A cellular and spatial atlas of *TP53*-associated tissue remodeling in lung adenocarcinoma

Authors

William Zhao[1,2], Benjamin Kepecs[1,2], Navin R. Mahadevan[6,8,10], Asa Segerstolpe[4], Jason L. Weirather[8], Naomi R. Besson[9], Bruno Giotti[1,2], Brian Y. Soong[1,2], Chendi Li[12,13], Sebastien Vigneau[7], Michal Slyper[4,16], Isaac Wakiro[7,14], Judit Jane-Valbuena[4,16], Orr Ashenberg[4], Asaf Rotem[7,15], Raphael Bueno[11], Orit Rozenblatt-Rosen[4,16], Kathleen Pfaff[9], Scott Rodig[10], Aaron N. Hata[12,13], Aviv Regev[4,5,16], Bruce E. Johnson[7], & Alexander M. Tsankov[1,2,3,4*]

Affiliations

[1]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai (ISMMS), New York, NY, USA

[2]The Precision Immunology Institute, ISMMS, New York, NY, USA

[3]The Tisch Cancer Institute, ISMMS, New York, NY, USA

[4]Klarman Cell Observatory, Broad Institute of Massachusetts Institute of Technology (MIT) and Harvard, Cambridge, MA, USA

[5]Department of Biology, MIT, Cambridge, MA, USA

[6]Department of Medical Oncology, Dana-Farber Cancer Institute (DFCI), Boston, MA, USA

[7]Center for Cancer Genomics, Department of Medical Oncology, DFCI, Boston, MA, USA

[8]Department of Data Science, DFCI, Boston, MA, USA

[9]Center for Immuno-Oncology, DFCI, Boston, MA, USA

[10]Department of Pathology, Brigham and Women's Hospital (BWH), Boston, MA, USA

[11]Department of Thoracic Surgery, BWH, Boston, MA, USA

[12]Massachusetts General Hospital Cancer Center, Charlestown, MA, USA

[13]Department of Medicine, Harvard Medical School, Boston, MA, USA

[14]Present address: Louis V. Gerstner Jr. Graduate School of Biomedical Sciences at Memorial Sloan Kettering Cancer Center, New York, NY, USA

[15]Present address: Astrazeneca, 35 Gatehouse Dr. Waltham, MA, USA

[16]Present address: Genentech, 1 DNA Way, South San Francisco, CA, USA

*Correspondence: alexander.tsankov@mssm.edu (A.M.T.)

Abstract

TP53 is the most frequently mutated gene across many cancers and is associated with shorter survival in non-small cell lung cancer (NSCLC). To understand how *TP53*-mutant (*TP53*^{mut}) malignant cells interact with the tumor microenvironment (TME) at a molecular, cellular, and tissue level, we built a multi-omic cellular and spatial tumor atlas of 23 treatment-naïve NSCLC human tumors. We identified significant differences in malignant expression programs and spatial cell-cell interactions between *TP53*^{mut} and *TP53*^{WT} tumors and found that highly-entropic *TP53*^{mut} malignant cells lose alveolar identity and coincide with an increased abundance of exhausted T cells and immune checkpoint interactions with implications for response to checkpoint blockade. We also identified a multicellular, pro-metastatic, hypoxic tumor niche, where highly-plastic, *TP53*^{mut} malignant cells expressing epithelial to mesenchymal transition (EMT) programs associate with *SPP1*$^{+}$ myeloid cells and collagen-expressing cancer-associated fibroblasts. Our approach can be further applied to investigate mutation-specific TME changes in other solid tumors.

INTRODUCTION

Non-small-cell lung cancer (NSCLC) is a heterogeneous disease that has routinely been classified into different histological subtypes in the clinical setting.[1] More recently, genomic alterations have been used to characterize molecular subtypes of NSCLC that are used to stratify patients for treatment.[2-4] Tyrosine kinase inhibitors (TKIs) and other targeted therapies against oncogenic mutations and chromosomal rearrangements (*e.g.*, *EGFR*, *KRAS-G12C, ALK*) have led to a recent improvement in survival in NSCLC patients.[5] Immune checkpoint inhibitors (ICIs) (*e.g.*, anti-PD-1/PD-L1) with or without chemotherapy have proven efficacious in a subset of patients[6] and are currently part of the first-line treatment in patients with stage IV NSCLC.[7] However, due to high rates of cancer recurrence, novel approaches toward patient stratification and precision therapy are urgently needed.

Tumor protein p53 (*TP53*) is the most commonly mutated gene in NSCLC (~50%) as well as across many other cancers.[8] *TP53* mutations are more common in smokers[9] and have been linked to shorter survival in patients with NSCLC.[10] Despite being the most well-studied tumor suppressor, *TP53* has been challenging to target therapeutically due to the diversity of mutations reported and its involvement in a multitude of intracellular pathways.[11-13] Interestingly, *TP53* is known to have an impact on the tumor microenvironment (TME) during tumor development and progression in mouse models.[14,15] Mutant *TP53* is also associated with increased expression of PD-L1 and improved response to anti-PD1 therapy in human lung adenocarcinoma (LUAD),[16-21] motivating us to investigate TME differences in LUAD patients with *TP53*-mutant ($TP53^{mut}$) *vs.* *TP53*-wild type ($TP53^{WT}$) tumors.

While single cell and spatial profiling of multiple treatment-naïve, resectable NSCLC patient tumors could help advance our fundamental understanding of *TP53*-mutant cancer cells and their relationship to the TME and provide insights that may yield new therapeutic opportunities, prior studies have not tackled this question. Single-cell RNA-sequencing (scRNA-seq) has previously been used to characterize malignant cells and the TME in early stage NSCLCs, focusing on gene expression differences between the primary tumor and adjacent normal tissue,[22-25] or across metastatic sites.[26,27] Studies have also leveraged TME characteristics to stratify NSCLC patients into clinically-relevant subgroups that could predict response to immunotherapy.[19,28] One study focused on characterizing the TME and intra-tumoral heterogeneity in *EGFR*^{mut} LUAD,[29] but these previous scRNA-seq studies either did not determine the mutational status of *TP53* in tumors or did not investigate its impact on all cellular compartments in the TME.

Here, we generated the most comprehensive spatial multi-omic atlas of NSCLC to date, combining whole-exome sequencing, scRNA-seq, and spatial transcriptomics from *TP53*^{mut} and *TP53*^{WT} human tumors, to identify *TP53* mutation-associated changes in the TME and its tissue context. Our integrated atlas allowed us to identify spatially-validated cell-cell interactions that may contribute to shaping the tumor architecture of NSCLC. In doing so, we uncovered a hypoxic expression niche in *TP53*^{mut} LUAD and identified possible mechanisms by which malignant tumor cells may limit the growth of blood vessels in the TME. We further found a T cell compartment with highly exhausted features in *TP53*^{mut} *vs.* *TP53*^{WT} LUAD, which was accompanied by an enrichment in multiple immune checkpoint interactions between malignant, T, and myeloid cells. Finally, using a systematic approach to detect cellular landmarks in spatial transcriptomics data,

we observed a potential pro-metastatic signaling niche consisting of *SPP1*-expressing tumor associated macrophages (TAMs) and a collagen-expressing population of cancer-associated fibroblasts (CAFs) linked to hypoxia in the *TP53*^{mut} LUAD TME. Overall, our approach provides a framework for leveraging multi-omic technologies to understand the impact of specific genomic alterations on the cellular and spatial architecture that can enable discovery of new biomarkers and more optimal methods for patient stratification in lung cancer and other solid tumors.

RESULTS

A multi-omic atlas of NSCLC

To characterize the molecular, cellular, and spatial heterogeneity in NSCLC, we performed whole-exome sequencing (WES), scRNA-seq, and spatial transcriptomics (ST; 10X Visium) on primary tumor resections from 23 treatment-naïve, resectable NSCLC patients (**Figure 1A**), of which 18 were diagnosed as LUAD (**Supplementary Table 1)**. Twenty-one of the 23 patients had a history of cigarette smoking (Median 40, Range 12-100 pack years). WES analysis allowed us to determine the total mutational burden, copy-number alterations (CNAs), and mutational status of genes commonly altered in NSCLC, including *TP53*, *KRAS*, *EGFR*, and *STK11* in each tumor. There was high concordance between predicted CNA profiles from tumor-matched WES and scRNA-seq samples (**Figure 1B; Supplementary Figure 1A**), many of which overlapped with commonly deleted (*e.g.*, *CDKN2A*, *PTPRD*, *B2M*) and amplified (*e.g.*, *NKX2-1*, *TERT*, *MCL1*, *EGFR*) genes in LUAD.[8] After filtering, we obtained a total of 167,193 high-quality cell profiles, used CNA inference to distinguish between malignant and non-malignant epithelial cells (**Methods**), and then integrated,[30] clustered, and annotated them into 11 broad cell classes based

on coherent expression of known marker genes (**Figure 1C-D; Supplementary Figure 1B-C; Supplementary Table 2**). Distinguishing tumors by their *TP53* mutation status, we observed that the different *TP53* mutations had comparable effects in downregulating the average expression of p53 target genes in malignant cells relative to *TP53*^{WT} LUAD samples (**Figure 1E**). The remaining five tumors consisted of squamous, mucinous, and colloid histologies, and were not used for comparisons between *TP53*^{mut} and *TP53*^{WT} tumors because *TP53* mutations have been found to have histology-dependent effects on patient outcome.[31]

TP53^{mut} LUAD is associated with a distinct TME composition

The TME in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumors had distinguishable cell-intrinsic expression profiles (**Figure 1D, right**) and cell compositions (**Figure 1F, left; Supplementary Figure 1D**). In particular, there was a pronounced decrease in the proportion of endothelial cells and pericytes in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumors, which we confirmed by scoring the corresponding cell type signatures in 510 bulk RNA-seq profiles of primary LUAD tumors from The Cancer Genome Atlas (TCGA) (**Figure 1F, right**). Endothelial cell abnormalities have been linked to tumor cell progression and metastasis,[32] and pericyte depletion has been linked to hypoxia-associated epithelial to mesenchymal transition (EMT) in breast cancer.[33]

These compositional changes in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD were also supported by spatial transcriptomics of 14 tissue sections from 6 patients in our cohort (2 with *TP53*^{mut} and 4 with *TP53*^{WT} LUAD), where 6 slides from 2 patients with other NSCLC histology or with insufficient number of malignant cells were excluded from downstream analyses (**Figure 1G, H**; **Supplementary Table 1**). We used Tangram,[34] a deep learning framework, to map tumor-matched

scRNA-seq to the corresponding spatial transcriptomics data and generate a probabilistic measure of the cellular composition of each spatial transcriptomics spot (**Figure 1G**). These were highly concordant with independent, manual annotations of the corresponding hematoxylin and eosin (H&E) stains by a pathologist (**Figure 1G**, **Supplementary Figure 1E-G**). Consistent with our scRNA-seq, there was also a consistent decrease in the proportion of endothelial cells and pericytes across spots in *TP53*^{mut} *vs*. *TP53*^{WT} samples (**Figure 1H**).

Spatial interactions known to inhibit vascularization are enriched in *TP53*^{mut} LUAD

To understand the decreased proportion of vascular cells in in *TP53*^{mut} LUAD, we further used expression of established markers to identify 9 endothelial cell subsets (**Figure 2A-B, Supplementary Figure 2A-B**), mostly resembling endothelial cell subsets in the normal lung.[35] These included: aerocytes, a specialized capillary cell responsible for gas exchange in the alveolus[36] (expressing *HPGD*, *EDNRB*, and *SOSTDC1*); arterial cells (*GJA5*, *ENPP2*); capillary cells (*CA4*, *FCN3*); vascular endothelial cells (VEC), including *COL4A1*$^{+}$ VECs (VEC.COL4A1), cycling (VEC.Cycling) and interferon-stimulated (VEC.IFN) subsets; lymphatic cells; pulmonary venous cells; and systemic venous cells.

The proportion of aerocytes out of all endothelial cells was decreased in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 2C**), which was further supported by bulk RNA-seq data deconvolved using scRNA-seq-derived markers (**Supplementary Figure 2C**). This suggests a corresponding decrease in gas exchange in *TP53*^{mut} LUAD and a resulting decrease in oxygen diffusion to the area directly surrounding the alveoli at the tumor site, providing a selective advantage for tumor cell survival.[37]

Analyzing ligand-receptor pairs (**Methods**), several interactions between malignant and endothelial cells known to inhibit endothelial cell growth and function were induced in *TP53*^{mut} LUAD (**Figure 2D**), including *SEMA3A-NRP1* and several Ephrin interactions (*EPHA1-EFNA1*, *EPHB2-EFNB1*).[38-40] *EFNB1* is more highly expressed in aerocytes compared to other endothelial populations (**Supplementary Figure 2D**), while *EFNB2* is predominantly expressed in arterial cells. *NRP1* is more highly expressed in aerocytes as well as in VEC.COL4A1 cells (**Supplementary Figure 2D**). Spatial transcriptomics from matched tumor samples showed a significantly higher spatial correlation in the expression of *SEMA3A-NRP1*, *EPHA1-EFNA1* and *EPHB2-EFNB1* ligand-receptor pairs in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 2E, Supplementary Figure 2E**). Taken together, there is enriched, spatially correlated expression of known inhibitory interactions between malignant and endothelial cells in *TP53*^{mut} LUAD, which may explain the accompanying decrease in aerocyte and endothelial cell abundance.

Dichotomous activity of malignant cell programs in LUAD associates with survival

While malignant cells from one tumor can occupy a gradient of expression states rather than discrete cell subsets,[41,42] different cancer expression programs and their usage across cells and patients have not been comprehensively defined in LUAD at single-cell resolution. To identify shared cell states across tumors despite predominant, patient-specific differences in malignant cells,[41,42] we used canonical correlation analysis (CCA)[43] to integrate and cluster malignant cells into 18 subsets shared across patients (**Figure 3A**; **Supplementary Figure 3A**). These subsets were defined by marker genes associated with hallmark cancer processes (*e.g.*, cell cycle, hypoxia, glycolysis, partial EMT (pEMT), interferon gamma (IFNG) and TNFA signaling) and with lung

epithelial cell identity (*e.g.*, alveolar type 2 (AT2-like), ciliated, secretory) (**Figure 3B**, **Supplementary Figure 3B**). Other malignant clusters showed marked expression of antigen presentation (MHCII), stress response,[44] metallothionein, and senescence-related genes,[45] which have not been previously defined using scRNA-seq in NSCLC primary tumors. We thus defined malignant LUAD programs by the marker genes associated with each malignant cells cluster (**Methods**), and scored them across all the cells.

