasticity.

Previous studies have highlighted CAF heterogeneity in various solid tumors, characterized by differential marker expression and functional roles, including the promotion of cancer invasion and metastasis through the secretion of soluble factors and matrix remodeling[73]. To our knowledge, CAFs have not yet been studied in detail in SCLC. In the present study, we find CAF subpopulations with heterogenous marker expression even within the same tumor. The complexity and heterogeneity of CAFs appear to be closely interconnected with the tumor state. Particularly, the hybrid-NE TME is characterized by an abundant presence of FAP/POSTN-expressing CAFs, akin to previously described CAF subtype S3[52]. These CAFs exhibit enrichment for extracellular matrix and pathways associated with collagen organization and degradation. Correspondingly, we observed pronounced stromal expression of both nuclear and cytoplasmic YAP1 in hybrid-NE SCLC, consistent with the recognized role of YAP1 in the establishment and maintenance of CAFs[74], as well as its activation in cancer cells in close proximity to a rigid matrix[74,75]. The factors driving CAF heterogeneity and the implications of CAF infiltration in SCLCs, particularly regarding their immune regulatory functions, warrant further investigation. Our findings align with a recent study that reported that SCLC display the highest FAP expression compared to various other solid tumors[76].

Targeting the distinctive immunosuppressive mechanisms inherent to each cell state holds promise for reinstating immunosurveillance and enhancing SCLC responses to immunotherapy. This approach is particularly pertinent in SCLC, which, despite exhibiting a highly mutated genome, displays only modest responses to immunotherapy[77,78]. TME-targeted therapies may be broadly relevant for neuroendocrine cancers across tissues of origin, which are enriched for immunosuppressive CAFs. While early efforts to target CAF-cancer cell interactions in certain cancer types have hinted at the potential drawbacks of stromal ablation, such as the development of more aggressive cancers[79,80], a nuanced understanding of CAF diversity and their influence on immunosurveillance could inform the development of personalized treatment strategies for SCLC.

Our work provides crucial insights into the role of the TME in shaping the SCLC phenotype, an area that has remained poorly understood primarily due to the limited availability of patient tumor samples for research. Tumor samples obtained through rapid research autopsies from patients with metastatic SCLC provide the most compelling evidence to date of extensive transcriptional heterogeneity heavily influenced by the TME. In contrast to the NE and hybrid-NE subtypes, we find intimate interactions between the non-NE subtype and the adjacent TME, actively reprogramming the tumor states. Notably, previous studies in model systems have implicated FGF signaling as a crucial component mediating the cross-talk between NE and non-NE cell states[8,81], as well as a defining feature of chemotherapy-resistant persister cells[82]. Our findings demonstrate that, in human tumors, FGF signaling can originate from the TME, sustaining a chemo-resistant non-NE state. These results align with the previously established role of FGF signaling in driving SCLC development, particularly non-NE cells[83], and offer an explanation for the observed sensitivity of chemo-resistant non-NE SCLC cells to FGFR inhibition[84,85]. Furthermore, a dependency on FGF signaling has been reported in a subset of castration-resistant prostate cancers characterized by limited or absent androgen receptor expression and non-NE features[86]. Collectively, these findings underscore TME-derived FGF signaling as a shared pathway underlying the lineage plasticity of NE cancers across various contexts and may guide future endeavors to modulate tumor-TME crosstalk and restrain tumor evolution.

Limitations of the study

To our knowledge, this study represents the first comprehensive report of tumor-extrinsic drivers of SCLC plasticity in human tumors and provides unprecedented insights into CAFs as key TME elements in SCLC. Nevertheless, several notable limitations should be acknowledged. These include the lack of single-cell resolution at a spatial level, which prevents definitive conclusions regarding the precise localization of the hybrid NE state within the tumor edge. We recognize that this manuscript has not validated the functional impact of the newly described hybrid-NE state and the CAFs, including their impacts on chemo-resistance and metastasis. The study was limited by the relatively small number of samples, but the validation using orthogonal proteomic, and sc-RNA based approaches in independent datasets strengthen our conclusions. Further studies using tumors from a wider range of metastatic sites are needed to understand how the TME programs are influenced by the site of metastasis.

Acknowledgement

We gratefully acknowledge the contributions of our patients and their families. This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). We thank NanoString GeoMx's early access whole transcriptome atlas program (Jingjing Gong, Stephan Phelan).

Supplemental Figure legends

Figure S1

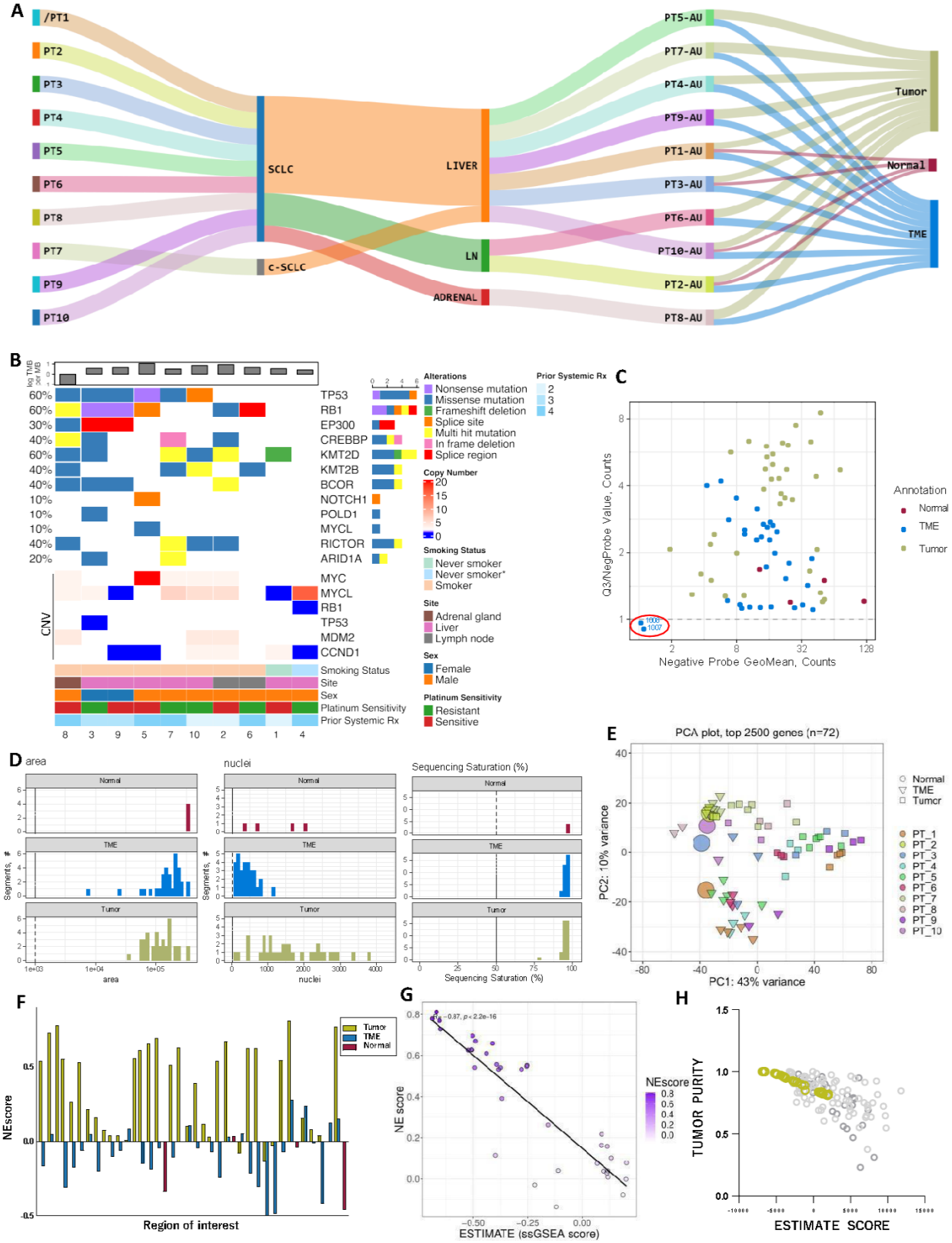


Figure S1

A. Overall workflow of ST approach for spatially profiled tumors (n=10).

B. Mutational (above) and CNV (below) landscape of SCLC tumors (whole genome sequencing) profiled using ST.

C. Quality metric of data assessed using mean Q3 value (of all 18,776 genes) to the negative probe Q3 value (y-axis) indicating TME segments 1007 and 1008 as outliers with relatively low-quality data (removed from subsequent analyses).

D. Bar-plots showing segment area of capture (left), number of nuclei (center), and sequencing saturation (right) for each tumor, TME and normal segments. Color code as Fig. S1C.

E. PCA plot showing PC1 vs PC2 like Fig 1B, additionally highlighting the precise location of normal (n=4), tumor(n=36) and TME(n=30) segments and colored by each patient origin. Overall normal segments clustered close to patient matched TME segments as opposed to tumor segments.

F. Bar plot showing NE score (ssGSEA) for each tumor, TME and normal segments profiled for each region.

G. Correlation between tumor segment NE scores and their stromal and immune scores[36]. (Spearman correlation coefficient, r= -0.87).

H. Scatter plot showing the correlation between the tumor purity score projections (derived from linear regression modeling of bulk RNA and WGS sequencing data[5,9] using ABSOLUTE approach[38]to calculate tumor purity estimates and the stromal and immune scores[36]for spatially profiled tumor segment. Color code as Fig. S1C.

Abbreviations: ST- spatial transcriptomics; c-SCLC, combined small cell lung carcinoma; AU, autopsy; LN, Lymph node; CNV, copy number variations; Q3, 3rd Quantile normalized count, TME- tumor microenvironment, PC- principal component, PCA- principal component analysis, NE- neuroendocrine, Rx, treatment; ssGSEA, single sample gene set enrichment analysis; *non-smoker but very heavy exposure to asbestos.

Figure S2

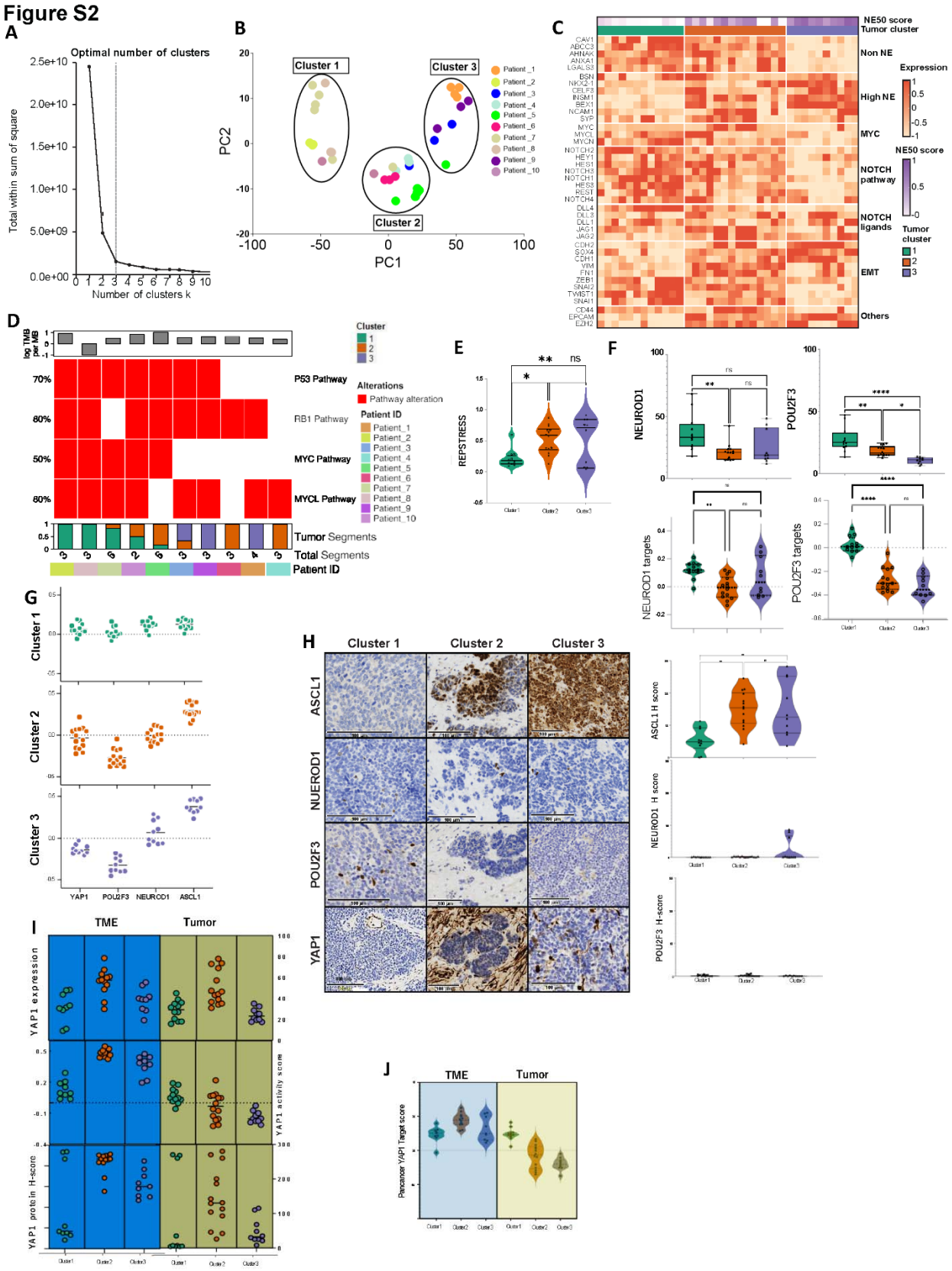

A
Optimal number of clusters
Total within sum of square
Number of clusters k
B
Cluster 1
Cluster 2
Cluster 3
PC1
PC2
C
NE50 score
Tumor cluster
Non NE
High NE
MYC
NOTCH pathway
NOTCH ligands
EMT
Others
Expression
D
log TMB per MB
Cluster
P53 Pathway
RB1 Pathway
MYC Pathway
MYCL Pathway
Alterations
Pathway alteration
Patient ID
Tumor Segments
Total Segments
E
REPSTRESS
F
NEUROD1
POU2F3
NEUROD1 targets
POU2F3 targets
G
Cluster 1
Cluster 2
Cluster 3
YAP1
POU2F3
NEUROD1
ASCL1
H
ASCL1
NUERODI
POU2F3
YAP1
ASCL1 H score
NEUROD1 H score
POU2F3 H-score
I
TME
Tumor
YAP1 expression
YAP1 activity score
YAP1 protein H-score
J
Pancancer YAP1 Target score

Figure S2 (continued)

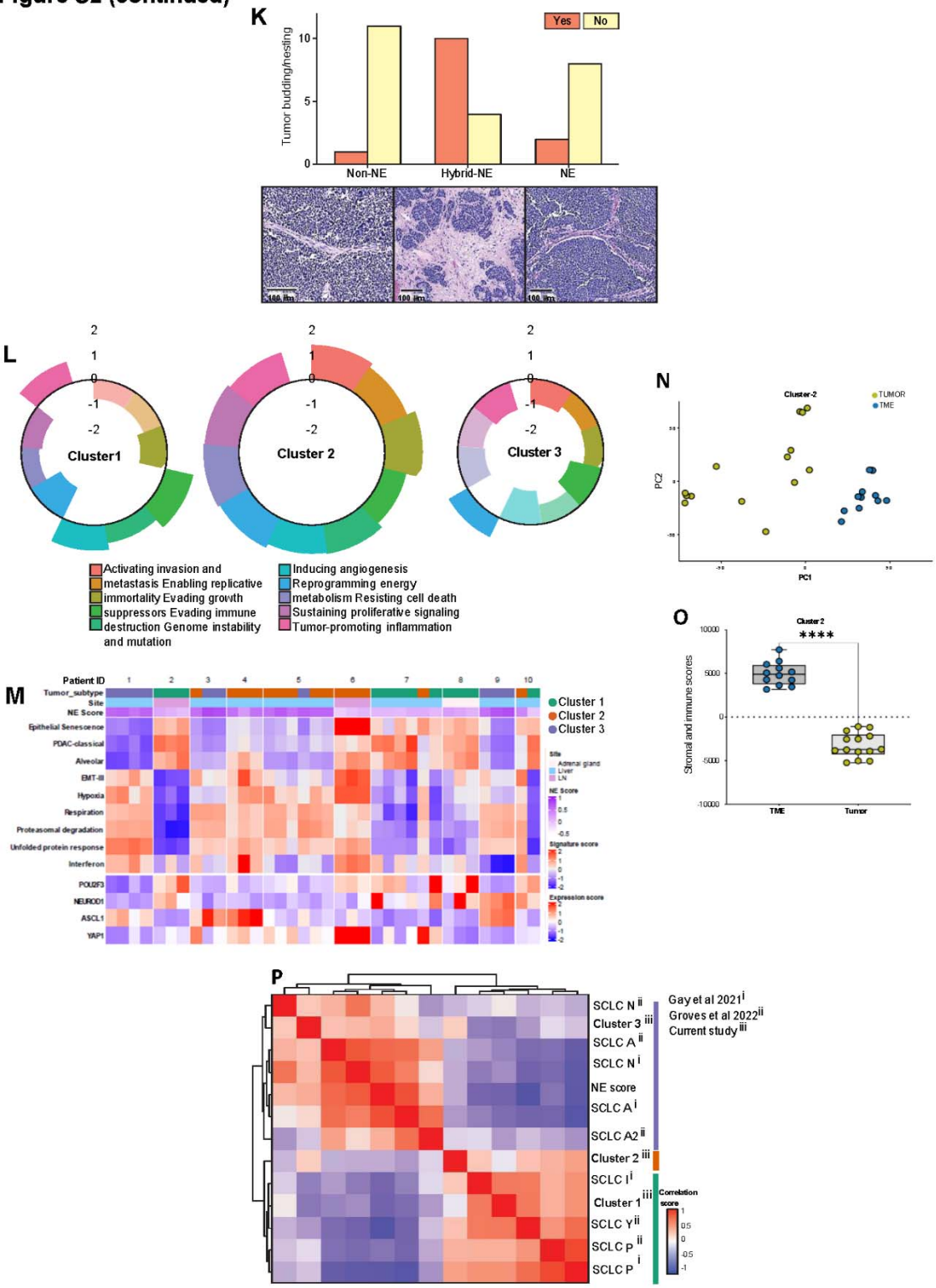


Figure S2:

A) Intra-cluster variation of the tumor segments (n=36) (as total within-sum of squares) (y-axis) plotted for each partitioning constant (k) for k=1-10. Relatively flattening of curve noted at k=3.

B) PCA as Fig 2A colored by patient origin to demonstrate intra-tumoral heterogeneity of tumor segments (n=36).

C) Heatmap showing distribution of SCLC related genes across the 3 clusters.

D) Oncoplot demonstrating spatially determined tumor cluster phenotype proportions in each (patient) tumor and presence of *RB1, TP53* pathway alterations (loss of function events) as well as *MYC* and *MYCL* alterations (amplification events) determined by bulk WGS.

E) Box- plots showing REPSTRESS[43] scores (ssGSEA derived) in three clusters.[#]

F) Gene expression counts (above, 3rd quantile normalized values) and target activity scores (below, ssGSEA) of *NEUROD1* (left) and *POU2F3* (right) in three clusters.[#]

G) Cluster-wise landscape of SCLC lineage-defining transcription factor activity scores (ssGSEA derived).

H) Representative images at high power (40X magnification) showing protein expression of SCLC lineage-defining transcription factors (ASCL1,NEUROD1, POU2F3 and YAP1) across the tumor clusters. Scale bar at 100 µm. Quantification (right) showing expression of ASCL1, NEUROD1 and POU2F3 protein in 3 tumor clusters as H-scores (range from 0-300).[#]

I) Dot plots showing *YAP1* RNA expression (top, 3rd quantile normalized values), *YAP1* TF activity score (middle, ssGSEA derived) and IHC H-scores (bottom) across both TME and tumor segments demonstrating increased YAP1 activity in TME segments of cluster 2 followed by cluster 3.

J) Validation of distinct *YAP1* TF activity patterns across tumor and TME segments using independent pan-cancer *YAP1* signature[46].

K) Frequency of tumor budding in cluster 2 compared with cluster 1 and Cluster 3 SCLC. Representative images shown below. Scale bar at 100 µm.

L) Cancer hallmarks differentially enriched (GSEA derived) across the spatially profiled tumor clusters. The height of each bar shows NES.

M) Expression heatmap of transcript-defined cancer meta-programs[39] and SCLC lineage-defining transcription factors across different tumor segments from individual patient tumors (n=10).

N) PCA plot subsetted for only Cluster 2 regions colored by tumor and TME segments showing distinct clustering of tumor and TME (2500 most differentially expressed genes).

O) Stromal and immune (ESTIMATE) scores[36] showing negative enrichment of these scores in Cluster 2 tumor segments as opposed to Cluster 2 TME segments.

P) Correlation matrix showing pairwise correlation for different SCLC related signatures with signatures generated for spatially profiled tumor segment clusters for SCLC cell lines (n=52)[87](like Fig. 2I).