The programs clustered into two main groups by the correlation of their scores across cells (**Figure 3C**) and patients (**Supplementary Figure 3C**): one consisting of cell cycle, hypoxia, pEMT, and glycolysis programs, and another of alveolar-like, antigen-presentation, and stress response programs. Malignant program scores in these two groups were strongly anticorrelated, suggesting that they were regulated by two distinct and opposing pathways (**Figure 3C**). This is consistent with a previous LUAD study where cells from patient-derived xenografts consisted of two distinct intra-tumoral subgroups, distinguished by a gene module consisting primarily of cell cycle genes.[46]

Moreover, expression of several malignant programs (*e.g.*, pEMT, hypoxia, cell-cycle, metallothionein, glycolysis) are associated with shorter survival across LUAD patients in TCGA, while expression of the antigen-presentation, MHCII program was associated with survival benefit, independent of *TP53* status (**Figure 3D**, **Supplementary Figure 3D**; Cox proportional-hazards regression test). Previously, various bulk TCGA-derived signatures have been shown to be predictive of outcome in NSCLC[47,48] but our *de novo* approach allowed for more granular and less biased program discovery compared to previous studies.

Loss of AT2 identity is associated with increased cellular entropy in *TP53*^{mut} LUAD

To identify cell-intrinsic differences between malignant cells in *TP53*^{mut} and *TP53*^{WT} LUAD, we compared the mean activity score of different malignant programs between the two groups and found an increased expression in *TP53*^{mut} LUAD of several programs associated with poor outcome (*e.g.*, cell cycle, glycolysis, metallothionein, pEMT), and a decreased expression of the AT2-like program and *TP53* targets (**Figure 3E, left**). These trends were corroborated in bulk RNA-seq data from 263 *TP53*^{mut} and 247 *TP53*^{WT} LUAD primary tumors (**Figure 3E, right**). Notably, spatial co-expression of cell cycle with glycolysis and of hypoxia with pEMT programs was higher in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD tumors (**Figure 3F**).

The decreased expression of the AT2-like program in *TP53*^{mut} LUAD was consistent with a loss of AT2 cell identity. Scoring each malignant cell profile across patients for normal lung epithelial cell signatures,[49] *TP53*^{WT} LUAD malignant cells scored highly for AT2 markers (**Figure 3G**), whereas cells from tumors with squamous, neuroendocrine, or mucinous/colloid subtypes expressed the expected corresponding basal, neuroendocrine, or goblet signatures.[50] Interestingly, AT2 scores for *TP53*^{mut} LUAD malignant cells were dramatically lower than *TP53*^{WT} LUAD samples and were more similar to scores in the tumors from non-adenocarcinoma subtypes, suggesting a loss of alveolar cell identity. This loss of AT2 cell score was not associated with differences in TMB in our data (**Supplementary Figure 3E**).

We have previously reported loss of alveolar cell identity and reversion to a progenitor-like, highly-plastic cell state (HPCS) during LUAD progression in mice with somatic activation of oncogenic **K**RAS-G12D, which was greater in those where the **p**53 tumor suppressor was deleted

(KP model) than wild-type p53 (K model).[51] Consistent with these observations in the KP mouse model, we found an increase in entropy[52] in *TP53*^{mut} LUAD malignant cell profiles and bulk RNA-seq profiles compared to *TP53*^{WT} LUAD malignant single cell and bulk profiles, respectively, and compared to non-malignant single cells or adjacent normal bulk samples (**Figure 3H**). Cell cycle and pEMT programs displayed the highest entropy amongst malignant subsets (**Figure 3I**) and the greatest similarity to the HPCS cell state described in the KP mouse model (**Figure 3J**). Taken together, malignant cells from *TP53*^{WT} LUAD tumors were enriched for antigen presentation and AT2-like cell programs, while those from *TP53*^{mut} LUAD tumors were highly entropic, plastic, and associated with programs predictive of poor outcome, including cell cycle, glycolysis, and pEMT.

Functional differences in mesenchymal cells between *TP53*^{mut} and *TP53*^{WT} LUAD

The heterogeneity of mesenchymal cells and their potential role in more rapid tumor progression and shorter patient survival has been documented in NSCLC.[23,53] Our mesenchymal cells were partitioned into 13 subsets (**Figure 4A-B**, **Supplementary Figure 4A-B**). Two subsets most closely resembled normal lung myofibroblasts[49] (**Figure 4C**), each associated with distinct functional pathways (**Supplementary Figure 4C**): CAF.COLs expressing high levels of collagens (*COL1A2*, *COL3A1*, *COL1A1*), which shares markers with recently described *LRRC15*$^{+}$ myofibroblasts,[54] and myofibroblasts, expressing both fibroblast and smooth muscle cell markers. We also annotated two pericyte-like subsets, one expressing canonical pericyte markers (*HIGD1B*, *COX4I2*), and one expressing *COL4A2* and EMT-related genes (**Supplementary Figure 4A**); airway and vascular smooth muscle subsets, and alveolar (CAF.ADH1B) and adventitial (CAF.Adventitial) fibroblasts, which correspond to recently described human lung alveolar and

adventitial fibroblasts;[54] and fibroblasts expressing lipid metabolism (CAF.APOE), ribosomal (CAF.Ribo), complement and chemotaxis (CAF.Complement), and interferon-stimulated (CAF.ISGs) genes. Of note, we did not observe subsets similar to normal lung fibromyocytes or lipofibroblasts (**Figure 4C**). The different mesenchymal subsets were associated with different active regulons (**Figure 4D**), as inferred by pySCENIC,[55] with high regulon specificity (**Supplementary Figure 4D**). As expected, cycling CAFs had an active *E2F1* regulon, a cell cycle regulator; CAF.ISGs activated the *IRF7* and *STAT1* regulons, controlling interferon responsive genes; and CAF.COLs activated a *TWIST1* regulon, previously linked to EMT in cancer.[56]

Based on the inferred composition of CAF subsets in spatial transcriptomics spots across sections (**Methods**), pericytes were significantly spatially correlated with endothelial cells (**Supplementary Figure 4E**, Pearson correlation P value $< 10^{-15}$); CAF.COLs with vascular smooth muscle and CAF.ADH1B, CAF.Complement with CAF.Adventitial, pericytes with CAF.ISGs, and airway smooth muscle cells with CAF.APOE and CAF.Ribo. Spatially, most mesenchymal subsets were negatively correlated with malignant cells, indicating occlusion of fibroblasts from the tumor core (**Supplementary Figure 4E**).

Comparing the relative proportions of mesenchymal cell subsets (out of all mesenchymal cells) between *TP53*^{mut} *vs*. *TP53*^{WT} LUAD, CAF.ADH1B and pericytes were decreased in *TP53*^{mut} LUAD (as observed for broad cell classes) and CAF.ISGs and CAF.Ribo were increased (**Figure 4E**), also supported by deconvolved bulk RNA-seq (**Supplementary Figure 4F**). *TGFB2-TGFBR2* and *TGFBR2-TGFB3* interactions between mesenchymal and malignant cells were enriched in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD, and *TGFB2-TGFBR2* were spatially correlated in *TP53*^{mut}

LUAD (**Figure 4F-G**). *VEGFA* expression in myofibroblasts is higher in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Supplementary Figure 4G**), consistent with previous reports that myofibroblasts can regulate angiogenesis in response to hypoxia.[57] Overall, the mesenchymal compartment is remodeled in *TP53*^{mut} LUAD in both cell composition and spatial organization with an enrichment of TGF-β interactions between malignant and mesenchymal cells.

Increased *SPP1* expression in myeloid cells in *TP53*^{mut} LUAD is linked to hypoxia and EMT

Myeloid cells play an important role in tumor progression and may be a compelling therapeutic target in cancer.[58] While NSCLC myeloid cell populations were generally characterized,[19,24] the impact of *TP53* mutations on myeloid cells, and their interactions with malignant, stromal, and immune compartments in the TME, has not been studied in detail in humans.

Myeloid cells partitioned into 14 subsets, largely consistent with previous NSCLC atlases[19] (**Figure 5A-B**, **Supplementary Figure 5A**), but with significant compositional differences between *TP53*^{mut} and *TP53*^{WT} LUAD. These include a significant decrease in *TP53*^{mut} LUAD in the proportion of TAM.FABP4 cells, which resemble tissue-resident alveolar macrophages (TAMs), and an increase in *SPP1*-expressing TAMs (TAM.SPP1) (**Figure 5C**). In agreement, *SPP1* expression was significantly increased both in TCGA bulk RNA-seq data and in myeloid pseudobulk comparisons between *TP53*^{mut} and *TP53*^{WT} LUAD tumors (**Figure 5D, left**). The proportion of *CXCL9/10/11* expressing TAMs (TAM.CXCLs) was also higher in *TP53*^{mut} LUAD patients both in the scRNA-seq data and in deconvolution analysis of TCGA bulk RNA-seq data (**Supplementary Figure 5B**).

Myeloid cell subsets in *TP53*^{mut} and *TP53*^{WT} LUAD were also associated with distinct cell intrinsic gene programs, such that genes differentially expressed between myeloid subsets from *TP53*^{mut} *vs*. *TP53*^{WT} tumors were enriched for EMT, hypoxia, and interferon response pathways (**Supplementary Figure 5C**), especially in TAM.SPP1 cells (**Figure 5E**). *SPP1* expression has been previously associated with EMT, and with early lymph node metastasis in LUAD,[59] as well as with shorter survival in TCGA LUAD independent of *TP53* mutational status (**Supplementary Figure 5E**). Notably, the genes most correlated with *SPP1* expression in myeloid cells were enriched for EMT, hypoxia and glycolysis functions (**Figure 5D, right**) and included VEGF receptors *NRP1* and *FLT1* (**Supplementary Figure 5D**). Thus, the increased expression of *SPP1* in *TP53*^{mut} LUAD myeloid cells is tightly linked with induction of genes involved in EMT and hypoxia programs.

To identify ligand-receptor interactions that may contribute to a distinct myeloid compartment in *TP53*^{mut} LUAD, we conducted differential ligand-receptor interaction analysis. Putative interactions enriched in *TP53*^{mut} LUAD include *VEGFA-FLT1* and *NRP1-SEMA3A* (**Figure 5F**), involved in recruitment of myeloid cells to tumor sites,[60] and *TGFB2-TGFBR1* interactions between malignant and myeloid cells, which are known to promote monocyte differentiation into TAMs,[61] and may explain the relative increase in the monocyte-macrophage population in *TP53*^{mut} LUAD. Additionally, *NRP1-VEGFA*, *NRP1-VEGFB*, *FLT1-VEGFB*, and *FLT1-PGF* interactions between myeloid and mesenchymal cells, induced in *TP53*^{mut} LUAD, suggest that mesenchymal cells may contribute to monocyte recruitment to the TME (**Figure 5F**).[60] In addition, *CXCL11-CXCR3* interactions between myeloid and B/T cells were enriched in *TP53*^{mut} LUAD (**Figure 5F**) with *CXCR3* specifically expressed by NK, B, and T cells (**Figure 5G**) and *CXCL11* induced in

multiple myeloid subsets in *TP53*^{mut} LUAD (**Figure 5H**), most highly in TAM.CXCLs (**Supplementary Figure 5F**), which constitute a higher proportion of myeloid cells in *TP53*^{mut} LUAD. Thus, myeloid cells may enhance T cell recruitment in *TP53*^{mut} LUAD. Finally, there was a significant increase in putative *SPP1*-mediated interactions between myeloid and endothelial cells or pericytes in *TP53*^{mut} LUAD (**Figure 5F**), also supported by increased spatial correlation of *SPP1* and *CD44* expression in *TP53*^{mut} tumors (**Figure 5I**). Thus, *TP53*^{mut} LUAD has increased expression of putative interactions involved in monocyte and lymphocyte recruitment, TAM differentiation, and downstream effector functions via *SPP1*.

Enrichment of T cells with exhaustion features and putative immune checkpoint interactions in *TP53*^{mut} LUAD

Lymphoid cells partitioned into 13 subsets of T and NK cells (**Figure 6A-B, Supplementary Figure 6A**), follicular and marginal zone B cells, IgA and IgG plasma cells, plasmacytoid DCs, and mast cells (**Supplementary Figure 6B**). While there were no significant differences in B cell composition, *TP53*^{mut} LUAD had a decreased proportion of CD4^{+} tissue-resident memory T cells (CD4.TRM) and an increase in pre-exhausted-like,[62] *GZMK*-expressing CD8^{+} (CD8.GZMK), exhausted-like (T.Exhausted), and follicular helper (TFH) T cell populations out of all T and NK cells (**Figure 6C**), supported by deconvolved bulk RNA-seq from samples in TCGA (**Figure S6C**). Expression of *CXCL13*, a top predictor for response to checkpoint inhibition therapy,[63] was significantly increased in exhausted-like and TFH subsets from *TP53*^{mut} patients (**Supplementary Figure 6E**). Taken together, increased *CXCL13* expression and proportions of *GZMK*-expressing CD8^{+}, exhausted-like, and TFH T cells all suggest a heightened immunogenic potential in the *TP53*^{mut} LUAD TME that may contribute to a more favorable ICI immunotherapy response.[63-66]

Ligand-receptor analysis of putative immune checkpoint interactions[67] between T, myeloid, and malignant cells showed an enrichment in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD in *PDCD1-CD274* and *TIGIT-PVR* putative interactions between T and malignant cells (**Figure 6D**) and in *CTLA4-CD86* and *HAVCR2-LGALS9* putative interactions between T and myeloid cells. *PDCD1*, *CTLA4*, and *TIGIT* expression was higher in *TP53*^{mut} bulk RNA-seq LUAD samples from TCGA (**Supplementary Figure 6D**) and in multiple T cell subsets in our scRNA-seq data (**Supplementary Figure 6E**), where *TIGIT* was the most commonly expressed ICI target across T cells, compared to *CTLA4* and *PDCD1* (**Figure 6E,** in both *TP53*^{mut} and *TP53*^{WT} tumors). *PVR* expression was higher in malignant cells from *TP53*^{mut} *vs*. *TP53*^{WT} tumors (**Figure 6F left**) and highly correlated with the expression of the cell cycle program in malignant cells, suggesting that *PVR* may be activated during cancer cell proliferation (**Figure 6F right**). Finally, there was significantly higher spatial colocalization of *TIGIT* and *PVR* expression in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 6G**). Taken together, increased expression of *TIGIT* in T cells and *PVR* in malignant cells both contribute to increased spatial co-expression of *PVR* and *TIGIT* in *TP53*^{mut} LUAD.

Spatially-defined fibroblast-macrophage niche enriched in hypoxia and EMT programs in *TP53*^{mut} LUAD

Finally, to comprehensively map inter-compartment cellular colocalization, we deconvolved the proportion of cell classes and subsets across all spatial transcriptomics spots and calculated the average correlation of cell type proportions across all LUAD sections, and their difference between *TP53*^{mut} and *TP53*^{WT} LUAD (**Figure 7A, Supplementary Figure 7A**). Interestingly, there was a

significantly lower colocalization of endothelial cells with mesenchymal cells and of endothelial cells with myeloid cells in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 7B**), in agreement with the increase in hypoxia related programs in mesenchymal and myeloid subsets from *TP53*^{mut} tumors. In addition, spatial colocalization of cells typically found in tertiary lymphoid structures (TLSs) (B cells, plasma cells, TFH cells, and other CD4^{+} cell populations) was also significantly decreased in *TP53*^{mut} LUAD and was accompanied by increased colocalization of B, Plasma, and malignant cells (**Supplementary Figure 7B**). B cell infiltration has been associated with increased PD1/PD-L1 expression, TMB, and overall survival in NSCLC[68] and likely contributes to the heightened immunogenicity of *TP53*^{mut} LUAD tumors.