Abbreviations: ST- spatial transcriptomics; WGS- whole genome sequencing, ssGSEA- single sample gene set enrichment analysis, IHC- immunohistochemistry, SCLC- small cell lung cancer; REPSTRESS, Replication stress; TF, transcription factor; *statistical significance at $p<0.05$; **statistical significance at $p<0.001$; ***statistical significance at $p<0.001$; ****statistical significance at $p<0.0001$; [#] Tukey's-multiple comparison test; NES-normalized enrichment score

Figure S3

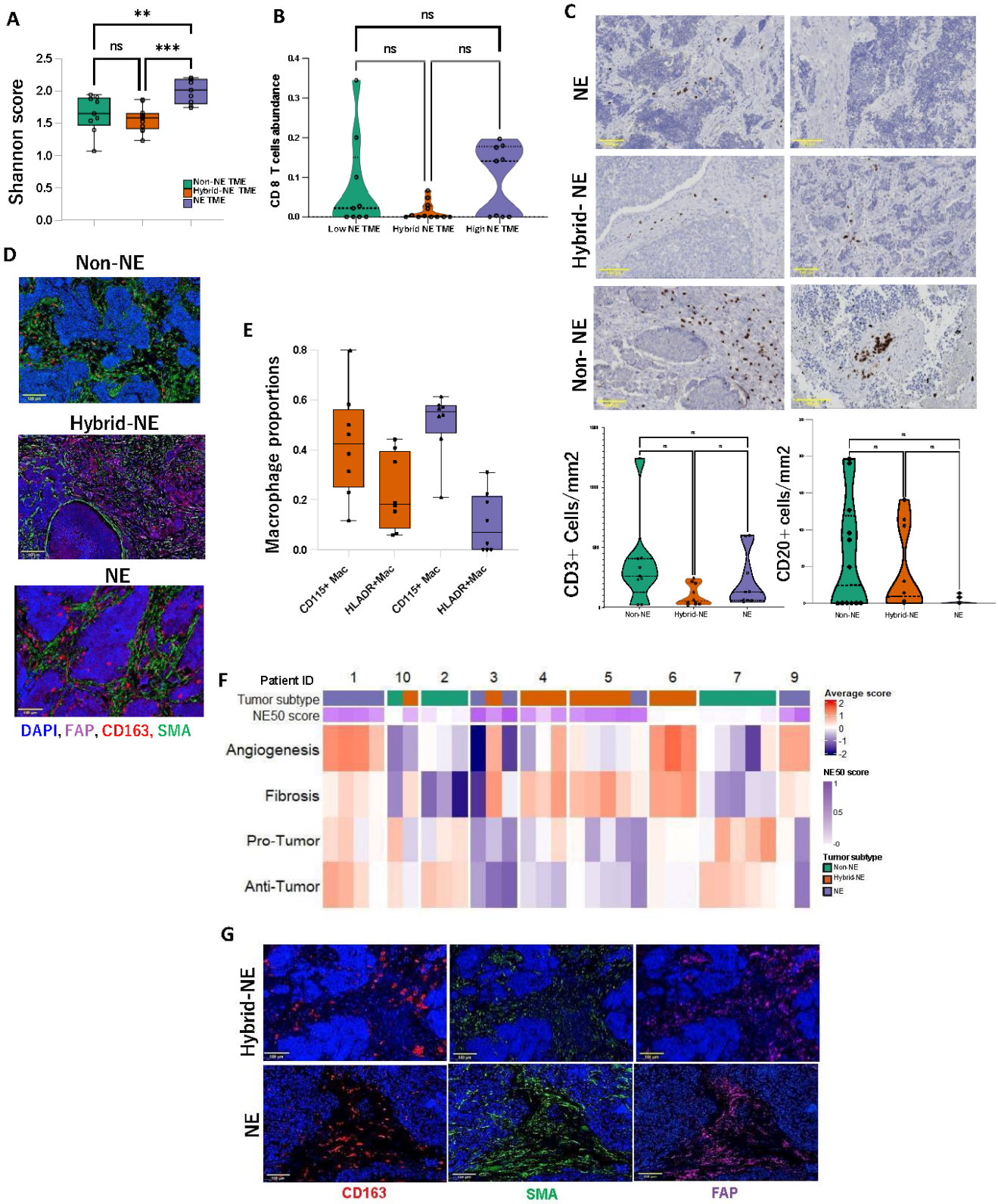

A
Shannon score
Non-NE TME
Hybrid-NE TME
NE TME
B
CD 8 T cells abundance
Low NE TME
Hybrid NE TME
High NE TME
C
NE
Hybrid- NE
Non- NE
CD3+ Cells/mm2
CD20+ cells/mm2
D
Non-NE
Hybrid-NE
NE
DAPI, FAP, CD163, SMA
E
Macrophage proportions
CD115+ Mac
HLADR+Mac
F
Patient ID
Tumor subtype
NE50 score
Angiogenesis
Fibrosis
Pro-Tumor
Anti-Tumor
Average score
Tumor subtype
G
Hybrid-NE
NE
CD163
SMA
FAP

Figure S3:

A) Shannon scores for CIBERSORT deconvoluted cell types for TME segments. [#]

B) CD8+ T cells abundance score (CIBERSORT) for each TME subtype. Color codes as Fig. S3A.[#]

C) Representative IHC images performed on sub-level sections (40x magnification) for CD3+ T cells (left) and CD20+ B cells (right) in different corresponding NE subtype tumor segments. Violin plots (below) demonstrating abundance of CD3+ T cells (left) and CD20+ B cells (right) [#]. Scale bar is set at 100µm.

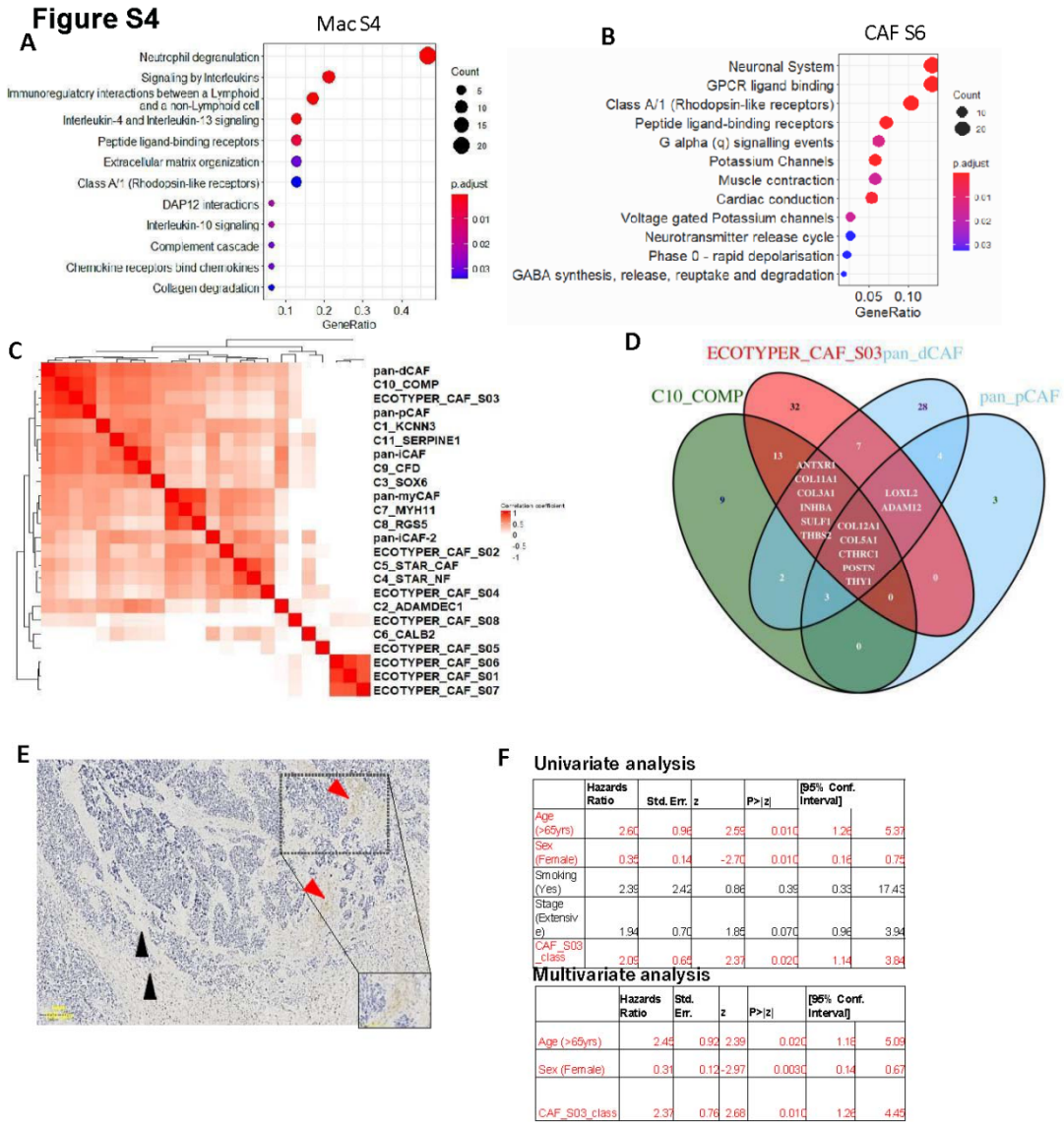
D) Representative multispectral IF images corresponding to Fig. 3E (filters on for DAPI, CD163, SMA and FAP), non-NE (top), hybrid-NE (center), NE (bottom). Scale bar is set at 100µm.

E) Macrophage subtype proportions in NE and hybrid-NE TME.

F) Heatmap showing average expression scores of pan-cancer TME features[51] (reduced to 4 major TME features- fibrosis, angiogenesis, pro-tumor immune factors and anti-tumor immune factors) and clustered patient wise to demonstrate TME ITH. Color code of TME subtypes as Fig. S3A. Number labels on top indicate patient ID.

G) Single component (40x magnification) mIF images corresponding to Fig. 3G (patient#5 tumor) demonstrating individual staining of CD163 (left), SMA (center) and FAP (right). DAPI (blue) filter is on in all the images. Scale bar set at 100µm.

Abbreviations: TME, tumor microenvironment; IHC, immunohistochemistry; ITH, Intra-tumoral heterogeneity; DAPI,4□,6-diamidino-2-phenylindole; NE, neuroendocrine; FAP, Fibroblast activation protein- alpha; SMA, smooth muscle actin; Mac, macrophages; ns, statistically non-significant; **statistical significance at $p<0.001$; ***statistical significance at $p<0.001$; [#] Tukey's- multiple comparison test.



	Hazards Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Age (>65yrs)	2.60	0.96	2.59	0.010	1.26	5.37
Sex (Female)	0.35	0.14	-2.70	0.010	0.16	0.75
Smoking (Yes)	2.39	2.42	0.86	0.39	0.33	17.43
Stage (Extensive)	1.94	0.70	1.85	0.070	0.96	3.94
CAF_S03_class	2.05	0.65	2.37	0.020	1.14	3.84

	Hazards Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Age (>65yrs)	2.45	0.92	2.39	0.020	1.16	5.09
Sex (Female)	0.31	0.12	-2.97	0.0030	0.14	0.67
CAF_S03_class	2.37	0.76	2.68	0.010	1.26	4.45

Figure S4:

A) Programs enriched in Mac S4 cell state.

B) Programs enriched in CAF S6 cell state.

C) Pairwise-correlation plot of ssGSEA-derived enrichment scores of CAF signatures from different studies[52,57,58] in TME segments of spatially profiled tumors of the current study, like Fig 4C.

D) Common and distinct genes of different CAF types clustering together with ECOTYPER CAF S3.

E) Low power (20x magnification) IHC image of TEM8 (ANTXR1) in patient #10 tumor with inset showing positive TEM8 staining in TME of hybrid-NE segment (red arrows) same area corresponding to Fig. 2H and Fig. 2I. Negative staining in areas corresponding to non-NE TME segments (black arrows). Scale bar set at 100µm.

F) Univariate (above) and multivariate (below) survival analysis of bulk transcriptome dataset in SCLC with available survival data[5] considering CAF S03 high and low class (see methods). Cox proportional hazard algorithm used for survival analyses.

Abbreviations: Mac, macrophage; Mac S4- Macrophage cell state 4; CAF, cancer associated fibroblasts, Endo S2, Endothelial cell state 2; ssGSEA- single sample gene set enrichment analysis, TME- tumor microenvironment, SCNC- small cell neuroendocrine carcinoma, NE- neuroendocrine, IHC- immunohistochemistry, TEM8- tumor endothelial marker 8, std. error- standard error, **** statistical significance $p<0.0001$.

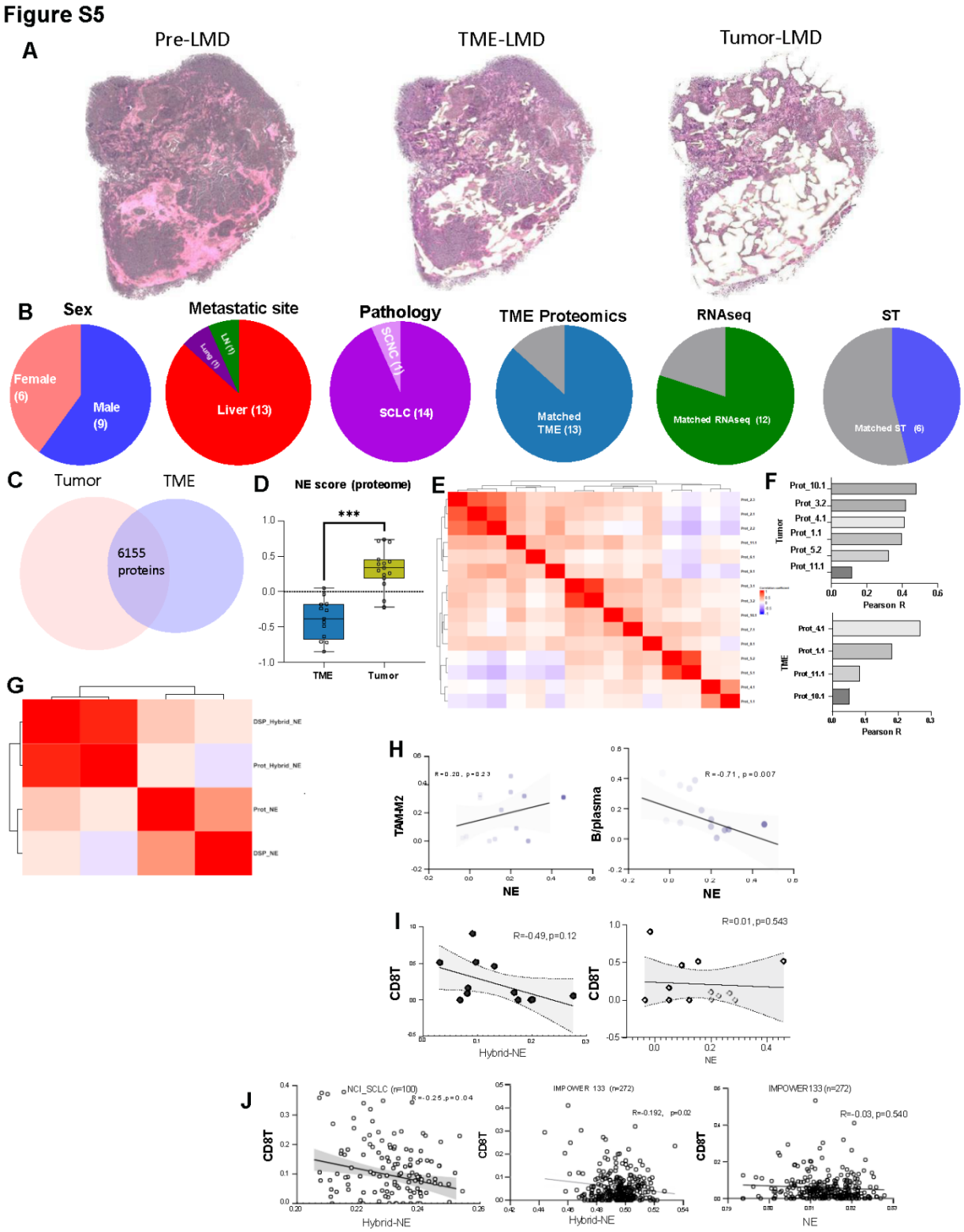


Figure S5
A
Pre-LMD
TME-LMD
Tumor-LMD
B
Sex
Female (6)
Male (9)
Metastatic site
Liver (13)
Pathology
SCLC (14)
TME Proteomics
Matched TME (13)
RNAseq
Matched RNAseq (12)
ST
Matched ST (6)
C
Tumor
TME
6155 proteins
D
NE score (proteome)
***
TME
Tumor
E
F
Prot_10.1
Prot_3.2
Prot_4.1
Prot_1.1
Prot_5.2
Prot_11.1
Tumor
Pearson R
TME
G
DSP_Hybrid_NE
Prot_Hybrid_NE
Prot_NE
DSP_NE
H
TAM-M2
B/plasma
NE
I
CD8T
Hybrid-NE
R=-0.49, p=0.12
R=0.01, p=0.543
J
NCI_SCLC (n=100)
IMPOWER 133 (n=272)
IMPOWER133 (n=272)
R=-0.03, p=0.540

Figure S5:

A) Low power (4X), H&E image of a representative tumor tissue (prot#7.1) subjected to LMD for proteomics processing with pre-LMD (left), post-TME LMD (middle) and post-Tumor LMD (right).

B) Clinical and experimental distribution of proteomics profiled tumors (n=15) from tumors collected during rapid autopsy.

C) Venn diagram showing proteomic capture landscape of tumor and TME proteins in our dataset with 6155 common proteins.

D) Proteomics derived NE signature score[7,9] between tumor and TME enriched regions. [#]

E) Pairwise correlation of 1000 proteins with the highest variance in SCLC tumor proteome.

F) Transcript to protein correlation data for tumors with both proteome and spatial transcriptomics data available (n=6 for tumor enriched regions, n=4 for TME enriched regions).

G) Tumor transcript-protein correlation for NE and hybrid NE subtypes. Pairwise correlation of matched tumor proteome and ST tumor segments derived NE and hybrid-NE signatures (n=6).

H) Correlation of tumor proteome NE signature (x-axis; ssGSEA) with TME proteome-derived TAM-M2 (left) and B/plasma cells (right) (CIBERSORT-derived proportions).

I) Correlation of tumor proteome NE (right) and Hybrid-NE (left) signature (ssGSEA) with TME proteome-derived CD8T signatures (CIBERSORT- derived).

J) Correlation of Hybrid-NE (left, middle) and NE signature (right) (ssGSEA) with CD8 T signatures (CIBERSORT-derived) in larger SCLC bulk-RNA sequencing dataset[9,17].

Abbreviations: LMD, laser capture microdissection; TME, tumor microenvironment; H&E, hematoxylin, and eosin; ssGSEA, single sample gene set enrichment analysis; ST, Spatial transcriptomics; DSP, digital spatial profiling- Spatial transcriptomics; CAF, cancer associated fibroblasts; ****statistical significance at $p<0.0001$, [#] student t-test.

Figure S6

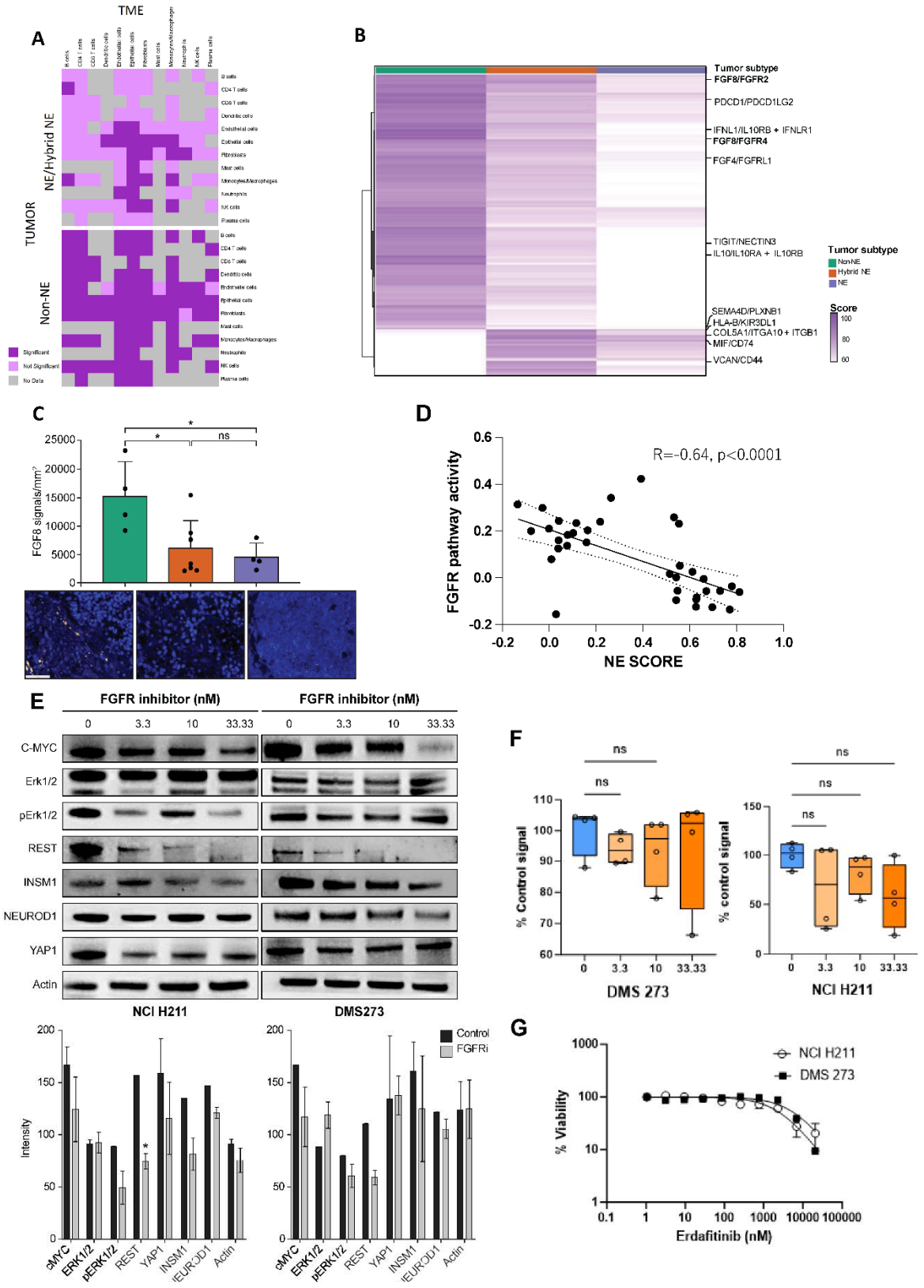


Figure S6:

A) Heatmap of global interactions between Tumor (receiver) and TME (sender) using an orthogonal cell-cell interaction approach (Cellphone DB)[65] confirming increased and varied significant interactions between tumor and TME regions in non-NE ecosystems compared to NE/Hybrid-NE ecosystems.

B) Heatmap of ligand-receptor interaction pairs using iCELLNET[64] showing most differentially enriched interactions (TME→ tumor). Clinically relevant and potentially targetable interactions are highlighted. FGF8 related interactions in bold.