We further identified multicellular communities by an NMF-based method (**Methods**) to derive 15 factors with varying participation of different subsets (**Figure 7C**). Factor 7 (NMF7) contained multiple myeloid and mesenchymal subtypes with known significance in tumor progression, including TAM.SPP1, CAF.COLs, and myofibroblasts, which were highly colocalized with each other, as well as with hallmark hypoxia and EMT programs (**Figure 7D**). Spots high in NMF7 were low in malignant cell proportions and high in myeloid and mesenchymal cell proportions (**Supplementary Figure 7C**). NMF7 was also more positively correlated with hypoxia and EMT signature expression in malignant-low spots and more negatively correlated with endothelial cells across all spots in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD (**Figure 7E**). Since CAFs and macrophages are both thought to induce EMT in tumors,[69,70] NMF7 may represent a hypoxic fibroblast-macrophage niche that resides on the periphery of the tumor core and may contribute to the early EMT phenotype in malignant cells in *TP53*^{mut} LUAD.

DISCUSSION

Targeted therapies for NSCLC are currently directed at oncogenic mutations and chromosomal rearrangements (e.g., *EGFR*, *KRAS-G12C, ALK*) yet *TP53* mutations, despite being common in tumors and extensively studied, currently have no therapies commercially available. Restoration of *Trp53* expression has been shown to cause tumor regression in mice,[71] leading to efforts to treat *TP53*^{mut} tumors by restoring wild-type p53 activity.[12] We hypothesized that cellular and spatial profiling of the TME in *TP53*^{mut} LUAD could provide a more detailed understanding of how *TP53* mutations could contribute to shorter survival and suggest new therapeutic opportunities. To this end, we combined scRNA-seq of 167,193 cells, 23 tumor and matched normal WES, and 20 spatial transcriptomics sections across 23 NSCLC patients to characterize the LUAD TME of *TP53*^{mut} *vs*. *TP53*^{WT} tumors. Together, these represent the most comprehensive map of NSCLC to date and provide a valuable resource of cell annotations, markers, and integrated data sets for the scientific community.

Our integrative analysis revealed a substantial remodeling of the TME associated with *TP53* mutations (**Figure 7F**), including decreased presence of pericytes and endothelial aerocytes. The causes of TME remodeling are likely complex but our study uncovers multiple mechanisms that may play a role. We highlight three putative ligand-receptor interactions between malignant and endothelial cells (*SEMA3A-NRP1*, *EPHA1-EFNA1*, *EPHB2-EFNB1*) that were enriched in *TP53*^{mut} tumors and have known roles in inhibiting vascularization, and which were significantly more colocalized in the *TP53*^{mut} spatial data. *TP53*^{mut}-specific depletion of endothelial cells was accompanied by increased hypoxia and pEMT program expression in malignant cells. Hypoxia

has been shown to induce EMT through HIF-1α as well as other mediators[72,73] and to positively select for p53-deficient cells that have lost apoptotic potential in mice.[74] Thus, a hypoxic TME may be providing a survival advantage for highly plastic *TP53*^{mut} LUAD malignant cells, which lose alveolar identity and experience increased proliferative capacity and metastatic potential. Furthermore, *de novo* discovery of spatial multicellular communities revealed a highly hypoxic, EMT-promoting spatial niche enriched for *SPP1*$^{+}$ TAMs, myofibroblasts, and collagen producing CAFs. *SPP1* expression was highly upregulated in *TP53*^{mut} monocytes and TAMs and tightly linked with regulation of hypoxia and EMT. Increased TGF-β signaling in CAFs likely also help shape an EMT-supporting niche in *TP53*^{mut} samples, where *TGFB2-TGFBR2* interactions between malignant cells and CAFs were significantly more enriched in both scRNA-seq and spatial data. Previous studies linking mutant p53 to metastatic potential have focused on malignant cell-intrinsic processes,[75] as the effects of mutant p53 on tumor-stromal and tumor-immune crosstalk have been described only recently in model systems.[76] Our study is the first to profile the cellular and spatial context of *TP53* mutations in LUAD to provide insight into how *TP53*^{mut} malignant cells could remodel the TME to promote tumor survival and metastasis in patients.

Interestingly, one aspect of TME remodeling in *TP53*^{mut} LUAD is a higher tumor infiltration by B cells and CD8^{+} T cells. Increased lymphocyte recruitment may be in part due to enriched *CXCL11-CXCR3* ligand-receptor interactions between myeloid and B/T cells in *TP53*^{mut} tumors. We observed some overlap in myeloid and T cell phenotypes in *TP53*^{mut} LUAD with the anti-tumor immunity hub we reported in a previous study in colorectal cancer,[77] but additionally report B/Plasma tumor infiltration, T cell exhaustion, and the enrichment of *SPP1*-expressing TAMs. Our findings of a potentially more immunogenic TME in *TP53*^{mut} LUAD is consistent with a previous

study reporting higher levels of anti-tumor immune signatures in *TP53*^{mut} vs *TP53*^{WT} LUAD using bulk RNA-seq data from TCGA.[78] Previous studies have reported that p53 mutations result in increased expression of PD-L1 through decreased inhibition *via* miR-34[79] and patients with *KRAS-TP53* co-mutations in LUAD have higher expression of PD-L1 and response to anti-PD1 treatment compared to patients with *KRAS*-only mutant tumors,[16,18,20,80] as opposed to worse response observed in patients with *STK11* and *KEAP1* mutations.[81] In agreement, our computational screen for immune checkpoint interactions showed enrichment in *PDCD1-CD274* in *TP53*^{mut} LUAD patients. Moreover, we observed an enrichment in *CTLA4-CD86* and *TIGIT-PVR* interactions in *TP53*^{mut} LUAD patients, which have not been studied in the context of *TP53* mutations to the best of our knowledge. Increased expression of *TIGIT* in T cells and *PVR* in malignant cell both contribute to the spatially-resolved enrichment of *TIGIT-PVR* interactions, which we observed in all 8 *TP53*^{mut} resectable LUAD patients in our cohort. Recent trials demonstrate benefit to combining anti-PD1 ICI with chemotherapy as a neo-adjuvant in resectable LUAD.[82] Future studies are needed to assess the predictive potential of *TP53* mutational status for response to these therapies.

More generally, the methodology introduced in this study can be used to study the effect of *TP53* or other mutations on the TME in lung cancer and other solid tumor types and to infer intercellular pathways and communities to target for tumor suppressors in other contexts, where targeted therapies currently do not exist.

ACKNOWLEDGMENTS

We thank Leslie Gaffney for extensive help in editing and graphical design of the figures in the manuscript, the Broad Genomics Platform, Broad Flow Cytometry Facility, Pathology and Surgery Departments at MGH and BWH; members of the Bueno, Hata, Regev, Johnson, and Tsankov labs. The project is part of the Human Tumor Atlas Pilot Project (HTAPP) consortium and National Institutes of Health (NIH) HTAN (Human Tumor Atlas Network) consortium paper package. This work was made possible by the generous support of Leidos and the Klarman Cell Observatory (to O.R.-R. and A. Regev). The project was also funded in part by federal funds from the National Cancer Institute (NCI), NIH Task Order HHSN261100039 under Contract HHSN261201500003, and U2CCA233195. A.M.T. was supported by Department of the Army Lung Cancer Research Program Career Development Award W81XWH2210079 and ISMMS seed funding. W.Z. and B.Y.S. were supported by NIH grant 2T32GM007280-41 as part of the ISMMS Medical Scientist Training Program. The graphical summary was created with Biorender.com.

FIGURE LEGENDS

Figure 1. Spatially resolved, multi-omic atlas of NSCLC.

(A) Schematic of study design, genomic profiling, cohort stratification, and downstream analyses.

(B) Inferred $\log_2$ copy number alteration (CNA) ratio in scRNA-seq malignant cells (averaged across malignant cells per patient) (left) and WES (right) data from matched patient samples. Highlighted are locations of common focal amplifications and deletions in LUAD.

(C) Expression of representative markers across each annotated broad cell class.

(D) Uniform Manifold Approximation and Projection (UMAP) visualization of 167,193 cells, colored by annotated cell classes (left) and *TP53* mutational status (right).

(E) Mean expression of p53 target genes in malignant cells isolated from *TP53*^{WT} (blue dots) and *TP53*^{mut} (red dots) LUAD primary tumor samples.

(F) Left: Proportion of endothelial cells and pericytes found in the scRNA-seq data from *TP53*^{WT} (blue dots) and *TP53*^{mut} (red dots) LUAD patients. Right: Average $\log_{10}$ expression of highly specific endothelial and pericyte markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(G) Left: Integration of ST data with tumor-matched scRNA-seq data using Tangram allows for deconvolution of the proportion of different cell classes and subtypes. Middle: Spatial distribution of cell classes using Tangram (top), and a representative H&E stain annotated by a pulmonary pathologist (bottom). Right: violin plots comparing proportion of computationally inferred cell classes (y-axis) within manually annotated categories (x-axis). Mann-Whitney-Wilcoxon Test *p*-values were computed; Kruskal-Wallis FDR, $p \leq 0.05$.

(H) Proportion of endothelial cells (top) and pericytes (bottom) in spatial transcriptomics spots from *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD sections. Mann-Whitney-Wilcoxon Test *p*-values were computed.

Figure 2. Spatially-resolved, inhibitory interactions between endothelial and malignant cells enriched in *TP53*^{mut} LUAD.

(A) UMAP of 15,320 endothelial cells, integrated across patients and colored by annotated subtype.

(B) Feature plots of select endothelial subtype markers.

(C) Comparison of endothelial subtype cell proportions in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon Test was used for computing *p*-values.

(D) Dot plot of differential cell-cell interactions between endothelial cells and other cell classes. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples. Black outline indicates *p*-value ≤ 0.05, as assessed by Fisher's exact test.

(E) Spatial feature plot of *SEMA3A* ligand and *NRP1* receptor in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *SEMA3A* and *NRP1* expression between *TP53*^{WT} and *TP53*^{mut} samples (right). Mann-Whitney-Wilcoxon Test *p*-value was computed.

Figure 3. Enrichment of hypoxia and pEMT programs, loss of AT2 identity and increased cellular entropy in *TP53*^{mut} LUAD malignant cells.

(A) UMAP of 33,377 malignant cells integrated across patients and colored by annotated cancer expression program.

(B) Dot plot of two selected markers representative of each malignant program.

(C) Pearson correlation of malignant program scores across all malignant cells.

(D) Lollipop plot of negative $\log_{10}$ *p*-values from Cox proportional-hazards regression analysis linking malignant programs scores to patient disease outcome, corrected for *TP53* mutational status.

(E) Comparison of malignant program mean expression between *TP53*^{mut} and *TP53*^{WT} patients in scRNA-seq (left) and in bulk RNA-seq data from TCGA (right).

(F) Left: Representative example of the spatial distribution of expression of four malignant programs in a *TP53*^{mut} patient. Right: Bar plots comparing colocalization (spatial correlation) of

malignant programs between *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(G) Module scores for normal lung epithelial cell signatures averaged across malignant cells, ordered by histology and then *TP53* mutational status. Log_{10} SNP count for each tumor is shown on the bottom.

(H) Distribution of entropy scores across non-malignant cells (grey), malignant *TP53*^{WT} cells (blue), and malignant *TP53*^{mut} cells (red) in the scRNA-seq data (left) and across adjacent normal (grey), *TP53*^{WT} (blue), and *TP53*^{mut} (red) bulk LUAD samples from TCGA (right). Mann-Whitney-Wilcoxon Test *p*-values were calculated.

(I) Heatmap of mean entropy score for malignant subsets. Mann-Whitney-Wilcoxon Test *p*-values were calculated comparing mean entropy score for each subset *vs*. mean entropy score of all malignant cells.

(J) Stacked bar plot of proportion of HPCS cells[51] classified into different malignant subsets from this study.

Figure 4. Functional differences of mesenchymal subtypes in *TP53*^{mut} *vs*. *TP53*^{WT} LUAD.

(A) UMAP of 31,868 mesenchymal cells integrated across patients and colored by annotated subtype.

(B) Feature plot of select mesenchymal subtype markers.

(C) Confusion matrix of overlap between de novo mesenchymal annotations and predicted identities using a classifier trained on annotated scRNA-seq data from the normal lung.[49]

(D) Z-scores of most specific regulon activity for each mesenchymal subtype, computed using pySCENIC.[55]

(E) Comparison of mesenchymal subtype cell proportions in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(F) Dot plot of differential cell-cell interactions between mesenchymal cells and other cell classes. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples. Black outline indicates *p*-value ≤ 0.05, as assessed by Fisher's exact test.

(G) Spatial feature plot of *TGFB2* ligand and *TGFBR2* receptor in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *TGFB2* and *TGFBR2* expression between *TP53*^{WT} and *TP53*^{mut} samples (right). Mann-Whitney-Wilcoxon Test *p*-values were computed.

Figure 5. *SPP1*$^{+}$ TAMs are enriched in *TP53*^{mut} LUAD myeloid compartment.

(A) UMAP of 18,086 myeloid cells integrated across patients and colored by annotated subtype.

(B) Feature plot of select myeloid subtype markers.

(C) Comparison of myeloid subtype cell proportions in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(D) Left: Comparison of *SPP1* expression in TCGA bulk RNA-seq (left) and scRNA-seq (middle) data from *TP53*^{mut} versus *TP53*^{WT} primary tumor samples. Mann-Whitney-Wilcoxon Test *p*-values were computed. Right: Horizontal bar plot of negative $\log_{10}$ *q*-values from gene set enrichment analysis of the top 100 positively correlated genes with *SPP1* in myeloid cells.

(E) Differential enrichment of hallmark hypoxia and EMT module score for each myeloid subtype, comparing pseudobulk subtype profiles for *TP53*^{mut} versus *TP53*^{WT} patients. Mann-Whitney-Wilcoxon Test *p*-values were calculated.

(F) Dot plot of differential cell-cell interactions between myeloid cells and other cell classes. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples. Black outline indicates *p*-value ≤ 0.05, as assessed by Fisher's exact test.

(G) Dot plot of the expression of *CXCR3* and *CXCL11* across all broad level cell classes.