C) *FGF8* RNA ISH signals across TME subtypes. Representative images (high power 40x, magnification) showing *FGF8* (yellow) in non-NE TME (left), hybrid-NE TME (middle) and NE TME (right)[#]. Nuclei are blue (DAPI). Scale bar set at 50μm.

D) *FGFR* activity scores (ssGSEA) in spatially resolved tumor segments transcriptomic (n=36) data.

E) FGF signaling intermediates, NE, and non-NE proteins in NCI-H211 and DMS-273 following treatment with erdafitinib at varying concentrations. (Below) Quantification of western blots of FGF signaling intermediates, NE, and non-NE proteins in NCI-H211 and DMS-273 after treatment with erdafitinib at varying concentrations.

F) Caspase-8 activation assay showing no significant increase in apoptosis at day 5 after erdafitinib treatment in DMS 273(left) and NCI H211 (right) SCLC cell lines.[&]

G) Cell titer glow viability assay showing no decrease in cell viability at erdafitinib concentrations used in this experiment.

Abbreviations: ssGSEA- single sample gene set enrichment analysis, FGF- fibroblast growth factor, * statistical significance at $p<0.05$; R= spearman correlation co-efficient; [#] student t-test; [&] Tukey's multiple comparison test

STAR Methods:

Lead Contact: Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Anish Thomas, anish.thomas@nih.gov.

Materials availability: This study did not generate new unique reagents.

Data and code availability: Processed transcriptomic and immunohistochemistry data from spatial transcriptomics profiled tumors and DMS273 cell line data using FGFR inhibitor (erdafitinib at 33.33nM concentration) and control (untreated cell line) with metadata file is available in the supplemental information. Whole genome sequencing data deposited in _____. The mass spectrometry proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[88] with the dataset identifier_____. This paper does not report any original code. Any additional information including microscopy data required to reanalyze the data reported in this paper is available from the lead contact upon request.

Ethics approval and patient consent: *NCT01851395* is approved by NIH Institutional ethics committee and Institutional review board (IRB). Detailed informed consents were obtained from all patients and their next of kin (after demise of the subject) as per protocol *NCT01851395.* Rapid autopsy was performed at NIH Clinical center laboratory of Pathology (LP) Autopsy suite by LP pathologists. Tissue was collected, stored and processed as per the detailed plan laid out in protocol NCT01851395.

Resource table:

REAGENT/RESOURCE	SOURCE	IDENTIFIER	SPECIFICATIONS (if any)
Antibodies: Immunofluorescence (mIF)			
HLADR	Abcam	Abcam ab7856; cr3-43	1:250
CD163	Abcam	Abcam ab74604; 10d6	1:500
CD115	Abcam	Abcam 183316; SP211	1:100
CD11b	Abcam	Abcam ab52478; EP1345Y	1:300
SMA	DAKO	Dako MO851; 1A4	1:500
FAP	Abcam	Abcam ab240989; SP325	1:500
C3C	LSBIO	LSbio LS-B7932	1:800
DAPI	Akoya bio	Akoya FP1490	1:1000
C-MYC	Santa Cruz biotech.	SC-40	1:1000
REST	Abcam	Ab211537	1:500
INSM1	Santa Cruz	Sc-271408	1:500
Antibodies: Immunohistochemistry (IHC)			

ASCL1	Santa Cruz	D7	1:1000
NEUROD1	Abcam	EPR17084	1:400
YAP1	Santa Cruz	63.7	1:1000
POU2F3	Novus biologics	Rabbit polyclonal	1:200
INSM1	Santa Cruz	A-8	1:1000
TEM 8	Abcam	EPNCI-R173-37	1:100
Experimental Models: Cell lines			
DMS 273	Millipore Sigma	95062830	NA
NCI H211	ATCC	CRL-5824	NA
Chemicals, peptides, and recombinant proteins			
Erdafitinib	Janssen Pharmaceuticals	JNJ-42756493	
Antibodies: Western blot			
C-MYC	Santa Cruz biotech.	SC-40	1:1000
INSM1	Santa Cruz biotech.	A-8	1:1000
YAP1	Proteintech	13584-1-AP	1:1000
NEUROD1	Cell Signaling tech	D90G12	1:1000
ERK	Cell Signaling tech	9102	1:1000
pERK (p-44/42 MAPK)	Cell Signaling tech	9101	1:1000
ACTIN	Millipore Sigma	A2066	1:3000
REST	Abcam	Ab211537	1:500
Software(s) and algorithms			
R	Open source	V4.2.1	
R-studio	Open source	Build 554 2022.07.01	
GraphPad prism	GraphPad Software	V9.3.1 (471)	
Qu-Path	Open source	V0.3.0	
Qlucore omics explorer	Qlucore AB	V3.8	
Halo	Indica labs	v3.4.15916.175	
STATA	Stata-Corp	V16.0	
Maftools5	Bioconductor	v. 2.12.0	
Complex heatmap	Bioconductor	v. 2.12.1	
Trimmomatic	CRAN	v.0.33	
BWA2	Bioconductor	v. 0.7.17	
Picard	Broad institute	v. 2.17.11	
GATK	Broad institute	v.4.2.2.0	

Mutect2	Broad institute	v4.1.0.0	
maftools	Bioconductor	v.2.12.0	
edgeR	Bioconductor	v.3.40.2	

Method details-

Tumor(s) selection for spatial transcriptomics tumor and TME profiling-

Samples were obtained from the first 10 consecutive patients who died of metastatic small cell lung cancer and were enrolled into the rapid autopsy protocol ***NCT01851395.*** Multiple sampled tumor sites from every patient were studied for the presence or absence of necrosis and/or autolytic changes and tumors with the least amount of these and preserved morphology were selected for spatial profiling. One tumor was selected per patient. Clinical details were recorded (table S1). Multiple 4–5-micron serial sections from each tumor FFPE block were taken. The first two sections were processed for spatial transcriptomics profiling (see below), immediately followed by separate unstained sections for mIF, IHC and mass spectrometry-based proteomics in that order. Following these sections, DNA extraction (below) for whole genome sequencing as well as bulk RNA sequencing was performed wherever feasible.

Tumor(s) selection for Mass spectrometry proteomics profiling

15 samples were obtained from 11 unique patients who died of metastatic small cell lung cancer and were enrolled into the rapid autopsy protocol ***NCT01851395.*** **Table S7 includes important clinical and site information regarding these samples. Three patients had >1 tumor sample profiled, and 6/15 tumor samples profiled had matched spatial transcriptomics data as profiled above albeit on a different serial section. Majority (12/15) of the tumors had matched bulk-RNAseq profiled and used to find gene-protein correlation.**

Spatial transcriptomics experiment and analysis

GeoMx^{R} Human Whole Transcriptome atlas (GeoMx Hu WTA) was used for the tumors selected for spatial transcriptomic analysis under a special early access program. Two consecutive sections of 4µm -thick slides were prepared from parent FFPE blocks. 1st slide was H&E stained to visualize the tumor and TME regions. Random but non-necrotic and non-hemorrhagic regions of interest (ROI) were selected by a board-certified pathologist. 2nd slide was deparaffinized, heated in ER2 solution (Leica) at 100 degrees C for 20 minutes, and treated with 1 microgram/ml of proteinase K (Ambion) at 37 degrees C for 15 minutes on a BOND Rxm autostainer (Leica). An overnight in situ hybridization was performed with a probe concentration of 4nM per probe as described previously[89]. Slides were washed twice at 37 degrees for 25 min with 50% formamide/2X SSC buffer to remove unbound probes. Prepared slides were stained further with pan-cytokeratin (AE1+AE3, Novus Biologicals,1:500) for epithelial/tumor cells, CD45 for immune cells (D9M8I, CST, 1:100), Syto83 for nucleus (stainSyto83, Thermofisher, 1:25). Stained slides

were loaded onto GeoMx instrument and scanned. Forty (40) circular ROIs of 500µm diameter were selected randomly across different areas of 10 tumor slides. Using the information from pan-CK and morphological features, tumor cells/areas were selected and marked as “pan-CK positive/Tumor” areas, and the rest of the area was marked as “pan-CK negative/TME” areas to perform segmentation of most of the ROIs(n=32/40) where the clear distinction of “TME” and “Tumor” segments was possible. In addition, 4 regions with no/minimal visible TME were profiled as “tumor only” segments and other 4 regions without any tumor in near distance were profiled as “normal” segments.

After sequencing, reads were trimmed, merged, and aligned to retrieve the probe identity. The unique molecular identifier region of each read was used to remove PCR duplicates and duplicate reads, thus converting reads into digital counts. The sequencing saturation was sufficient for RNA and was >80% for all the ROIs. For each gene in each sample, the reported count value is the mean of the individual probe counts after the removal of outlier probes. The limit of quantification (LOQ) was set at the geometric mean plus two standard deviations of the negative probes. 18676 genes were targeted and captured by DSP. Logarithmic (base 2) conversions of normalized third Quantile expression values were used for downstream analysis. The Euclidean distance metric was used for principal component analysis (PCA).

Whole genome sequencing (WGS) and Bulk RNA sequencing

DNA and RNA were extracted from FFPE tumor tissues used for spatial and proteomics profiling for WGS and bulk RNA-sequencing using standardized DNA extraction kits. Matched normal DNA was extracted from stored blood. Samples were pooled and sequenced on Novaseq 6000 using S4 flow cell configuration using Truseq Nano DNA library prep and 150bp paired end sequencing was done. All the samples have percent of Q30 bases above 88%. All the samples have yields between 193 and 389 million pass filter reads. Human DNA libraries were sequenced with the aim of obtaining coverage of a minimum 70X for tumor DNA and 30xmatched normal DNA.

QC and Alignment -For all whole genome and whole exome data, raw FASTQ reads were trimmed for adapter and low-quality control using Trimmomatic1 (v. 0.33) prior to alignment[90]. Alignment was performed using BWA2 (v. 0.7.17) mapping to the human reference hg38 genome. Duplicated reads were marked using Picard (v. 2.17.11) followed by indel realignment and base quality score recalibration using GATK (v. 4.2.2.0)[91].

Variant Calling and Annotation-Somatic variant calling was performed using Mutect2, in both tumor-normal and tumor-only mode using the GATK best practices (v. 4.2.2.0). Variants were annotated using the Ensembl Variant Effect Predictor and the vcf2maf tool (v. 102)[92]. MAF files were used for all downstream annotation and visualization. Variants were filtered removing variants with tumor read depths <5, an alternate allele count <2, and normal alternate allele counts >1. Additionally, to filter commons polymorphisms, variants were removed with a frequency >0.001 in the ExAC, gnomAD, or 1000 Genomes population databases.

Copy Number Variant Analysis-For all samples, cn.MOPS4 was used for CNV calling. Aligned and processed BAM files were converted to read count matrices used as input. The software models the depths of coverage across samples at each genomic position to account for read count biases along chromosomes. For WES samples, the exomecn.mops function to adjust for the varying window spans across the target regions.

RNA was extracted from FFPE tumor cores (n= 12 samples) using RNeasy FFPE kits according to the manufacturer's protocol (QIAGEN, Germantown, MD). RNA-seq libraries were generated using Truseq RNA Access Library Prep Kits (TruSeq RNA Exome kits; Illumina) and sequenced on NextSeq500 sequencers using 75□bp paired-end sequencing method (Illumina, San Diego, CA). For transcriptomic analyses, raw RNA-Seq count data were normalized for inter-gene/sample comparison using FPKM, followed by log2(x□+□1) transformation, as implemented in the edgeR R/Bioconductor package[93]

Multiplex Immunofluorescence (mIF)-

mIF was performed through our collaboration with the Human Immune Monitoring Shared Resource (HIMSR) at the University of Colorado School of Medicine. We performed 6 colors multispectral imaging using the Akoya Biosciences Vectra Polaris instrument. Unstained FFPE-derived slides were on the Leica Bond RX autostainer according to standard protocols provided by Leica and Akoya Biosciences and performed routinely by the HIMSR. Slides from spatially profiled tumors were stained consecutively with antibodies specific for the following proteins: antibodies (see above table) and DAPI counterstain. Briefly, the slides were deparaffinized, heat treated in antigen retrieval buffer, blocked, and incubated with primary antibody, followed by horseradish peroxidase (HRP)-conjugated secondary antibody polymer, and HRP-reactive OPAL fluorescent reagents that use TSA chemistry to deposit dyes on the tissue immediately surrounding each HRP molecule. To prevent further deposition of fluorescent dyes in subsequent staining steps, the slides were stripped in between each stain with heat treatment in antigen retrieval buffer. Whole slide motif scans were collected using the 20x objective. The 6 color images were analyzed with inForm software to unmix adjacent fluorochromes, subtract autofluorescence, segment the tissue, compare the frequency and location of cells in the tumor and stromal areas, to segment cellular membrane, cytoplasm, and nuclear regions, score each cellular compartment for expression of any scoring markers you include in your panel, and phenotype infiltrating immune cells according to morphology and cell marker expression. Segments that were profiled using GeoMx spatial transcriptomics and retained the tumor and TME morphology on the sublevel section (as deemed by board certified pathologist) used for mIF were marked and quantified for cells/density parameter.

Immunohistochemistry-

All stains were run on our Leica Bond Max auto-stainer with standard DAB protocol. IHC stains (antibodies listed in the resource table) were performed at the National Institutes of Health (NIH), Laboratory of Pathology and at Molecular histopathology laboratory, Frederick national laboratory, NCI

according to the manufacturer's instruction. IHC-stained slides were scanned using the Carl Zeiss AxioScan.Z1 microscope equipped with a Plan-Apochromat 20x NA 0.8 objective. For Nuclear markers-H-score was calculated based on the equation: 1 x (% of weakly stained nuclei) + 2 x (% of moderately stained nuclei) + 3 x (% of strongly stained nuclei). For Cytoplasmic and Membranous markers % positive cells out of total cells in a given segment (based on nuclei scoring) were used to report. Segments that were profiled using GeoMx spatial transcriptomics and largely retained the tumor and TME morphology on the section (as deemed by board certified pathologist) used for IHC were marked and quantified for cells/density parameter. Nuclear and cytoplasmic IHC scoring was done using Qu-Path (open source v0.3.0) software.

Mass-spectrometry based quantitative proteomics-

Consecutive tissue thin sections (8 µm) were generated by microtome, placed onto polyethylene naphthalate (PEN) membrane slides, and stained using hematoxylin and eosin (H&E). Cellular enrichment via laser microdissection (LMD) was performed on the LMD7 (Leica Microsystems) for selective harvest of tumor and TME regions from each of the 15 tissue specimens. LMD enriched samples in 20 µl 100 mM triethylammonium bicarbonate / 10% acetonitrile were pressure-assisted, trypsin-digested using pressure cycling technology (PCT; Pressure Biosciences, Inc.), as previously described[62]. Briefly, LMD samples were incubated at 99 °C for 30 min followed by 50 °C for 10 min with SMART trypsin (ThermoFisher Scientific). SMART trypsin was added at a ratio of 1 µg per 30 mm^2 tissue. Lysis and digestion were performed in a Barocycler 2320EXT (Pressure Biosciences, Inc.) by cycling between 45 kpsi for 50 sec and atmospheric pressure for 10 sec, for 60 cycles at 50 °C. Peptide digests were quantified by colorimetric assay (Pierce BCA Protein Assay Kit). Digested peptides (10 µg) from each sample (n=15 LMD enriched tumor and n=13 LMD enriched TME samples) were labeled using isobaric tandem mass tags (TMT), per manufacturer's instructions (TMTpro 18-plex Isobaric Label Reagent Set, ThermoFisher Scientific)[94]. A reference sample was generated by pooling equivalent amounts of peptide digests from each of the patient samples and included in each TMTpro-18 multiplex. TMT-labeled samples were pooled within respective multiplexes, cleaned using the EasyPep Maxi MS Sample Prep Kit (ThermoFisher Scientific) and fractionated offline to 36 pooled fractions via basic reversed-phase liquid chromatography (bRPLC). The TMTpro-18 bRPLC fractions were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) employing a nanoflow LC system (EASY-nLC 1200, ThermoFisher Scientific) coupled online with a Q Exactive HF-X MS (ThermoFisher Scientific), as previously described[62].

RNA Insitu-Hybridization

FGF8 expression was detected by staining 5 um FFPE tissue sections with RNAscope® 2.5 LS Probe Hs-FGF8-C2 (ACD, Cat# 415798-C2) and the RNAscope LS Multiplex Fluorescent Assay (ACD, Cat# 322800) using the Bond RX auto-stainer (Leica Biosystems) with a tissue pretreatment of 15 minutes at

95°C with Bond Epitope Retrieval Solution 2 (Leica Biosystems), 15 minutes of Protease III (ACD) at 40°C, and 1:750 dilution of TSA-Cyanine 3 Plus (AKOYA). The RNAscope® 3-plex LS Multiplex Negative Control Probe (Bacillus subtilis dihydrodipicolinate reductase (*dapB*) gene in channels C1, C2, and C3, Cat# 320878) was used as a negative control. The RNAscope® LS 2.5 3-plex Positive Control Probe-Hs was used as a technical control to ensure the RNA quality of tissue sections was suitable for staining. Slides were digitally imaged using an Aperio ScanScope FL Scanner (Leica Biosystems). All image analysis was performed using HALO imaging analysis software (v3.4.2986.246; Indica Labs, Corrales, NM), and image annotations were performed by one pathologist (BK). The analysis was performed using FISH V3.1.3 in HALO to determine percent positive FGF8 positive cells. Areas of artifact such as folds, and tears were excluded from analysis.

Cell-lines treatment and analysis:

Briefly, established SCLC cell lines (DMS 273 and NCI H211) were grown in RPMI-1640 media supplemented with 10% FBS, Pen-Strep and L-Glutamine at 37˚C and 5% CO2. Cell aliquots were treated with either DMSO or FGFR inhibitor (Erdafitinib) dissolved in DMSO dosed at 3.3, 10 and 33.33uM concentrations once at the beginning of a 5-day interval period. At the end of 5 days, cells were subsequently analyzed for morphological analysis, RNA extraction using standard pipelines for RNA sequencing, and cell lysates preparation for western blot staining. RNA extraction was performed for cell lines exposed to DMSO and at 33.33uM FGFRi concentration (4 technical replicates each). Western blotting was performed for cell lines exposed to FGFR inhibitor at 0 (control), 3.3, 10 and 33.33nM concentrations and repeated twice. DMS 273 and NCI-H211 cells were plated at 1,000 cells per well in 384 well plates and treated with erdafitinib at the concentrations indicated. After 5 days cells were collected per manufacturer's instructions using either CellTiterGlo® (Promega, G7570) to measure viability by assessing ATP concentrations, or using Caspase-Glo® 8 Assay Systems® (Promega, G8200) to measure cleaved Caspase 8 activity.

<u>Bio-statistical and Bio-informatics analyses-</u>

Spatial transcriptomics analysis- Log(10) normalized data was used for subsequent analysis. Two TME segments that did not meet the sequencing quality metrics (Fig. S1C, segment 1007 and 1008) were excluded from further analyses.

Optimal Cluster determination in spatial transcriptomics dataset (tumors)-

In order to determine ideal number of clusters amongst the 36 tumor segments profiled we calculated "total within-cluster sum of squares" values for projected cluster numbers (k) from 1-10 using the *nbClust* R package[95]. Based on this, we found that partitioning the data into 3 (k=3) clusters would be the best approach in order to avoid over-clustering, as determined by the "elbow" of the curve. Subsequently using k=3, we performed k-means unsupervised partitioning clustering[96] (Iterations=50, Attempts=200) in order to annotate each tumor segment to a cluster.

Cluster-defining genes

For the tumor segments (n=36), 6000 highly variable genes were extracted Given that 3 well-defined clusters were visualized, PC1 and PC2 contribution in terms of coordinates axis of each gene was computed. positive contributors to PC1 were classified as cluster 1-specific signature (351 genes), the negative contributors to PC2 as the cluster 3-specific signature (149 genes), and the positive contributors to PC2 as the cluster 2-specific signature (244 genes).

Quantitative proteomics data processing pipeline-

Peptide identifications, normalization, and log_2-transformation for generation of protein-level quantitative data were performed, as previously described[62]. Briefly, raw data files were searched using Mascot (Matrix Science) and Proteome Discoverer (ThermoFisher Scientific, Inc). with a publicly available non-redundant human proteome database (Swiss-Prot, Homo sapiens, Proteome UP000005640, 20,257 sequences, downloaded 12-01-2017; http://www.uniprot.org/) appended with porcine trypsin (Uniprot: P00761) sequences. Peptide spectral match (PSM) results were filtered using a <1.0% false discovery rate (FDR). Log_2-transformed PSM abundances were calculated using z-score transformation for each TMT-18 channel relative to the pooled reference standard. Protein-level abundances were calculated from the normalized log_2-transformed TMT reporter ion ratio abundances from ≥2 PSMs corresponding to a single protein accession. Protein-level abundances were imputed for proteins absent from individual patient samples, but present in ≥50% of all patient samples.

Analysis platforms and statistical methods-

R studio version 4.2.1 (2021-11-01) (R Foundation for Statistical Computing), GraphPad Prism version 9.0.2 (GraphPad Software) and Qlucore Omics Explorer version 3.7 (Qlucore AB) were used to generate figures and perform statistical analyses. Partekflow version 10.0.22.0321 (Partek®) was used for processed single cell sequencing analysis and visualization. All tests were two-tailed and p-values less than 0.05 were considered significant. False discovery rate (FDR) of 5% was used wherever applicable. Correlation analyses was done using Pearson's correlation to obtain correlation co-efficient (r). Student's t-test was used to compare between two groups and one-way Anova test followed by Tukey's multiple comparison test was performed to compare numerical data between more than 2 groups unless otherwise specified where false discovery rate (set at q<0.05) using Benjamini and Hochberg correction method was used. Spearman correlation co-efficient was used when sample size was used as correlation constant.

R studio version 4.1.2 (2021-11-01) (R Foundation for Statistical Computing), GraphPad Prism version 9.0.2 (GraphPad Software) and Qlucore Omics Explorer version 3.7 (Qlucore AB) were used to generate figures and perform statistical analyses. Partekflow version 10.0.22.0321 (Partek®) was used for processed single cell sequencing analysis and visualization. All tests were two-tailed and p-values less than 0.05 were considered significant. False discovery rate (FDR) of 5% was used wherever applicable.