(H) Differential expression of ligand and receptor genes from the interactions displayed in **Figure 5F**, comparing pseudobulk subtype profiles for *TP53*^{mut} versus *TP53*^{WT} patients. Mann-Whitney-Wilcoxon Test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

(I) Spatial feature plot of *SPP1* ligand and *CD44* receptor expression in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *SPP1* and *CD44* expression between *TP53*^{WT} and *TP53*^{mut} samples (right). Mann-Whitney-Wilcoxon Test *p*-values were computed.

Figure 6. Immune checkpoint interactions and exhausted-like lymphoid cells are enriched in *TP53*^{mut} LUAD.

(A) UMAP of 40,293 T and NK cells integrated across patients and colored by annotated subtype.

(B) Feature plot of select T and NK cell subtype markers.

(C) Boxplots of the proportion of cells for select T cell subtypes in *TP53*^{mut} and *TP53*^{WT}. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(D) Enrichment of ICI targetable ligand-receptor interactions between T with malignant cells (top), and T with myeloid cells (bottom). Horizontal bar plots show negative $\log_{10}$ *p*-value of comparison between *TP53*^{mut} and *TP53*^{WT}. Fisher's exact test *p*-values were computed.

(E) Stacked bar plot depicting proportion of T and NK cells expressing different combination of ICI targets.

(F) Left: Boxplot of *PVR* expression in *TP53*^{WT} (blue) and *TP53*^{mut} (red) patients. Right: Correlation of *PVR* expression with cell cycle score averaged across malignant cells for each patient, where blue dots represent *TP53*^{WT} and red dots represent *TP53*^{mut} patients.

(G) Spatial feature plot of *PVR* and *TIGIT* expression in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *PVR* and *TIGIT* expression between *TP53*^{WT} and *TP53*^{mut} samples (right). Mann-Whitney-Wilcoxon Test *p*-values were computed.

Figure 7. Spatial multicellular hypoxic niche enriched in EMT activity in *TP53*^{mut} LUAD.

(A) Spatial correlation (left) and differential spatial correlation between *TP53*^{mut} and *TP53*^{WT} (right) in high level cell proportion across ST spots. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples for the right panel.

(B) Box plots comparing the spatial correlation (colocalization) of endothelial cells with mesenchymal (left), myeloid (middle), and malignant (right) cells across spots in *TP53*^{WT} and *TP53*^{mut} spatial transcriptomics samples.

(C) Heatmap of the 15 factors depicting multicellular programs, as calculated using NMF.

(D) Spatial distribution of the enrichment of different myeloid and mesenchymal subtypes comprising NMF7 (left) and of expression programs spatially correlated with NMF7 (right).

(E) Box plots comparing spatial correlation of NMF7 with hallmark hypoxia score (left), hallmark EMT score (middle) and endothelial cell proportion (right) in *TP53*^{WT} and *TP53*^{mut} ST samples.

(F) Graphical summary of predicted differences between the *TP53*^{WT} (left) and *TP53*^{mut} (right) TME.

Supplementary Figure 1. Correspondence of WES, scRNA-seq, and spatial transcriptomics data across patients.

(A) Spearman correlation shows correspondence between the $\log_2$ CNA ratio predicted by scRNA-seq and WES data from the same patient. Correlation was computed across genes for each patient.

(B) UMAP of 167,193 cells, colored by patient identity.

(C) Stacked bar plot displaying the composition of broad cell classes for each scRNA-seq patient sample.

(D) Top: Proportion of broad cell classes found in the scRNA-seq data from *TP53*^{WT} (blue dots) and *TP53*^{mut} (red dots) LUAD patients. Bottom: Average $\log_{10}$ expression of highly specific broad cell class markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(E) Stacked bar plot depicting mean cell composition across spatial transcriptomics samples, predicted using Tangram mapping.

(F) Spatial feature plots of broad cell classes predicted by Tangram mapping.

(G) Mean proportion of Tangram-predicted cell classes (y-axis) across different manual spot annotations.

Supplementary Figure 2. Endothelial subtypes and interactions.

(A) Dot plot of endothelial subtype-specific markers.

(B) Stacked bar plot displaying the composition of endothelial subtypes for each scRNA-seq patient sample.

(C) Average $\log_{10}$ expression of specific aerocytes (left) and arterial (right) markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(D) Left: Dot plot of genes involved in interactions displayed in **Figure 3D**. Right: Pseudobulk differential expression of the same genes *TP53*^{mut} *vs*. *TP53*^{WT} patients. Mann-Whitney-Wilcoxon Test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

(E) Left: Spatial feature plot of *EFNA1* ligand and *EPHA1* receptor in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *EFNA1* and *EPHA1* between *TP53*^{mut} and *TP53*^{WT} samples. Right: Spatial feature plot of *EFHB2* and *EFNB1* in a representative *TP53*^{mut} sample. Box plot comparing spatial correlation of *EFHB2* and *EFNB1* between *TP53*^{mut} and *TP53*^{WT} samples. Mann-Whitney-Wilcoxon Test *p*-values were computed.

Supplementary Figure 3. Characterization of the malignant compartment in NSCLC.

(A) UMAP of 33,377 malignant cells, integrated across patients and colored by patient.

(B) Dot plot of gene set enrichment outputs (Hallmark) of 100 most differentially expressed markers for each malignant subcluster.

(C) Pearson correlation of average malignant program scores across patients, analogous to Figure 1C.

(D) Kaplan Meier survival curves showing association of malignant program expression with survival in bulk TCGA LUAD. Cohort of 510 patients were stratified by program expression into high, medium, and low expression categories. Cox proportional-hazards model used to calculate *p*-value corrected for *TP53* mutational status.

(E) Comparison of $\log_{10}$ SNP and CNA abundance between *TP53*^{mut} *vs. TP53*^{WT} patients.

Supplementary Figure 4. Mesenchymal compartment differences between *TP53*^{mut} and *TP53*^{WT} LUAD.

(A) Dot plot of mesenchymal subtype-specific markers.

(B) Stacked bar plot displaying the composition of mesenchymal subtypes for each patient sample.

(C) Dot plot of gene set enrichment outputs (Hallmark) of the 100 most differentially expressed markers for each mesenchymal subcluster.

(D) Regulon specificity scores (RSS) calculated using pySCENIC plotted across 4 different mesenchymal subtypes.

(E) Spatial correlation (colocalization) of mesenchymal subtypes with other broad cell classes in *TP53*^{mut} *vs*. *TP53*^{WT} spatial samples. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

(F) Average $\log_{10}$ expression of specific CAF.ADH1B, CAF.ISG, and CAF.Ribo markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(G) Left: Dot plot of genes involved in interactions displayed in **Figure 4F**. Right: Pseudobulk differential expression of the same genes *TP53*^{mut} versus *TP53*^{WT} patients. Mann-Whitney-Wilcoxon Test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

Supplementary Figure 5. Characterization of myeloid subtypes and expression of ligand-receptor genes.

(A) Stacked bar plot displaying the composition of myeloid subtypes for each patient sample.

(B) Left: Proportion of TAM.CXCLs of all myeloid cells in scRNA-seq. Right: Average $\log_{10}$ expression of markers specific for TAM.CXCLs derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(C) Dot plot of gene set enrichment outputs (Hallmark) of 40 most differentially upregulated or downregulated markers for each myeloid subtype between *TP53*^{mut} or *TP53*^{WT}.

(D) Horizontal bar plot showing top correlated genes with *SPP1* in myeloid cells, ranked by Pearson's correlation coefficient.

(E) Kaplan-Meier survival curve showing association between *SPP1* expression and overall survival in bulk LUAD data from TCGA. Patients (n = 510) were stratified into high or low expression of *SPP*1. Cox proportional-hazards model used to calculate *p*-value corrected for *TP53* mutational status.

(F) Dot plot showing expression of genes involved in interactions shown in **Figure 5E.**

Supplementary Figure 6. Lymphoid immune checkpoint molecule expression and spatial distribution.

(A) Dot plot of T and NK subtype-specific markers.

(B) UMAP of 18,520 B, plasma, and mast cells, integrated by patient and colored by subset.

(C) Average $\log_{10}$ expression of specific T.Exhausted, and CD8.GZMK markers derived from the scRNA-seq data for *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test *p*-values were computed.

(D) Log_{10} expression of *PDCD1, CTLA4,* and *TIGIT* in *TP53*^{WT} (blue) and *TP53*^{mut} (red) LUAD bulk RNA-seq samples from TCGA. Mann-Whitney-Wilcoxon Test used.

(E) Left: Dot plot of genes involved in interactions displayed in **Figure 6E**. Right: Pseudobulk differential expression of the same genes *TP53*^{mut} *vs*. *TP53*^{WT} patients. Mann-Whitney-Wilcoxon Test *p*-values were calculated. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

Supplementary Figure 7. Spatial organization and cellular landmarks in *TP53*^{mut} and *TP53*^{WT} LUAD.

(A) Mean spatial correlation of broad cell classes in *TP53*^{WT} (left) and *TP53*^{mut} (right).

(B) Differential spatial correlation heatmap of lymphoid (B, plasma, T, NK) subsets with other broad cell classes between *TP53*^{mut} and *TP53*^{WT}. Red indicates enrichment in *TP53*^{mut} and blue enrichment in *TP53*^{WT} samples.

(C) Line plot showing proportion of malignant, myeloid, and mesenchymal cells ordered from NMF7 low spots to NMF7 high spots (x-axis).

Supplementary Table 1. Clinical data for patient cohort

Supplementary Table 2. High level cell class markers and subset markers

Methods

Human subjects

Fresh solid primary, treatment-naive tumor tissue was collected using an IRB-approved protocol (IRB protocol 98-063) at DFCI/BWH and MGH. Patients were consented to the study prior to collection. Immediately after resection, the tissue was reviewed by the clinical pathology team and high-quality portions (determined based on tumor content, necrosis, calcification, fat and hemorrhage) were allocated for WES, scRNA-seq, and spatial transcriptomics. The protocol was designed to reduce the time between surgical resection, anatomic pathology review, placement in media, and processing at the Broad Institute. Blood was drawn from the same patients and cryo-preserved at -80 °C for subsequent processing. Patient demographic information, including age, sex, and smoking history, are included in **Table S1**.

Whole-exome sequencing sample processing

DNA was extracted from fresh frozen lung cancer tissue embedded into OCT (TissueTek, Sakura) and from PBMCs isolated from preserved patient-matched blood (AllPrep DNA/RNA extraction kit, Qiagen). Library construction was performed as previously described.[83] In brief, DNA input for shearing was diluted to a final concentration of 20–250 ng in 50 µl of solution. Adapter ligation was performed using palindromic forked adapters (Integrated DNA Technologies), containing unique dual-indexed molecular barcode sequences to improve downstream pooling. End repair/A-tailing, adapter ligation, and library enrichment PCR were carried out using a 96-reaction kit format (Kapa HyperPrep). During solid-phase reversible immobilization (SPRI), a final elution volume of 30 µl was produced to maximize library

concentration, and vortexing was performed to maximize the effluent. Constructed libraries were first pooled and hybridization and capture were performed (Illumina's Nextera Exome Kit) using the recommended protocol from the manufacturer, using a skirted PCR plate to facilitate automation (Agilent Bravo liquid handling system). Library pools then underwent qPCR quantification and libraries were adjusted to a concentration of 2nM. DNA libraries were cluster-amplified using exclusion amplification chemistry in patterned flow cells (Illumina) according to the manufacturer's recommended protocol. Sequencing of flow cells was performed using sequencing-by-synthesis chemistry and analyzed using RTA v.2.7.3 or later. Library pools were sequenced on 76 cycle runs using two 8-cycle index reads across the appropriate number of lanes to attain coverage for all libraries.

Tissue processing, CD45 sorting, and scRNA sequencing

Tumor tissue resections were transferred in RPMI on ice from the operating room and washed in cold PBS in the laboratory, and transferred to 5ml Eppendorf tubes containing 3ml dissociation mixture (NSCLC PDEC protocol).[84] Samples were minced in the Eppendorf tube using scissors into pieces smaller than ~0.4 mm and incubated at 37 °C, rotating at ~14 rpm. for 10 min. Each sample was pipetted 20 times with a 1ml pipette tip at room temperature and placed back into incubation for 10 min. The sample was pipetted again 20 times using a 1 ml pipette tip, transferred to a 1.7 ml Eppendorf tube and centrifuged at 300–580*g* for 4–7 min at 4 °C. The pellet was resuspended in 200–500 µl of ACK (ammonium-chloride-potassium) RBC lysis buffer (Thermo Fisher Scientific, A1049201) and incubated for 1 min on ice. Twice the volume of cold PBS was added to stop the reaction, and cells were pelleted by a short centrifugation for 8 s at 4 °C, using the short spin setting and ramping up to 11,000*g*. If RBCs remained, the RBC lysis

step was performed up to two additional times. To assess cell count and viability, 5 µl of Trypan blue (Thermo Fisher Scientific, cat. no. T10282) was mixed with 5 µl of the sample and loaded onto an INCYTO C-Chip Disposable Hemocytometer, Neubauer Improved (VWR, cat. no. 82030-468). Depletion of $CD45^+$ cells in NSCLC samples was performed using CD45 MicroBeads (Miltenyi Biotec, cat. no. 130-045-801) according to the manufacturer's protocol. The single-cell suspension was centrifuged at 500*g* for 4 min at 4 °C and the pellet was resuspended in 80 µl of MACS buffer (PBS supplemented with 0.5% BSA, and 2 mM EDTA) per 10^6 cells. MACS CD45 microbeads were added to the cell suspension (20 µl per 10 million cells) and the cells were incubated on ice for 15 min. During incubation, a MS column was attached to a MidiMACS separator and rinsed with 3 ml MACS buffer. After incubation, the cells-bead conjugate was washed with 900 µl MACS buffer per 10 million cells. The cells were then centrifuged at 500*g* for 4 min at 4 °C and the pellet was resuspended in 500 µl MACS buffer. The cell suspension was transferred to the column and the effluent was collected ($CD45^-$ fraction). The column was washed three times with 3 ml MACS buffer. The $CD45^-$ fraction was centrifuged at 500*g* for 4 min at 4 °C. The pellet was resuspended in 50 µl of 0.4% BSA (Ambion, cat. no. AM2616) in PBS. Cells were re-counted and adjusted to a final range of 200–2,000 cells per µl.

NSCLC tissue dissociation protocol

The digestion mix contained 2692 µl HBSS (Thermo Fisher Scientific, cat. no. 14170112), 187.5 µl of 20 mg ml^{-1} pronase (Sigma Aldrich, cat. no. 10165921001) diluted to a final concentration of 1,250 µg ml^{-1}, 27.6 µl of 1 mg ml^{-1} elastase (Thermo Fisher Scientific, cat. no. NC9301601) diluted to a final concentration of 9.2 µg ml^{-1}, 30 µl of 10 mg ml^{-1} DNase I (Sigma

Aldrich, cat. no. 11284932001) diluted to a final concentration of 100 μg ml^{-1}, 30 μl of 10 mg ml^{-1} Dispase (Sigma Aldrich, cat. no. 4942078001) diluted to a final concentration of 100 μg ml^{-1}, 30 μl of 150 mg ml^{-1} collagenase A (Sigma Aldrich, cat. no. 10103578001) diluted to a final concentration of 1,500 μg ml^{-1} and 3 μl of 100 mg ml^{-1} collagenase IV (Thermo Fisher Scientific, cat. no. NC9836075) diluted to a final concentration of 100 μg ml^{-1}.

scRNA-seq library preparation and sequencing

A total of 8,000 cells were loaded onto each channel of the Chromium Controller (10X Genomics) to generate single-cell gel bead-in-emulsions (GEMs). Libraries were constructed using the Chromium Single Cell 3′ Library & Gel Bead Kit v.2 or v.3 (PN-120237, 10X Genomics): barcoded reverse transcription of RNA, cDNA amplification, fragmentation and adapter and sample index attachment were all performed according to the manufacturer's recommended protocol. Barcoded libraries from four 10x channels were pooled together and sequenced on either one lane of an Illumina HiSeq X or one flow cell of a NextSeq, with paired end reads accordingly: read 1, 26 nt; read 2, 55 nt; index 1, 8 nt; index 2, 0 nt.

Tissue processing for Spatial Transcriptomics

Fresh frozen lung cancer tissue samples were embedded into OCT (TissueTek, Sakura) and shipped to the Broad Institute for cryosectioning and generation of hematoxylin and eosin-stained (H&E) tissue sections. H&E-stained tissues were subject to pathology review to assess tissue quality, structural preservation, cellular viability, tumor content, inflammation, fibrosis, and necrosis. Samples were excluded based on small tissue size, low cellularity, or extensive fibrosis. Regions of interest (ROIs) were then selected based on the criteria mentioned, and

marked on H&E images for subsequent tissue sectioning and mounting on Visium slides. Visium sectioning and processing were performed on the eight samples that passed QC. Tissues were cryosectioned at 10 um thickness at -22°C and placed in the capture areas of Visium Tissue Optimization slides (3000394, 10X Genomics) and Visium Spatial Gene Expression slides (2000233, 10X Genomics). The tissue sections were adhered by warming the backside of the slides, and placed at -80°C for 1-3 days.

Visium spatial gene expression library generation

The tissue optimization sample slide and spatial gene expression slide were processed according to the manufacturer's protocols (CG000238_VisiumSpatialTissueOptimizationUserGuide_Rev_A.pdf and CG000239_VisiumSpatialGeneExpression_UserGuide_Rev_C.pdf). In short, following tissue methanol fixation and H&E staining, brightfield morphology images were obtained with a Zeiss Axio microscope using the Metafer slide-scanning platform (Metasystems) at 10x resolution. The images were joined together with the VSlide software (Metasystems) and exported as tiff files. Optimal tissue permeabilization times were tested with eight different time points (0, 3, 6, 12, 18, 24, 30, 36 minutes) on the tissue optimization sample slide for one of the samples. 12 minutes of permeabilization was set as the optimal time point and used for the spatial gene expression slide for all samples. The released RNAs from the tissue were reversely transcribed into complimentary DNA by priming to the spatial barcoded primers on the glass in the presence of template switching oligo. After second strand synthesis, a denaturation step released the cDNAs, which were then amplified with 14-15 cycles of PCR amplification. Finally, indexed sequencing-ready, spatial gene expression libraries were constructed according to the

manufacturer's protocol. The libraries were pooled together to generate >50 000 reads per spatial spot and sequenced on Illumina NovaSeq sequencers with S1 and SP kits and the settings: read 1, 28 cycles; read 2, 90 cycles; index 1, 10 cycles; index 2, 10 cycles.

WES data processing

The Picard pipeline (http://picard.sourceforge.net/) was used to align the tumor and normal whole exome sequences to the hg19 reference human genome build and generate BAM files. Somatic short variant discovery was performed using the Mutect 2 algorithm (https://software.broadinstitute.org/gatk/documentation/tooldocs/current/org_broadinstitute_gatk_tools_walkers_cancer_m2_MuTect2) from the GAT4K pipeline run on the Terra cloud-based platform. Somatic mutations were annotated using Funcotator (https://gatk.broadinstitute.org/hc/en-us/articles/360035889931-Funcotator-Information-and-Tutorial). Copy number ratios for each exon was calculated by comparing mean exon coverage with expected coverage based on a panel of normal samples, and were then segmented for downstream analysis. Maf and vcf files were generated as outputs, which were then subsequently processed with maftools[85] and cnvkit[86] packages.

Large-scale CNA inference and correlation

The inferCNV package[87] (https://github.com/broadinstitute/inferCNV) was applied to infer large-scale CNVs from scRNA-seq data. Normal lung epithelial cells from a separate dataset were used as a reference. Raw UMI counts, a gene ordering file, and an annotation file specifying normal and tumor cells and patient sample of origin were used as input. The function infercnv::run() was performed with a cutoff of 0.1, cluster_by_groups = T, and HMM = T. CNV

values were predicted using a 6-state Hidden Markov Model (HMM) and were normalized to the 3rd quantile of all values per sample. Cells with > 0.7 normalized CNV values were assigned as malignant cells. True CNVs called from WES data were used as a positive control to assess the accuracy of the inferCNV analysis. Spearman correlation was run between inferred CNVs from scRNA and called from WES data, and plotted as a heatmap.

scRNA-seq preprocessing

CellRanger v3.1 was used to align 10X Chromium reads, to the GRCh38 human genome reference and generate barcode, gene, and count matrices. Preprocessed matrices were then loaded into Seurat v4[88] in R v4.1.0 using the Read10X() function. Cells with less than 1000 unique molecular identifiers (UMIs), less than 400 detected genes (1000 genes for malignant cells), or greater than 25% mitochondrial genes were excluded from further analyses.

scRNA-seq integration, clustering, annotation, and doublet removal

The UMI counts matrices were batch-corrected by patient sample, using the Seurat v4 SCTransform() (using 2000 variable features and regressing out percentage of mitochondrial genes), FindIntegrationAnchors(), and IntegrateData(), using the largest sample as the reference dataset. PCA was then performed in the variable gene space and 15 principal components were used for graph-based clustering and UMAP dimensionality reduction. Cell type markers for broad cell classes were identified using the FindAllMarkers() function, using a maximum of 1,000 cells per cluster. Cells were annotated based on expression of specific markers that were established by previous literature. Individual cell classes were then subject to another series of

integration, clustering, and annotation based on expression of subtype-specific markers. Doublets were removed based on expression of markers specific to > 1 broad cell classes.

scRNA-seq signature scoring

To identify broad cell classes signatures, we used the FindAllMarkers() function in Seurat v4 and generated gene sets using the most specific genes for each classes, based on $\log_2$ fold change and Bonferroni-corrected *p*-values compared to the other clusters. We identified malignant program signatures by performing Louvain clustering on malignant cells at resolution 1.2 and using the FindAllMarkers() function to derive the top 10 most specific genes from each cluster, based on $\log_2$ fold change and Bonferroni-corrected *p*-values. Signatures for normal lung epithelial cells were derived from published healthy lung scRNA-seq data using the same method. The lists of broad cell classes and cell subtype specific genes, as well as final gene set signatures are provided in **Table S2**. To define hallmark hypoxia and EMT scores, we downloaded the gene sets "HALLMARK_HYPOXIA" "HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION" from MSigDB.[89] Signature scores were assigned per cell in scRNA-seq data using the AddModuleScore() function in Seurat v4, which compares the average expression of the genes in each gene set to the average expression of a control gene set of 100 randomly selected genes within the same expression bin.

scRNA-seq differential expression analysis

The data was first split into different subtypes, and pseudobulk profiles were generated by setting the identity of the objects to patient sample ID and running the AverageExpression() function in Seurat v4. A Mann-Whitney Wilcoxon test was used to test for Bonferroni-corrected significance

in gene differential expression between groups. The average $\log_2$ fold change was plotted using the ComplexHeatmap package in R, and asterisks were added for genes that were significantly differentially expressed.

Gene Set Enrichment Analysis

Using the top 100 DEGs between conditions per subtype or the top 100 specific markers per subtype, we performed gene-set enrichment analysis using the R package clusterProfiler[90] (v4.0), focusing on GO terms, KEGG, Reactome, and Hallmark Cancer gene sets. Gene ratios and adjusted *p*-values were visualized as dot plots, after applying BH correction and setting *p*-value and *q*-value cutoffs of 0.05.

Annotation using normal lung reference dataset

To identify transcriptomic similarities between cell subtypes in NSCLC and the normal lung, the Seurat v4 function TransferData() was used. The broad cell compartment object in the NSCLC scRNA-seq dataset was assigned to the query parameter and the corresponding compartment in a published normal lung cell atlas[49] was assigned to the reference parameter, and the remaining parameters were set to defaults. Each cell in the NSCLC object was assigned a predicted identity based on the normal lung subtype it was most similar to transcriptionally. We created a confusion matrix displaying the number of cells that overlapped between predicted identities from the normal lung and annotations assigned de novo.

Signaling Entropy calculation

Entropy was inferred for cell from the epithelial classes using the R package SCENT,[91] using the default protein-protein interaction (PPI) network ("net17Jan2016.m"). Raw UMI counts were used as the scRNA-seq input. Entropy calculation was done separately for malignant and non-malignant epithelial cells. Quality control was performed using the scater package in R, and low-quality cells based on library size, number of expressed genes, and mitochondrial gene proportion were removed. Differential potency estimation was estimated using the DoIntegPPI() function, which provided SR (Signaling Entropy Rate) as output. The same pipeline was used for bulk RNA data from TCGA. Tumor samples and adjacent normal samples were run separately. SR values were compared using the ggviolin() function using the ggpubr package in R.

Signature scoring of bulk expression profiles

Data from TCGA LUAD and somatic mutations were downloaded from cbioPortal[92] using the cgdsr package in R. Curated gene sets with highly specific markers for individual cell types, subtypes, or malignant programs were generated from the scRNA-seq data. The signature scores for these gene sets were defined as the mean $\log_{10}$ expression of all the genes in the gene set, and were used to estimate the relative frequency of cell types or subtypes, as well as average expression of gene signatures in TCGA LUAD.

Survival analysis

To predict the roles of individual genes and gene signatures on patient survival, TCGA LUAD patients were classified into high (top 25%), medium (top 50%), and low (bottom 25%); or high (top 50%) and low (bottom 50%) expression groups. Kaplan-Meier (KM) curves were generated using the survival package in R, and bivariate Cox proportional-hazards regression analysis was

used to calculate the significance of the contribution of different model parameters to overall survival (defined as time to death from any cause).

Cell-cell interaction analysis

CellPhoneDB[93] (v2.1.7) was used to predict significant ligand-receptor interactions using normalized raw counts and annotation of broad cell classes as input. The method was run with default parameters separately for each sample ID to account for batch effects between samples. Interactions with a *p*-value below 0.05 were classified as significant. The significant means output was converted into a binary matrix based on presence or absence of a significant interaction. Enrichment of a specific interaction was calculated as the proportion of patients with the interaction in group A minus the proportion of patients with the interaction in group B. Significance of enrichment for an interaction was calculated using a Fisher's exact test. Negative $\log_{10}$ *p*-values and enrichment scores for each interaction-cell pair combination were plotted using the ggplot2 package in R.

scRNA-seq regulon analysis

A scRNA-seq transcription factor (TF) gene regulatory network was constructed using pySCENIC[55] (v0.11.2). Raw UMI counts were used as input and regulons were predicted using gene module co-expression using the GRNBoost2 package. The number of regulons was pruned using the feather ranking databases (hg19-500bp-upstream-7species.mc9nr, hg19-tss-centered-10kb-10species.mc9nr, hg19-tss-centered-5kb-10species.mc9nr). Regulon activity, driving TFs, and weights for individual target genes were predicted with the cisTarget() function in pySCENIC. Regulon activity enrichment scores were predicted for each cellular subtype using

AUCell and regulon specificity scores were used to identify regulatory networks specific to each subtype. Z-score-transformed regulon x subtype matrices were visualized using the ComplexHeatmap package in R.

Pathologist annotation of hematoxylin and eosin (H&E) stains from spatial transcriptomics samples

A board-certified pathologist, blinded to all sample identities and spatial transcriptomics measurements, performed manual annotation of 13 H&E stains using Loupe browser. Annotations on the H&E were classified into multiple categories (*e.g.* lepidic adenocarcinoma, TLS/BALT, vascular endothelium) that were most representative of cells within barcoded spots. Spots in which there were cell types whose identities could not be readily visually resolved were not annotated. These categories were then mapped to three general cell classes (malignant, stromal, and immune) to minimize variability in pathology annotations across samples and allow for comparison with the computationally inferred cellular compositions.

Processing of spatial transcriptomics data, integration with paired scRNA-seq data, and colocalization analysis

Tangram version 0.4.0, scanpy 1.8.1, and anndata 0.7.6 was used to map scRNA-seq data to spatial locations on 10X Visium spatial transcriptomics samples as described in.[34] UMI counts for scRNA-seq and spatial transcriptomics data were converted into anndata objects and pre-processed using standard scanpy functions. The top 100 genes that best characterized each broad cell classes were selected by sc.tl.rank_genes_groups(). This subset of genes was used afterwards as training genes for Tangram alignment. In order to more accurately map single-cells onto

spatial transcriptomics spots, an estimate of cell density per spot was attained through watershed-based nuclei segmentation of the hematoxylin and eosin-stained tissue section. Images were first imported by squidpy 1.1.0, and smoothened using squidpy.im.process with a sigma = 4, then segmented by squidpy.im.segment with method = "watershed", using parameters channel = 0 to select the red color channel and thres = 120. A threshold of 120 was selected to best separate nuclei from background, and sq.im.calculate_features was used to count the number of segmented nuclei per spot after segmentation. Finally, mapping was performed by tg.map_cells_to_space using mode = "constrained" and density_prior = # of nuclei per spot / total # of nuclei to constrain alignment based on segmented nuclei density, returning probabilities of spatial location on a per-cell basis which were employed in later analysis. To project individual gene expression onto spatial transcriptomics spots, tg.project_genes was used. Cell type and subtype proportion were predicted per spot by summing up cell x spot probabilities and normalizing the total probability scores per spot to 1. Spatial colocalization was inferred using Pearson correlation between predicted cell type and subtype compositions, and additional continuous attributes from the metadata across spots.

NMF analysis for spatial transcriptomics data

NMF analysis was adapted from cell2location[94] using subtype proportions per spatial transcriptomics spot as inputs, enabling analysis of spatially colocalized cellular programs. NMF was trained five times for a range of k = 3 to 30 factors, and k = 14 was selected based on stability of training across the five restarts and elucidating discrete spatially-delimited compartments across spatial transcriptomics samples. Using scanpy, NMF weights were plotted

onto spatial transcriptomics data to investigate spatial-dependent patterning of NMF factors. NMF weights per subtype were visualized using the heatmap function in cell2location.