Student's t-test was used to compare between two groups and one- way Anova test followed by Tukey's multiple comparison test was performed to compare numerical data between more than 2 groups unless otherwise specified.

Survival analysis – Ecotyper cell state scores were computed for bulk RNA sequencing data (George et al). For Ecotyper_CAF_S03 scores, an arbitrary cut-off of 75% percentile was used to divide samples into "low CAF_S03 (<75 percentile)" and "high CAF_S03 (>75 percentile). Survival analysis was further done using the cox proportional hazards model and log-rank testing was done to calculate the survival distributions of the two types of samples. Statistical significance was noted at value <p. Multivariate analysis was done after factoring in other available and potentially confounding variables (age, sex, smoking status).

Tumor budding/nesting assignment.

All the tumor segments profiled (n=36) were analyzed for cyto morphological features on the initial Hematoxylin and Eosin stain. Tumor nesting was identified as feature where multiple tumor cell clusters were found to be surrounded by dense stromal tissue as previously reported[47]. Tumor budding was identified when tumor cells either single cells or small clusters (<5-10 tumor cells) were identified in a segment.

Pairwise correlation analysis

Normalized cell line and human bulk RNA data were downloaded and subsequently, ssGSEA based scores were computed for relevant gene signatures for each sample. Correlation metric pairwise correlation was done either for ssGSEA-derived gene sets or single normalized genes. Subsequently, a heatmap plot was constructed using Pearson's correlation coefficients, which was then hierarchically clustered to observe the correlation patterns between genes/gene sets of interest.

Tumor purity estimates:

For prior datasets[5,9] where matched bulk RNA-seq and Whole genome/exome sequencing data was available tumor purity scores were estimated using ABSOLUTE[38] algorithm. Subsequently ssGSEA derived stromal and immune scores[36] were computed for these transcriptomes. Using these two scores a linear model was generated. Subsequently, stromal and immune scores were calculated for tumor segments of spatial transcriptomics profiled segments and tumor purity estimates were predicted using the above model.

Single sample gene set enrichment analysis (ssGSEA)

ssGSEA enrichment scores were computed using GSVA R/Bioconductor package[97]. Additionally, data generated for Figure 2C was calculated using ssGSEA scores for each tumor segment for the 50 hallmark gene sets downloaded from MsigDB. Subsequently, for each hallmark set, q value was

calculated by comparing one cluster segments vs other two segments using two-way Anova followed by controlling for FDR method of Benjamini and Hochberg.

Gene set enrichment analysis (GSEA)

For each cluster, the marker genes were identified by comparing the cluster with the other clusters using the Wilcoxon rank-sum test. Then, the fold change in expression of each gene was computed by taking the taking the log2(median Q3 expression / median Q3 expression in other clusters). Genes lists were constructed using those genes with p-values < 0.05 and an absolute fold change of >2. Enriched hallmarks from the Molecular Signatures Database were identified by pre-ranked GSEA using clusterProfiler v.4.0.5 using the gene list ranked by log-transformed p-values with signs set to positive/negative for a fold change of >1 or <1, respectively.

Transcription factors downstream target activity

An SCLC context-specific gene regulatory network was computed using the ARACNe-AP software[98]. Its associated transcription factor (TF) target sets (regulons) were then input to the VIPER software package[99], to infer a TF x sample matrix of activity values. Each TF activity value is a normalized measure of regulon enrichment among sample-specific over/under expressed genes. ARACNe-AP was run using a bulk RNA-Seq data set derived from 127 SCLC tumor samples[9] excluding genes with zero expression across all samples. The regulatory network was restricted to TFs and TF co-regulators associated with the Gene Ontology terms GO:0003700 (DNA-binding transcription factor activity) and GO:0003712 (transcription coregulator activity), together with their inferred targets (using the ARACNe-AP '--tfs' parameter to specify the starting gene regulator set). The final network was derived by integrating 100 intermediate networks, constructed using bootstraps of the expression data set, as described[98].The p-value for mutual information significance, a parameter governing the number of associations in the network, was set to 10^-8.The VIPER activity matrix was computed using the viper R/Bioconductor package (https://www.bioconductor.org/packages/release/bioc/html/viper.html). In particular, the viper::viper::aracne2regulon () function was used to derive a regulon object from the ARACNe-AP output and gene expression data, using the parameter setting "format='3col'". Activity data were then computed using latter regulon and the expression data as input to the viper::viper () function, with parameter settings 'pleiotropy = TRUE' and 'nes = TRUE'.

TME cluster assignment

Each of the TME segment was assigned the cluster class of the tumor segment from which it was segmented (spatially proximate TME). Because TME segments of 2 regions (1007 and 1008) yielded low quality data, they were excluded from further analyses. Total of 30 TME segments were thus analyzed downstream.

CIBERSORT deconvolution

CIBERSORT tool developed by Newman et al[53] to quantify cell types in bulk derived RNA expression data was used to predict cell types in TME segments. The analysis was run on the CIBERSORT website at http://cibersort.stanford.edu. We applied safeTME[100] to the TME segments gene expression data. For each run, 100 permutations were performed, and quantile normalization was disabled.

Cellular ecotypes and cell states computation and evaluation

ECOTYPER is a recent tool developed by Luca et al[52] to deconvolve tumor-microenvironment cell types and interactions between them as "multi-cellular ecotypes". We ran the analysis on ECOTYPER website https://ecotyper.stanford.edu/carcinoma/ for TME segments. Both ecotypes and different cell states for each cell-type information was downloaded from the analyzed results and for each given TME segment a given Ecotype and a cell state with maximum enrichment was binned to that category. For the pan cancer tumors analysis, Ecotyper data for TCGA cell state assignments were downloaded from https://ecotyper.stanford.edu/carcinoma/ website. SCN scores of the TCGA tumors were obtained from publicly available data available with Balanis et al[34].

Cell-cell communication /ligand-receptor interaction

We used ICELLNET[64] to score sender-receiver (either TME segment->Tumor segment or Tumor segment->TME segment) expression profiles using a custom set of ligand-receptor (L/R) annotations from ECOTYPER[52]. Next, we computed a summary communication score for each sender-receiver cell population (as available in ECOTYPER) by taking the maximum ICELLNET score of all L/R pairs for each sender-receiver cell population.

To confirm, findings from ICELLNET we also computed similar scores using an orthogonal pipeline CellphoneDB[65]. Briefly, Cellphone was run with the "cpdb_statistical_analysis_method". Ligand-receptor interaction means, and p-values were generated, and ligand-receptor pairs were cross matched with annotations from ECOTYPER[52] converting gene-gene interaction to cell-cell interaction. Lowest ligand/receptor interaction values were assigned to each cell/cell interaction after which a mean score was computed and plotted as a heatmap.

Shannon diversity scores

Shannon biodiversity index measures the diversity of a population distribution. Higher biodiversity could indicate the presence of a higher number of species, or homogeneity in the distribution of a smaller number of species. We define the Shannon index as:

$$H' = -\Sigma pi * ln(pi) Ni = 1$$

Here, N is the total number of species and pi is the proportion of i'th species.

Additional Resources:

Rapid autopsy protocol link: URL https://clinicaltrials.gov/ct2/show/NCT01851395

Natural history protocol link: URL https://clinicaltrials.gov/ct2/show/NCT02146170

Funding:

This project has been funded in part with Federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health, under contract HHSN2612008000031 (BK, AW, DB). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613-628 (2017).
2. Gupta, P.B., Pastushenko, I., Skibinski, A., Blanpain, C. & Kuperwasser, C. Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. *Cell Stem Cell* **24**, 65-78 (2019).
3. Boumahdi, S. & de Sauvage, F.J. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat Rev Drug Discov* **19**, 39-56 (2020).
4. Rudin, C.M., Brambilla, E., Faivre-Finn, C. & Sage, J. Small-cell lung cancer. *Nat Rev Dis Primers* **7**, 3 (2021).
5. George, J. *et al.* Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53 (2015).
6. Gazdar, A.F., Carney, D.N., Nau, M.M. & Minna, J.D. Characterization of Variant Subclasses of Cell-Lines Derived from Small Cell Lung-Cancer Having Distinctive Biochemical, Morphological, and Growth-Properties. *Cancer Research* **45**, 2924-2930 (1985).
7. Zhang, W. *et al.* Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. *Translational Lung Cancer Research* **7**, 32-34 (2018).
8. Calbo, J. *et al.* A Functional Role for Tumor Cell Heterogeneity in a Mouse Model of Small Cell Lung Cancer. *Cancer Cell* **19**, 244-256 (2011).
9. Lissa, D. *et al.* Heterogeneity of neuroendocrine transcriptional states in metastatic small cell lung cancers and patient-derived models. *Nat Commun* **13**, 2023 (2022).
10. Thomas, A. *et al.* Therapeutic targeting of ATR yields durable regressions in small cell lung cancers with high replication stress. *Cancer Cell* **39**, 566-579 e7 (2021).
11. Roper, N. *et al.* Notch signaling and efficacy of PD-1/PD-L1 blockade in relapsed small cell lung cancer. *Nature Communications* **12**(2021).
12. Stewart, C.A. *et al.* Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nat Cancer* **1**, 423-436 (2020).
13. McColl, K. *et al.* Reciprocal expression of INSM1 and YAP1 defines subgroups in small cell lung cancer. *Oncotarget* **8**, 73745-73756 (2017).
14. Qu, S. *et al.* Molecular Subtypes of Primary SCLC Tumors and Their Associations With Neuroendocrine and Therapeutic Markers. *Journal of Thoracic Oncology* **17**, 141-153 (2022).