REFERENCES

1. Travis, W.D., Brambilla, E., Nicholson, A.G., Yatabe, Y., Austin, J.H.M., Beasley, M.B., Chirieac, L.R., Dacic, S., Duhig, E., Flieder, D.B., et al. (2015). The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol *10*, 1243-1260. 10.1097/JTO.0000000000000630.
2. Cancer Genome Atlas Research, N. (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519-525. 10.1038/nature11404.
3. Cancer Genome Atlas Research, N. (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature *511*, 543-550. 10.1038/nature13385.
4. Hanna, N.H., Robinson, A.G., Temin, S., Baker, S., Jr., Brahmer, J.R., Ellis, P.M., Gaspar, L.E., Haddad, R.Y., Hesketh, P.J., Jain, D., et al. (2021). Therapy for Stage IV Non-Small-Cell Lung Cancer With Driver Alterations: ASCO and OH (CCO) Joint Guideline Update. J Clin Oncol *39*, 1040-1091. 10.1200/JCO.20.03570.
5. Howlader, N., Forjaz, G., Mooradian, M.J., Meza, R., Kong, C.Y., Cronin, K.A., Mariotto, A.B., Lowy, D.R., and Feuer, E.J. (2020). The Effect of Advances in Lung-Cancer Treatment on Population Mortality. N Engl J Med *383*, 640-649. 10.1056/NEJMoa1916623.
6. Herbst, R.S., Morgensztern, D., and Boshoff, C. (2018). The biology and management of non-small cell lung cancer. Nature *553*, 446-454. 10.1038/nature25183.
7. National Cancer Institute (2021). Non-Small Cell Lung Cancer Treatment. https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq.

8. Campbell, J.D., Alexandrov, A., Kim, J., Wala, J., Berger, A.H., Pedamallu, C.S., Shukla, S.A., Guo, G., Brooks, A.N., Murray, B.A., et al. (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet *48*, 607-616. 10.1038/ng.3564.

9. Halvorsen, A.R., Silwal-Pandit, L., Meza-Zepeda, L.A., Vodak, D., Vu, P., Sagerup, C., Hovig, E., Myklebost, O., Borresen-Dale, A.L., Brustugun, O.T., and Helland, A. (2016). TP53 Mutation Spectrum in Smokers and Never Smoking Lung Cancer Patients. Front Genet *7*, 85. 10.3389/fgene.2016.00085.

10. Kaubryte, J., and Lai, A.G. (2022). Pan-cancer prognostic genetic mutations and clinicopathological factors associated with survival outcomes: a systematic review. NPJ Precis Oncol *6*, 27. 10.1038/s41698-022-00269-5.

11. Duffy, M.J., Synnott, N.C., O'Grady, S., and Crown, J. (2020). Targeting p53 for the treatment of cancer. Semin Cancer Biol. 10.1016/j.semcancer.2020.07.005.

12. Bykov, V.J.N., Eriksson, S.E., Bianchi, J., and Wiman, K.G. (2018). Targeting mutant p53 for efficient cancer therapy. Nat Rev Cancer *18*, 89-102. 10.1038/nrc.2017.109.

13. Zhu, G., Pan, C., Bei, J.X., Li, B., Liang, C., Xu, Y., and Fu, X. (2020). Mutant p53 in Cancer Progression and Targeted Therapies. Front Oncol *10*, 595187. 10.3389/fonc.2020.595187.

14. Uehara, I., and Tanaka, N. (2018). Role of p53 in the Regulation of the Inflammatory Tumor Microenvironment and Tumor Suppression. Cancers *10*. ARTN 219 10.3390/cancers10070219.

15. Shi, D., and Jiang, P. (2021). A Different Facet of p53 Function: Regulation of Immunity and Inflammation During Tumor Development. Front Cell Dev Biol *9*, 762651. 10.3389/fcell.2021.762651.

16. Skoulidis, F., Goldberg, M.E., Greenawalt, D.M., Hellmann, M.D., Awad, M.M., Gainor, J.F., Schrock, A.B., Hartmaier, R.J., Trabucco, S.E., Gay, L., et al. (2018). STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. Cancer Discov *8*, 822-835. 10.1158/2159-8290.CD-18-0099.

17. Biton, J., Mansuet-Lupo, A., Pecuchet, N., Alifano, M., Ouakrim, H., Arrondeau, J., Boudou-Rouquette, P., Goldwasser, F., Leroy, K., Goc, J., et al. (2018). TP53, STK11, and EGFR Mutations Predict Tumor Immune Profile and the Response to Anti-PD-1 in Lung Adenocarcinoma. Clin Cancer Res *24*, 5710-5723. 10.1158/1078-0432.CCR-18-0163.

18. Dong, Z.Y., Zhong, W.Z., Zhang, X.C., Su, J., Xie, Z., Liu, S.Y., Tu, H.Y., Chen, H.J., Sun, Y.L., Zhou, Q., et al. (2017). Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. Clin Cancer Res *23*, 3012-3024. 10.1158/1078-0432.CCR-16-2554.

19. Leader, A.M., Grout, J.A., Maier, B.B., Nabet, B.Y., Park, M.D., Tabachnikova, A., Chang, C., Walker, L., Lansky, A., Le Berichel, J., et al. (2021). Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. Cancer Cell *39*, 1594-1609 e1512. 10.1016/j.ccell.2021.10.009.

20. Davis, A.P., Cooper, W.A., Boyer, M., Lee, J.H., Pavlakis, N., and Kao, S.C. (2021). Efficacy of immunotherapy in KRAS-mutant non-small-cell lung cancer with comutations. Immunotherapy-Uk *13*, 941-952. 10.2217/imt-2021-0090.

21. Scalera, S., Mazzotta, M., Corleone, G., Sperati, F., Terrenato, I., Krasniqi, E., Pizzuti, L., Barba, M., Vici, P., Gallo, E., et al. (2021). KEAP1 and TP53 Frame Genomic, Evolutionary, and Immunologic Subtypes of Lung Adenocarcinoma With Different Sensitivity to Immunotherapy. J Thorac Oncol *16*, 2065-2077. 10.1016/j.jtho.2021.08.010.

22. Lavin, Y., Kobayashi, S., Leader, A., Amir, E.D., Elefant, N., Bigenwald, C., Remark, R., Sweeney, R., Becker, C.D., Levine, J.H., et al. (2017). Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. Cell *169*, 750-765 e717. 10.1016/j.cell.2017.04.014.

23. Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwe, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat Med *24*, 1277-1289. 10.1038/s41591-018-0096-5.

24. Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity *50*, 1317-1334 e1310. 10.1016/j.immuni.2019.03.009.

25. Zhang, L., Zhang, Y., Wang, C., Yang, Y., Ni, Y., Wang, Z., Song, T., Yao, M., Liu, Z., Chao, N., et al. (2022). Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. Signal Transduct Target Ther *7*, 9. 10.1038/s41392-021-00824-9.

26. Kim, N., Kim, H.K., Lee, K., Hong, Y., Cho, J.H., Choi, J.W., Lee, J.I., Suh, Y.L., Ku, B.M., Eum, H.H., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun *11*, 2285. 10.1038/s41467-020-16164-1.

27. Laughney, A.M., Hu, J., Campbell, N.R., Bakhoum, S.F., Setty, M., Lavallee, V.P., Xie, Y., Masilionis, I., Carr, A.J., Kottapalli, S., et al. (2020). Regenerative lineages and immune-mediated pruning in lung cancer metastasis. Nat Med *26*, 259-269. 10.1038/s41591-019-0750-6.

28. Bischoff, P., Trinks, A., Obermayer, B., Pett, J.P., Wiederspahn, J., Uhlitz, F., Liang, X., Lehmann, A., Jurmeister, P., Elsner, A., et al. (2021). Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. Oncogene *40*, 6748-6758. 10.1038/s41388-021-02054-3.

29. He, D., Wang, D., Lu, P., Yang, N., Xue, Z., Zhu, X., Zhang, P., and Fan, G. (2021). Single-cell RNA sequencing reveals heterogeneous tumor and immune cell populations in early-stage lung adenocarcinomas harboring EGFR mutations. Oncogene *40*, 355-368. 10.1038/s41388-020-01528-0.

30. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902 e1821. 10.1016/j.cell.2019.05.031.

31. Wang, P., Wang, F., He, H., Chen, Y., Lin, H., Chen, P., Chen, X., and Liu, S. (2021). TP53 and CDKN2A mutations in patients with early-stage lung squamous cell carcinoma: an analysis of the correlations and prognostic outcomes. Ann Transl Med *9*, 1330. 10.21037/atm-21-3709.

32. Dudley, A.C. (2012). Tumor endothelial cells. Cold Spring Harb Perspect Med *2*, a006536. 10.1101/cshperspect.a006536.

33. Cooke, V.G., LeBleu, V.S., Keskin, D., Khan, Z., O'Connell, J.T., Teng, Y., Duncan, M.B., Xie, L., Maeda, G., Vong, S., et al. (2012). Pericyte depletion results in hypoxia-associated epithelial-to-mesenchymal transition and metastasis mediated by met signaling pathway. Cancer Cell *21*, 66-81. 10.1016/j.ccr.2011.11.024.

34. Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, A., Zhang, M., et al. (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat Methods *18*, 1352-1362. 10.1038/s41592-021-01264-7.

35. Schupp, J.C., Adams, T.S., Cosme, C., Jr., Raredon, M.S.B., Yuan, Y., Omote, N., Poli, S., Chioccioli, M., Rose, K.A., Manning, E.P., et al. (2021). Integrated Single-Cell Atlas of Endothelial Cells of the Human Lung. Circulation *144*, 286-302. 10.1161/CIRCULATIONAHA.120.052318.

36. Gillich, A., Zhang, F., Farmer, C.G., Travaglini, K.J., Tan, S.Y., Gu, M., Zhou, B., Feinstein, J.A., Krasnow, M.A., and Metzger, R.J. (2020). Capillary cell-type specialization in the alveolus. Nature *586*, 785-789. 10.1038/s41586-020-2822-7.

37. Muz, B., de la Puente, P., Azab, F., and Azab, A.K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. Hypoxia (Auckl) *3*, 83-92. 10.2147/HP.S93413.

38. Acevedo, L.M., Barillas, S., Weis, S.M., Gothert, J.R., and Cheresh, D.A. (2008). Semaphorin 3A suppresses VEGF-mediated angiogenesis yet acts as a vascular permeability factor. Blood *111*, 2674-2680. 10.1182/blood-2007-08-110205.

39. Guttmann-Raviv, N., Shraga-Heled, N., Varshavsky, A., Guimaraes-Sternberg, C., Kessler, O., and Neufeld, G. (2007). Semaphorin-3A and semaphorin-3F work together to repel endothelial cells and to inhibit their survival by induction of apoptosis. J Biol Chem *282*, 26294-26305. 10.1074/jbc.M609711200.

40. Pasquale, E.B. (2010). Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nat Rev Cancer *10*, 165-180. 10.1038/nrc2806.

41. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189-196. 10.1126/science.aad0501.

42. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396-1401. 10.1126/science.1254257.

43. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology *36*, 411.

44. Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White, R.M., and Yanai, I. (2020). The Stress-Like Cancer Cell State Is a Consistent Component of Tumorigenesis. Cell Syst *11*, 536-546 e537. 10.1016/j.cels.2020.08.018.

45. Kinker, G.S., Greenwald, A.C., Tal, R., Orlova, Z., Cuoco, M.S., McFarland, J.M., Warren, A., Rodman, C., Roth, J.A., Bender, S.A., et al. (2020). Pan-cancer single-cell

RNA-seq identifies recurring programs of cellular heterogeneity. Nat Genet *52*, 1208-1218. 10.1038/s41588-020-00726-6.

46. Min, J.W., Kim, W.J., Han, J.A., Jung, Y.J., Kim, K.T., Park, W.Y., Lee, H.O., and Choi, S.S. (2015). Identification of Distinct Tumor Subpopulations in Lung Adenocarcinoma via Single-Cell RNA-seq. PLoS One *10*, e0135817. 10.1371/journal.pone.0135817.

47. Tang, H., Wang, S., Xiao, G., Schiller, J., Papadimitrakopoulou, V., Minna, J., Wistuba, II, and Xie, Y. (2017). Comprehensive evaluation of published gene expression prognostic signatures for biomarker-based lung cancer clinical studies. Ann Oncol *28*, 733-740. 10.1093/annonc/mdw683.

48. Nagy, A., Munkacsy, G., and Gyorffy, B. (2021). Pancancer survival analysis of cancer hallmark genes. Sci Rep *11*, 6047. 10.1038/s41598-021-84787-5.

49. Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature *587*, 619-625. 10.1038/s41586-020-2922-4.

50. Ferone, G., Lee, M.C., Sage, J., and Berns, A. (2020). Cells of origin of lung cancers: lessons from mouse studies. Genes Dev *34*, 1017-1032. 10.1101/gad.338228.120.

51. Marjanovic, N.D., Hofree, M., Chan, J.E., Canner, D., Wu, K., Trakala, M., Hartmann, G.G., Smith, O.C., Kim, J.Y., Evans, K.V., et al. (2020). Emergence of a High-Plasticity Cell State during Lung Cancer Evolution. Cancer Cell *38*, 229-+. 10.1016/j.ccell.2020.06.012.

52. Teschendorff, A.E., and Enver, T. (2017). Single-cell entropy for accurate estimation of differentiation potency from a cell's transcriptome. Nat Commun *8*, 15599. 10.1038/ncomms15599.

53. Irvine, A.F., Waise, S., Green, E.W., Stuart, B., and Thomas, G.J. (2021). Characterising cancer-associated fibroblast heterogeneity in non-small cell lung cancer: a systematic review and meta-analysis. Sci Rep *11*, 3727. 10.1038/s41598-021-81796-2.

54. Buechler, M.B., Pradhan, R.N., Krishnamurty, A.T., Cox, C., Calviello, A.K., Wang, A.W., Yang, Y.A., Tam, L., Caothien, R., Roose-Girma, M., et al. (2021). Cross-tissue organization of the fibroblast lineage. Nature *593*, 575-579. 10.1038/s41586-021-03549-5.

55. Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc *15*, 2247-2276. 10.1038/s41596-020-0336-2.

56. Lee, K.W., Yeo, S.Y., Sung, C.O., and Kim, S.H. (2015). Twist1 is a key regulator of cancer-associated fibroblasts. Cancer Res *75*, 73-85. 10.1158/0008-5472.CAN-14-0350.

57. De Wever, O., Demetter, P., Mareel, M., and Bracke, M. (2008). Stromal myofibroblasts are drivers of invasive cancer growth. Int J Cancer *123*, 2229-2238. 10.1002/ijc.23925.

58. Engblom, C., Pfirschke, C., and Pittet, M.J. (2016). The role of myeloid cells in cancer therapies. Nat Rev Cancer *16*, 447-462. 10.1038/nrc.2016.54.

59. Dong, B., Wu, C., Huang, L., and Qi, Y. (2021). Macrophage-Related SPP1 as a Potential Biomarker for Early Lymph Node Metastasis in Lung Adenocarcinoma. Front Cell Dev Biol *9*, 739358. 10.3389/fcell.2021.739358.

60. Casazza, A., Laoui, D., Wenes, M., Rizzolio, S., Bassani, N., Mambretti, M., Deschoemaeker, S., Van Ginderachter, J.A., Tamagnone, L., and Mazzone, M. (2013). Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling

blockade inhibits angiogenesis and restores antitumor immunity. Cancer Cell *24*, 695-709. 10.1016/j.ccr.2013.11.007.

61. Zhang, F., Wang, H., Wang, X., Jiang, G., Liu, H., Zhang, G., Wang, H., Fang, R., Bu, X., Cai, S., and Du, J. (2016). TGF-beta induces M2-like macrophage polarization via SNAIL-mediated suppression of a pro-inflammatory phenotype. Oncotarget *7*, 52294-52306. 10.18632/oncotarget.10561.

62. van der Leun, A.M., Thommen, D.S., and Schumacher, T.N. (2020). CD8(+) T cell states in human cancer: insights from single-cell analysis. Nat Rev Cancer *20*, 218-232. 10.1038/s41568-019-0235-4.

63. Litchfield, K., Reading, J.L., Puttick, C., Thakkar, K., Abbosh, C., Bentham, R., Watkins, T.B.K., Rosenthal, R., Biswas, D., Rowan, A., et al. (2021). Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell *184*, 596-614 e514. 10.1016/j.cell.2021.01.002.

64. Hollern, D.P., Xu, N., Thennavan, A., Glodowski, C., Garcia-Recio, S., Mott, K.R., He, X., Garay, J.P., Carey-Ewend, K., Marron, D., et al. (2019). B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. Cell *179*, 1191-1206 e1121. 10.1016/j.cell.2019.10.028.

65. Gide, T.N., Quek, C., Menzies, A.M., Tasker, A.T., Shang, P., Holst, J., Madore, J., Lim, S.Y., Velickovic, R., Wongchenko, M., et al. (2019). Distinct Immune Cell Populations Define Response to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy. Cancer Cell *35*, 238-255 e236. 10.1016/j.ccell.2019.01.003.

66. Thommen, D.S., and Schumacher, T.N. (2018). T Cell Dysfunction in Cancer. Cancer Cell *33*, 547-562. 10.1016/j.ccell.2018.03.012.

67. Upadhaya, S., Hubbard-Lucey, V.M., and Yu, J.X. (2020). Immuno-oncology drug development forges on despite COVID-19. Nature reviews. Drug Discovery.

68. Chen, J., Tan, Y., Sun, F., Hou, L., Zhang, C., Ge, T., Yu, H., Wu, C., Zhu, Y., Duan, L., et al. (2020). Single-cell transcriptome and antigen-immunoglobin analysis reveals the diversity of B cells in non-small cell lung cancer. Genome Biol *21*, 152. 10.1186/s13059-020-02064-6.

69. Yu, Y., Xiao, C.H., Tan, L.D., Wang, Q.S., Li, X.Q., and Feng, Y.M. (2014). Cancer-associated fibroblasts induce epithelial-mesenchymal transition of breast cancer cells through paracrine TGF-beta signalling. Br J Cancer *110*, 724-732. 10.1038/bjc.2013.768.

70. Bonde, A.K., Tischler, V., Kumar, S., Soltermann, A., and Schwendener, R.A. (2012). Intratumoral macrophages contribute to epithelial-mesenchymal transition in solid tumors. Bmc Cancer *12*, 35. 10.1186/1471-2407-12-35.

71. Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., Newman, J., Reczek, E.E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature *445*, 661-665. 10.1038/nature05541.

72. Tam, S.Y., Wu, V.W.C., and Law, H.K.W. (2020). Hypoxia-Induced Epithelial-Mesenchymal Transition in Cancers: HIF-1alpha and Beyond. Front Oncol *10*, 486. 10.3389/fonc.2020.00486.

73. Joseph, J.P., Harishankar, M.K., Pillai, A.A., and Devi, A. (2018). Hypoxia induced EMT: A review on the mechanism of tumor progression and metastasis in OSCC. Oral Oncol *80*, 23-32. 10.1016/j.oraloncology.2018.03.004.

74. Graeber, T.G., Osmanian, C., Jacks, T., Housman, D.E., Koch, C.J., Lowe, S.W., and Giaccia, A.J. (1996). Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours. Nature *379*, 88-91. 10.1038/379088a0.

75. Tang, Q., Su, Z., Gu, W., and Rustgi, A.K. (2020). Mutant p53 on the Path to Metastasis. Trends Cancer *6*, 62-73. 10.1016/j.trecan.2019.11.004.

76. Capaci, V., Mantovani, F., and Del Sal, G. (2020). Amplifying Tumor-Stroma Communication: An Emerging Oncogenic Function of Mutant p53. Front Oncol *10*, 614230. 10.3389/fonc.2020.614230.

77. Pelka, K., Hofree, M., Chen, J.H., Sarkizova, S., Pirl, J.D., Jorgji, V., Bejnood, A., Dionne, D., Ge, W.H., Xu, K.H., et al. (2021). Spatially organized multicellular immune hubs in human colorectal cancer. Cell *184*, 4734-4752 e4720. 10.1016/j.cell.2021.08.003.

78. Li, L., Li M, Wang, X (2020). Cancer type-dependent correlations between TP53 mutations and antitumor immunity. DNA Repair *88*. https://doi.org/10.1016/j.dnarep.2020.102785.

79. Cortez, M.A., Ivan, C., Valdecanas, D., Wang, X., Peltier, H.J., Ye, Y., Araujo, L., Carbone, D.P., Shilo, K., Giri, D.K., et al. (2016). PDL1 Regulation by p53 via miR-34. J Natl Cancer Inst *108*. 10.1093/jnci/djv303.

80. Skoulidis, F., Byers, L.A., Diao, L., Papadimitrakopoulou, V.A., Tong, P., Izzo, J., Behrens, C., Kadara, H., Parra, E.R., Canales, J.R., et al. (2015). Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. Cancer Discov *5*, 860-877. 10.1158/2159-8290.CD-14-1236.

81. Ricciuti, B., Arbour, K.C., Lin, J.J., Vajdi, A., Vokes, N., Hong, L., Zhang, J., Tolstorukov, M.Y., Li, Y.Y., Spurr, L.F., et al. (2022). Diminished Efficacy of Programmed Death-(Ligand)1 Inhibition in STK11- and KEAP1-Mutant Lung Adenocarcinoma Is Affected by KRAS Mutation Status. J Thorac Oncol *17*, 399-410. 10.1016/j.jtho.2021.10.013.

82. Forde, P.M., Spicer, J., Lu, S., Provencio, M., Mitsudomi, T., Awad, M.M., Felip, E., Broderick, S.R., Brahmer, J.R., Swanson, S.J., et al. (2022). Neoadjuvant Nivolumab plus Chemotherapy in Resectable Lung Cancer. N Engl J Med. 10.1056/NEJMoa2202170.

83. Jerby-Arnon, L., Neftel, C., Shore, M.E., Weisman, H.R., Mathewson, N.D., McBride, M.J., Haas, B., Izar, B., Volorio, A., Boulay, G., et al. (2021). Opposing immune and genetic mechanisms shape oncogenic programs in synovial sarcoma. Nat Med *27*, 289-300. 10.1038/s41591-020-01212-6.

84. Slyper, M., Porter, C.B.M., Ashenberg, O., Waldman, J., Drokhlyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med *26*, 792-802. 10.1038/s41591-020-0844-1.

85. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res *28*, 1747-1756. 10.1101/gr.239244.118.

86. Talevich, E., Shain, A.H., Botton, T., and Bastian, B.C. (2016). CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol *12*, e1004873. 10.1371/journal.pcbi.1004873.

87. inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV.

88. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587 e3529. 10.1016/j.cell.2021.04.048.

89. Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425. 10.1016/j.cels.2015.12.004.

90. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (N Y) *2*, 100141. 10.1016/j.xinn.2021.100141.

91. Chen, W., Morabito, S.J., Kessenbrock, K., Enver, T., Meyer, K.B., and Teschendorff, A.E. (2019). Single-cell landscape in mammary epithelium reveals bipotent-like cells associated with breast cancer risk and outcome. Commun Biol *2*, 306. 10.1038/s42003-019-0554-8.

92. Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal *6*, pl1. 10.1126/scisignal.2004088.

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

94. Kleshchevnikov, V., Shmatko, A., Dann, E., Aivazidis, A., King, H.W., Li, T., Elmentaite, R., Lomakin, A., Kedlian, V., Gayoso, A., et al. (2022). Cell2location maps

fine-grained cell types in spatial transcriptomics. Nat Biotechnol. 10.1038/s41587-021-01139-4.

Figure 1

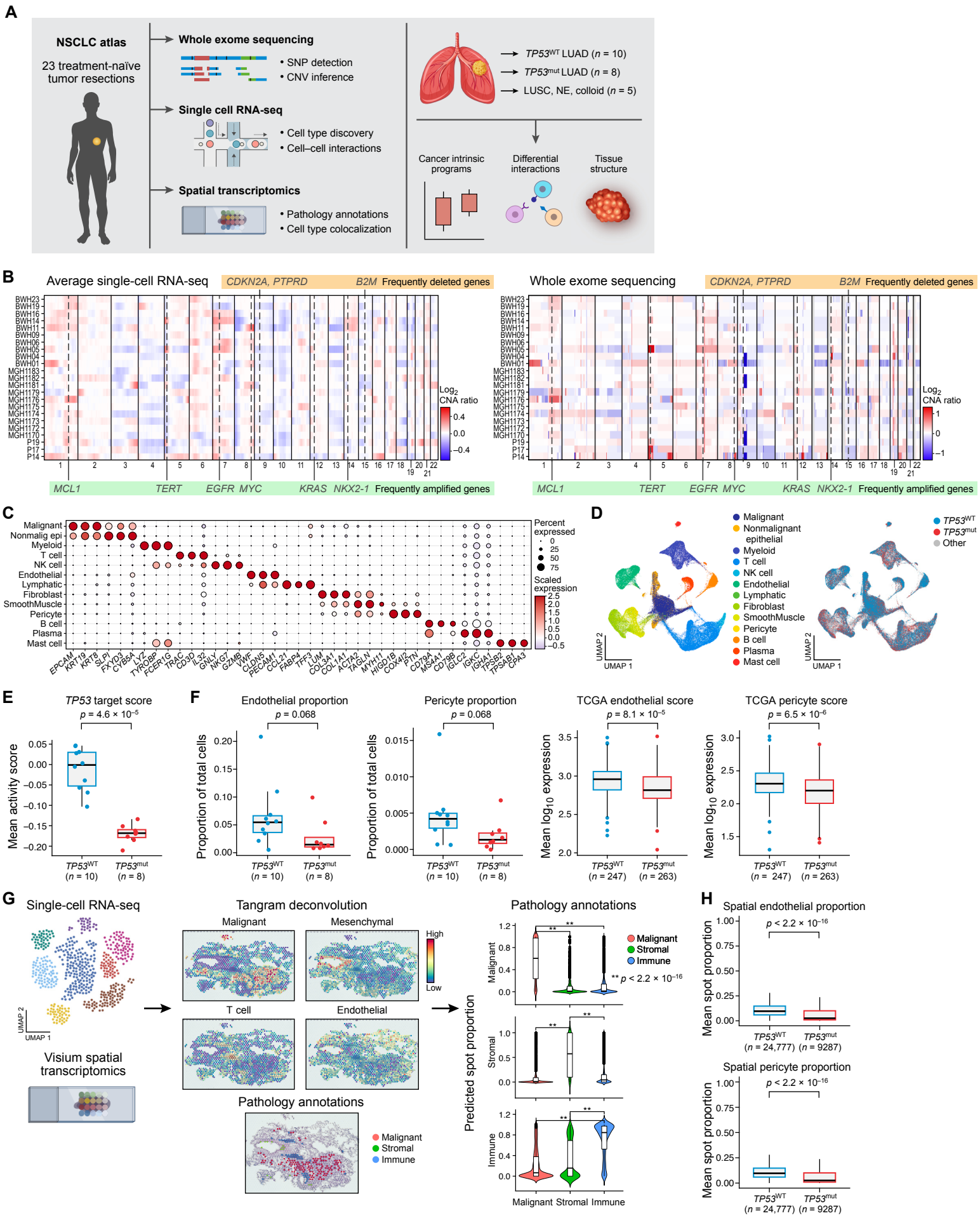

A
NSCLC atlas
23 treatment-naïve tumor resections
Whole exome sequencing
• SNP detection
• CNV inference
Single cell RNA-seq
• Cell type discovery
• Cell–cell interactions
Spatial transcriptomics
• Pathology annotations
• Cell type colocalization
TP53WT LUAD (n = 10)
TP53mut LUAD (n = 8)
LUSC, NE, colloid (n = 5)
Cancer intrinsic programs
Differential interactions
Tissue structure
B
Average single-cell RNA-seq
Whole exome sequencing
CDKN2A, PTPRD
B2M Frequently deleted genes
MCL1
TERT
EGFR MYC
KRAS NKX2-1 Frequently amplified genes
Log2 CNA ratio
C
Malignant
Nonmalig epi
Myeloid
T cell
NK cell
Endothelial
Lymphatic
Fibroblast
SmoothMuscle
Pericyte
B cell
Plasma
Mast cell
Percent expressed
Scaled expression
D
Malignant
Nonmalignant epithelial
Myeloid
T cell
NK cell
Endothelial
Lymphatic
Fibroblast
SmoothMuscle
Pericyte
B cell
Plasma
Mast cell
TP53WT
TP53mut
Other
UMAP 1
UMAP 2
E
TP53 target score
p = 4.6 × 10−5
Mean activity score
TP53WT (n = 10)
TP53mut (n = 8)
F
Endothelial proportion
p = 0.068
Pericyte proportion
p = 0.068
TCGA endothelial score
p = 8.1 × 10−5
TCGA pericyte score
p = 6.5 × 10−6
Proportion of total cells
Mean log10 expression
TP53WT (n = 247)
TP53mut (n = 263)
G
Single-cell RNA-seq
Visium spatial transcriptomics
Tangram deconvolution
Malignant
Mesenchymal
T cell
Endothelial
High
Low
Pathology annotations
Malignant
Stromal
Immune
Predicted spot proportion
** p < 2.2 × 10−16
H
Spatial endothelial proportion
p < 2.2 × 10−16
Spatial pericyte proportion
p < 2.2 × 10−16
Mean spot proportion
TP53WT (n = 24,777)
TP53mut (n = 9287)