15. Owonikoko, T.K. *et al.* YAP1 Expression in SCLC Defines a Distinct Subtype With T-cell-Inflamed Phenotype. *J Thorac Oncol* **16**, 464-476 (2021).
16. Roper, N. *et al.* Notch signaling and efficacy of PD-1/PD-L1 blockade in relapsed small cell lung cancer. *Nat Commun* **12**, 3880 (2021).
17. Gay, C.M. *et al.* Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell* **39**, 346-+ (2021).
18. Mahadevan, N.R. *et al.* Intrinsic Immunogenicity of Small Cell Lung Carcinoma Revealed by Its Cellular Plasticity. *Cancer Discov* **11**, 1952-1969 (2021).
19. Wagner, A.H. *et al.* Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. *Nature Communications* **9**(2018).
20. Sutherland, K.D., Ireland, A.S. & Oliver, T.G. Killing SCLC: insights into how to target a shapeshifting tumor. *Genes Dev* **36**, 241-258 (2022).
21. Raghavan, S. *et al.* Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* **184**, 6119-6137 e26 (2021).
22. Yang, D. *et al.* Intertumoral Heterogeneity in SCLC Is Influenced by the Cell Type of Origin. *Cancer Discovery* **8**, 1316-1331 (2018).
23. Ferone, G., Lee, M.C., Sage, J. & Berns, A. Cells of origin of lung cancers: lessons from mouse studies. *Genes Dev* **34**, 1017-1032 (2020).
24. Ireland, A.S. *et al.* MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. *Cancer Cell* **38**, 60-78 e12 (2020).
25. Lim, J.S. *et al.* Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature* **545**, 360-364 (2017).
26. Pongor, L.S. *et al.* Extrachromosomal DNA Amplification Contributes to Small Cell Lung Cancer Heterogeneity and is Associated with Worse Outcomes. *Cancer Discov* (2023).
27. Mollaoglu, G. *et al.* MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. *Cancer Cell* **31**, 270-285 (2017).
28. Thomas, A., Desai, P. & Takahashi, N. Translational research: A patient-centered approach to bridge the valley of death. *Cancer Cell* **40**, 565-568 (2022).
29. Bohan, P.K.M. *et al.* Correlation of tumor microenvironment from biopsy and resection specimens in untreated colorectal cancer patients: a surprising lack of agreement. *Cancer Immunology Immunotherapy* **70**, 1465-1474 (2021).
30. Thomas, A., Pattanayak, P., Szabo, E. & Pinsky, P. Characteristics and Outcomes of Small Cell Lung Cancer Detected by CT Screening. *Chest* **154**, 1284-1290 (2018).
31. Rudin, C.M. *et al.* Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat Rev Cancer* **19**, 289-297 (2019).
32. Megyesfalvi, Z. *et al.* The landscape of small cell lung cancer metastases: Organ specificity and timing. *Thorac Cancer* **12**, 914-923 (2021).
33. Brady, L. *et al.* Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling. *Nat Commun* **12**, 1426 (2021).
34. Balanis, N.G. *et al.* Pan-cancer Convergence to a Small-Cell Neuroendocrine Phenotype that Shares Susceptibilities with Hematological Malignancies. *Cancer Cell* **36**, 17-34 e7 (2019).
35. Schaefer, C.F. *et al.* PID: the Pathway Interaction Database. *Nucleic Acids Res* **37**, D674-9 (2009).
36. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* **4**, 2612 (2013).
37. Zhang, W. *et al.* Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. *Transl Lung Cancer Res* **7**, 32-49 (2018).

38. Carter, S.L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* **30**, 413-21 (2012).
39. Gavish, A. *et al.* Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. *Nature* **618**, 598-606 (2023).
40. Gavish, A. *et al.* The transcriptional hallmarks of intra-tumor heterogeneity across a thousand tumors. *bioRxiv*, 2021.12.19.473368 (2021).
41. Groves, S.M. *et al.* Cancer Hallmarks Define a Continuum of Plastic Cell States between Small Cell Lung Cancer Archetypes. *bioRxiv*, 2021.01.22.427865 (2021).
42. Jahchan, N.S. *et al.* Identification and Targeting of Long-Term Tumor-Propagating Cells in Small Cell Lung Cancer. *Cell Rep* **16**, 644-56 (2016).
43. Takahashi, N. *et al.* Replication stress defines distinct molecular subtypes across cancers. *Cancer Res Commun* **2**, 503-517 (2022).
44. Cai, L. *et al.* Cell-autonomous immune gene expression is repressed in pulmonary neuroendocrine cells and small cell lung cancer. *Commun Biol* **4**, 314 (2021).
45. Mukhopadhyay, S., Dermawan, J.K., Lanigan, C.P. & Farver, C.F. Insulinoma-associated protein 1 (INSM1) is a sensitive and highly specific marker of neuroendocrine differentiation in primary lung neoplasms: an immunohistochemical study of 345 cases, including 292 whole-tissue sections. *Mod Pathol* **32**, 100-109 (2019).
46. Wang, Y. *et al.* Comprehensive Molecular Characterization of the Hippo Signaling Pathway in Cancer. *Cell Rep* **25**, 1304-1317 e5 (2018).
47. Boxberg, M. *et al.* Tumor Budding and Cell Nest Size Are Highly Prognostic in Laryngeal and Hypopharyngeal Squamous Cell Carcinoma: Further Evidence for a Unified Histopathologic Grading System for Squamous Cell Carcinomas of the Upper Aerodigestive Tract. *Am J Surg Pathol* **43**, 303-313 (2019).
48. Alcala, N. *et al.* Integrative and comparative genomic analyses identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids. *Nat Commun* **10**, 3407 (2019).
49. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).
50. Chan, J.M. *et al.* Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell* **39**, 1479-1496 e18 (2021).
51. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845-865 e7 (2021).
52. Luca, B.A. *et al.* Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell* **184**, 5482-5496 e28 (2021).
53. Newman, A.M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* **12**, 453-7 (2015).
54. Shannon, C.E. A mathematical theory of communication. *THe Bell System Technical Journal* **27**, 379-423 (1948).
55. Ma, R.Y., Black, A. & Qian, B.Z. Macrophage diversity in cancer revisited in the era of single-cell omics. *Trends Immunol* **43**, 546-563 (2022).
56. Buchmann, K. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Front Immunol* **5**, 459 (2014).
57. Galbo, P.M., Jr., Zang, X. & Zheng, D. Molecular Features of Cancer-associated Fibroblast Subtypes and their Implication on Cancer Pathogenesis, Prognosis, and Immunotherapy Resistance. *Clin Cancer Res* **27**, 2636-2647 (2021).
58. Qian, J. *et al.* A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. *Cell Res* **30**, 745-762 (2020).
59. Yu, B. *et al.* Periostin secreted by cancer-associated fibroblasts promotes cancer stemness in head and neck cancer by activating protein tyrosine kinase 7. *Cell Death Dis* **9**, 1082 (2018).

60. Hoye, A.M. *et al.* Tumor endothelial marker 8 promotes cancer progression and metastasis. *Oncotarget* **9**, 30173-30188 (2018).
61. Miles, L.A. *et al.* Anthrax toxin receptor 1 is the cellular receptor for Seneca Valley virus. *Journal of Clinical Investigation* **127**, 2957-2967 (2017).
62. Lee, S. *et al.* Molecular Analysis of Clinically Defined Subsets of High-Grade Serous Ovarian Cancer. *Cell Rep* **31**, 107502 (2020).
63. Buccitelli, C. & Selbach, M. mRNAs, proteins and the emerging principles of gene expression control. *Nat Rev Genet* **21**, 630-644 (2020).
64. Noel, F. *et al.* Dissection of intercellular communication using the transcriptome-based framework ICELLNET. *Nat Commun* **12**, 1089 (2021).
65. Efremova, M., Vento-Tormo, M., Teichmann, S.A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506 (2020).
66. Figueiredo, C.R. *et al.* Blockade of MIF-CD74 Signalling on Macrophages and Dendritic Cells Restores the Antitumour Immune Response Against Metastatic Melanoma. *Front Immunol* **9**, 1132 (2018).
67. Barkal, A.A. *et al.* CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy. *Nature* **572**, 392-396 (2019).
68. Matozaki, T., Murata, Y., Okazawa, H. & Ohnishi, H. Functions and molecular mechanisms of the CD47-SIRPalpha signalling pathway. *Trends Cell Biol* **19**, 72-80 (2009).
69. Xie, Y. *et al.* FGF/FGFR signaling in health and disease. *Signal Transduct Target Ther* **5**, 181 (2020).
70. Zheng, X. *et al.* Proteomic Analysis for the Assessment of Different Lots of Fetal Bovine Serum as a Raw Material for Cell Culture. Part IV. Application of Proteomics to the Manufacture of Biological Drugs. *Biotechnology Progress* **22**, 1294-1300 (2008).
71. Caeser, R. *et al.* MAPK pathway activation selectively inhibits ASCL1-driven small cell lung cancer. *iScience* **24**, 103224 (2021).
72. Morrison, S.J. *et al.* Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells. *Cell* **101**, 499-510 (2000).
73. Chen, Y., McAndrews, K.M. & Kalluri, R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol* **18**, 792-804 (2021).
74. Calvo, F. *et al.* Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nature Cell Biology* **15**, 637-+ (2013).
75. Cordenonsi, M. *et al.* The Hippo Transducer TAZ Confers Cancer Stem Cell-Related Traits on Breast Cancer Cells. *Cell* **147**, 759-772 (2011).
76. Melero, I. *et al.* A first-in-human study of the fibroblast activation protein-targeted, 4-1BB agonist RO7122290 in patients with advanced solid tumors. *Sci Transl Med* **15**, eabp9229 (2023).
77. Horn, L. *et al.* First-Line Atezolizumab plus Chemotherapy in Extensive-Stage Small-Cell Lung Cancer. *N Engl J Med* **379**, 2220-2229 (2018).
78. Paz-Ares, L. *et al.* Durvalumab plus platinum-etoposide versus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): a randomised, controlled, open-label, phase 3 trial. *Lancet* **394**, 1929-1939 (2019).
79. Ozdemir, B.C. *et al.* Depletion of Carcinoma-Associated Fibroblasts and Fibrosis Induces Immunosuppression and Accelerates Pancreas Cancer with Reduced Survival. *Cancer Cell* **28**, 831-833 (2015).
80. Rhim, A.D. *et al.* Stromal Elements Act to Restrain, Rather Than Support, Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **25**, 735-747 (2014).

81. Kwon, M.C. *et al.* Paracrine signaling between tumor subclones of mouse SCLC: a critical role of ETS transcription factor Pea3 in facilitating metastasis. *Genes & Development* **29**, 1587-1592 (2015).
82. Shia, D.W. *et al.* Targeting PEA3 transcription factors to mitigate small cell lung cancer progression. *Oncogene* **42**, 434-448 (2023).
83. Ferone, G. *et al.* FGFR1 Oncogenic Activation Reveals an Alternative Cell of Origin of SCLC in Rb1/p53 Mice. *Cell Reports* **30**, 3837-+ (2020).
84. Chen, R. *et al.* FGFRL1 affects chemoresistance of small-cell lung cancer by modulating the PI3K/Akt pathway via ENO1. *Journal of Cellular and Molecular Medicine* **24**, 2123-2134 (2020).
85. Pardo, O.E. *et al.* The Fibroblast Growth Factor Receptor Inhibitor PD173074 Blocks Small Cell Lung Cancer Growth In vitro and In vivo. *Cancer Research* **69**, 8645-8651 (2009).
86. Bluemn, E.G. *et al.* Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. *Cancer Cell* **32**, 474-+ (2017).
87. Rajapakse, V.N. *et al.* CellMinerCDB for Integrative Cross-Database Genomics and Pharmacogenomics Analyses of Cancer Cell Lines. *iScience* **10**, 247-264 (2018).
88. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450 (2019).
89. Merritt, C.R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* **38**, 586-599 (2020).
90. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-20 (2014).
91. Van der Auwera, G.A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 1-11 10 33 (2013).
92. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C. & Koeffler, H.P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* **28**, 1747-1756 (2018).
93. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-40 (2010).
94. Li, J. *et al.* TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample Multiplexing. *J Proteome Res* **20**, 2964-2972 (2021).
95. Charrad, M., Ghazzali, . N., Boiteau, V., & Niknafs, A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software* **61**, 1-36 (2014).
96. Hartigan, J.A.a.W., M.A. A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society* **Series C (Applied Statistics)**, 100-108 (1979).
97. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
98. Lachmann, A., Giorgi, F.M., Lopez, G. & Califano, A. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics* **32**, 2233-5 (2016).
99. Alvarez, M.J. *et al.* Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet* **48**, 838-47 (2016).
100. Danaher, P. *et al.* Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat Commun* **13**, 385 (2022).