Figure 2

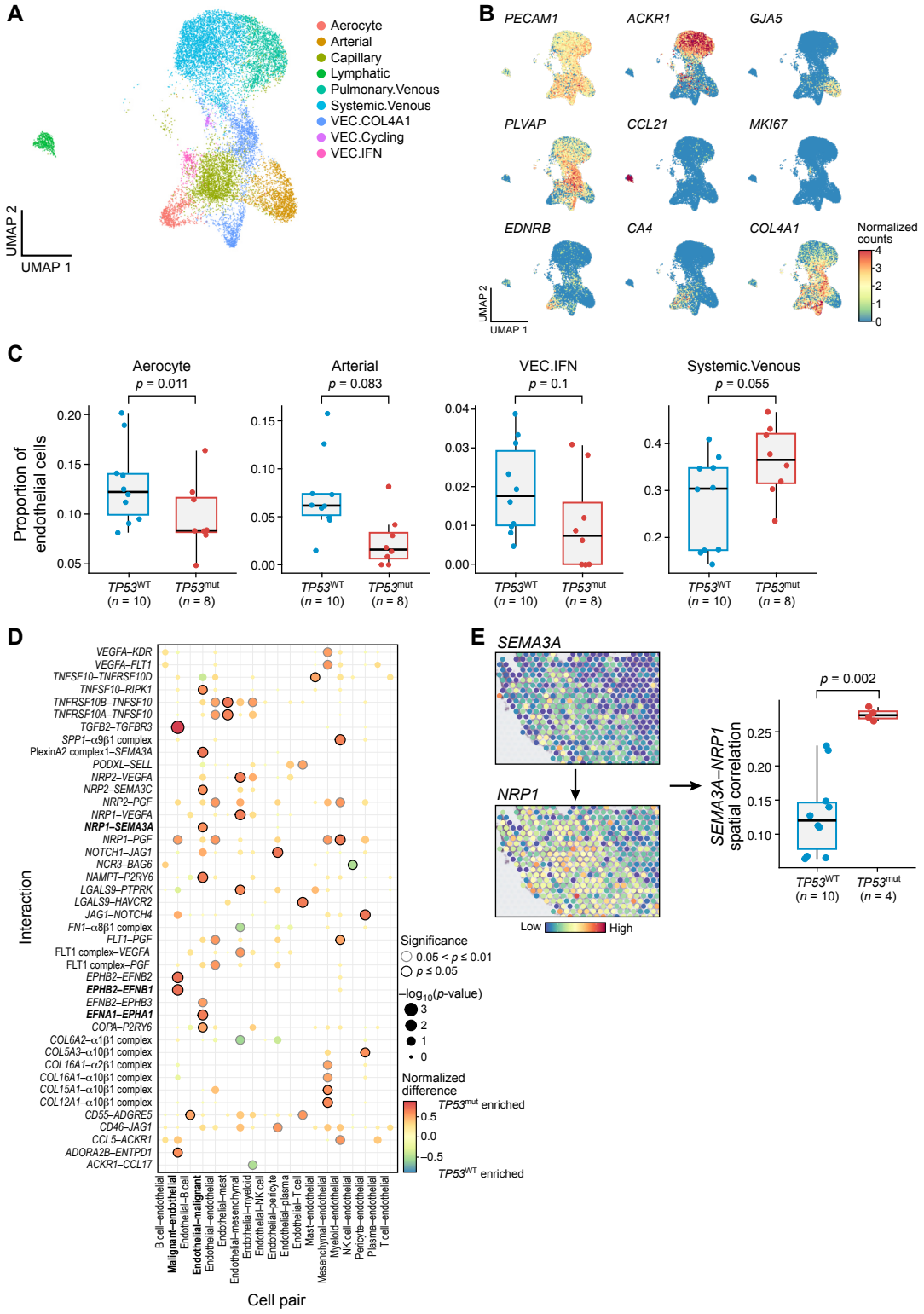

A
Aerocyte
Arterial
Capillary
Lymphatic
Pulmonary.Venous
Systemic.Venous
VEC.COL4A1
VEC.Cycling
VEC.IFN
UMAP 2
UMAP 1
B
PECAM1
ACKR1
GJA5
PLVAP
CCL21
MKI67
EDNRB
CA4
COL4A1
Normalized counts
C
Aerocyte
p = 0.011
Arterial
p = 0.083
VEC.IFN
p = 0.1
Systemic.Venous
p = 0.055
Proportion of endothelial cells
TP53WT (n = 10)
TP53mut (n = 8)
D
Interaction
Cell pair
Significance
0.05 < p ≤ 0.01
p ≤ 0.05
−log10(p-value)
Normalized difference
TP53mut enriched
TP53WT enriched
E
SEMA3A
NRP1
Low
High
p = 0.002
SEMA3A–NRP1 spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)

Figure 3

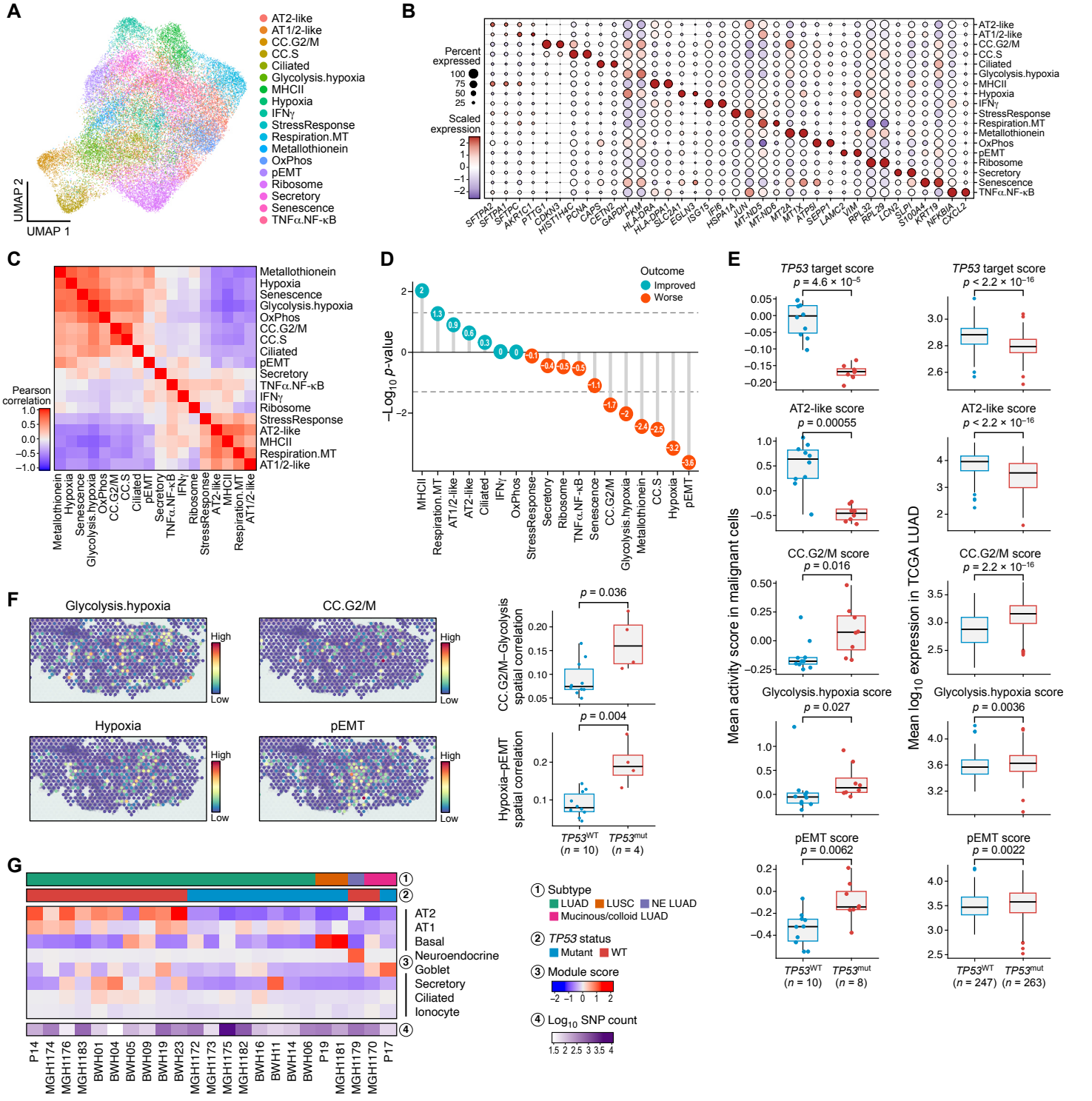

A
AT2-like
AT1/2-like
CC.G2/M
CC.S
Ciliated
Glycolysis.hypoxia
MHCII
Hypoxia
IFNγ
StressResponse
Respiration.MT
Metallothionein
OxPhos
pEMT
Ribosome
Secretory
Senescence
TNFα.NF-κB
UMAP 2
UMAP 1
B
Percent expressed
Scaled expression
C
Pearson correlation
D
Outcome
Improved
Worse
−Log10 p-value
E
TP53 target score
AT2-like score
CC.G2/M score
Glycolysis.hypoxia score
pEMT score
Mean activity score in malignant cells
Mean log10 expression in TCGA LUAD
TP53WT (n = 10)
TP53mut (n = 4)
TP53WT (n = 10)
TP53mut (n = 8)
TP53WT (n = 247)
TP53mut (n = 263)
F
Glycolysis.hypoxia
CC.G2/M
Hypoxia
pEMT
High
Low
CC.G2/M–Glycolysis spatial correlation
Hypoxia–pEMT spatial correlation
p = 0.036
p = 0.004
G
AT2
AT1
Basal
Neuroendocrine
Goblet
Secretory
Ciliated
Ionocyte
① Subtype
LUAD
LUSC
NE LUAD
Mucinous/colloid LUAD
② TP53 status
Mutant
WT
③ Module score
④ Log10 SNP count
H
Single cell entropy
Bulk entropy
Enntropy
Normal
TP53WT
TP53mut
* p = 0.00011
** p < 2.2 × 10−16
I
Mean entropy
J
Proportion HPCS
pEMT
Secretory
Ribosome
MHCII
IFNγ
CC.S
CC.G2/M

Figure 4

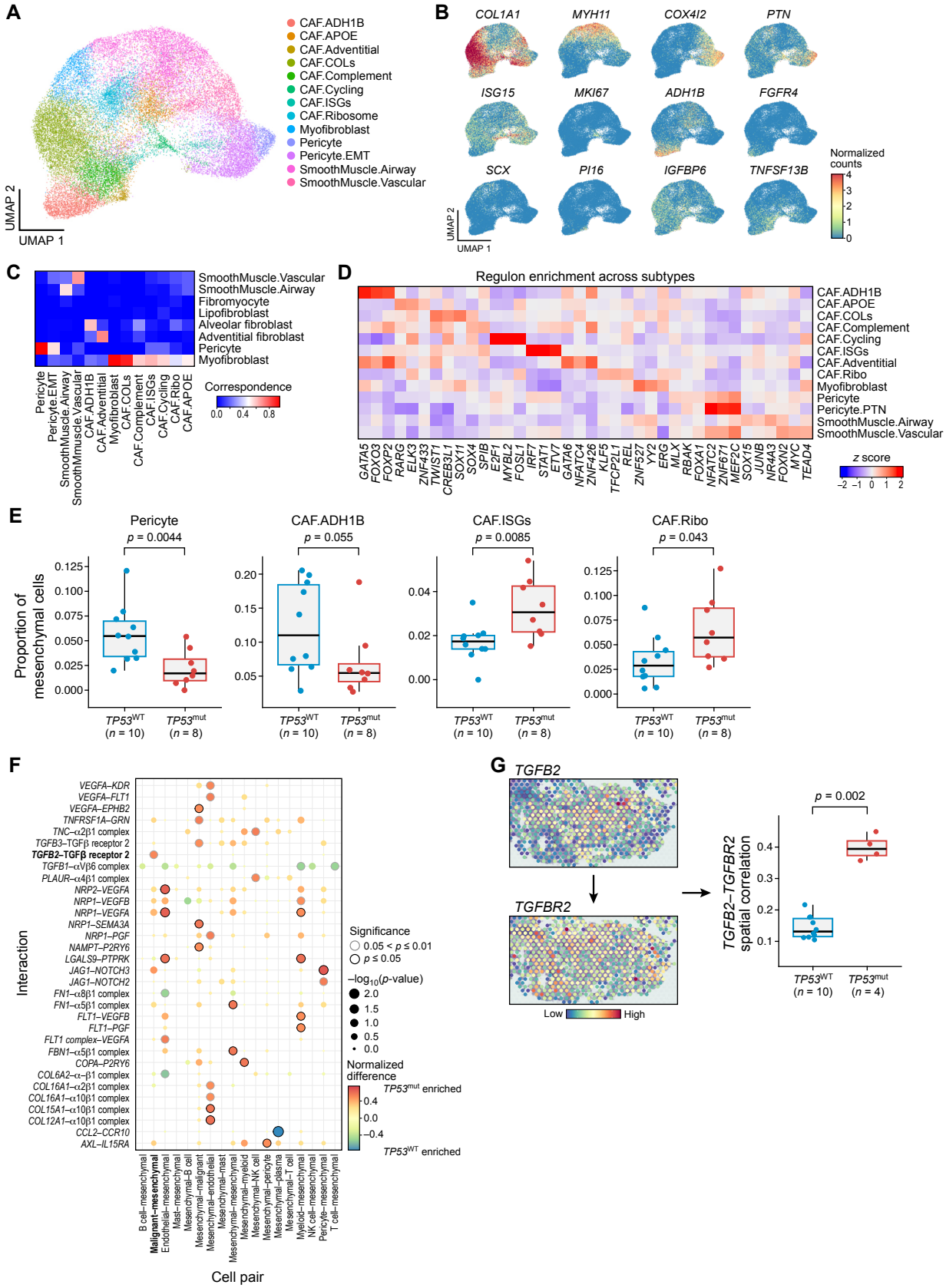

A
CAF.ADH1B
CAF.APOE
CAF.Adventitial
CAF.COLs
CAF.Complement
CAF.Cycling
CAF.ISGs
CAF.Ribosome
Myofibroblast
Pericyte
Pericyte.EMT
SmoothMuscle.Airway
SmoothMuscle.Vascular
UMAP 2
UMAP 1
B
COL1A1
MYH11
COX4I2
PTN
ISG15
MKI67
ADH1B
FGFR4
SCX
PI16
IGFBP6
TNFSF13B
Normalized counts
C
Correspondence
D
Regulon enrichment across subtypes
z score
E
Pericyte
CAF.ADH1B
CAF.ISGs
CAF.Ribo
Proportion of mesenchymal cells
F
Interaction
Cell pair
Significance
Normalized difference
G
TGFB2
TGFBR2
Low
High
TGFB2–TGFBR2 spatial correlation

Figure 5

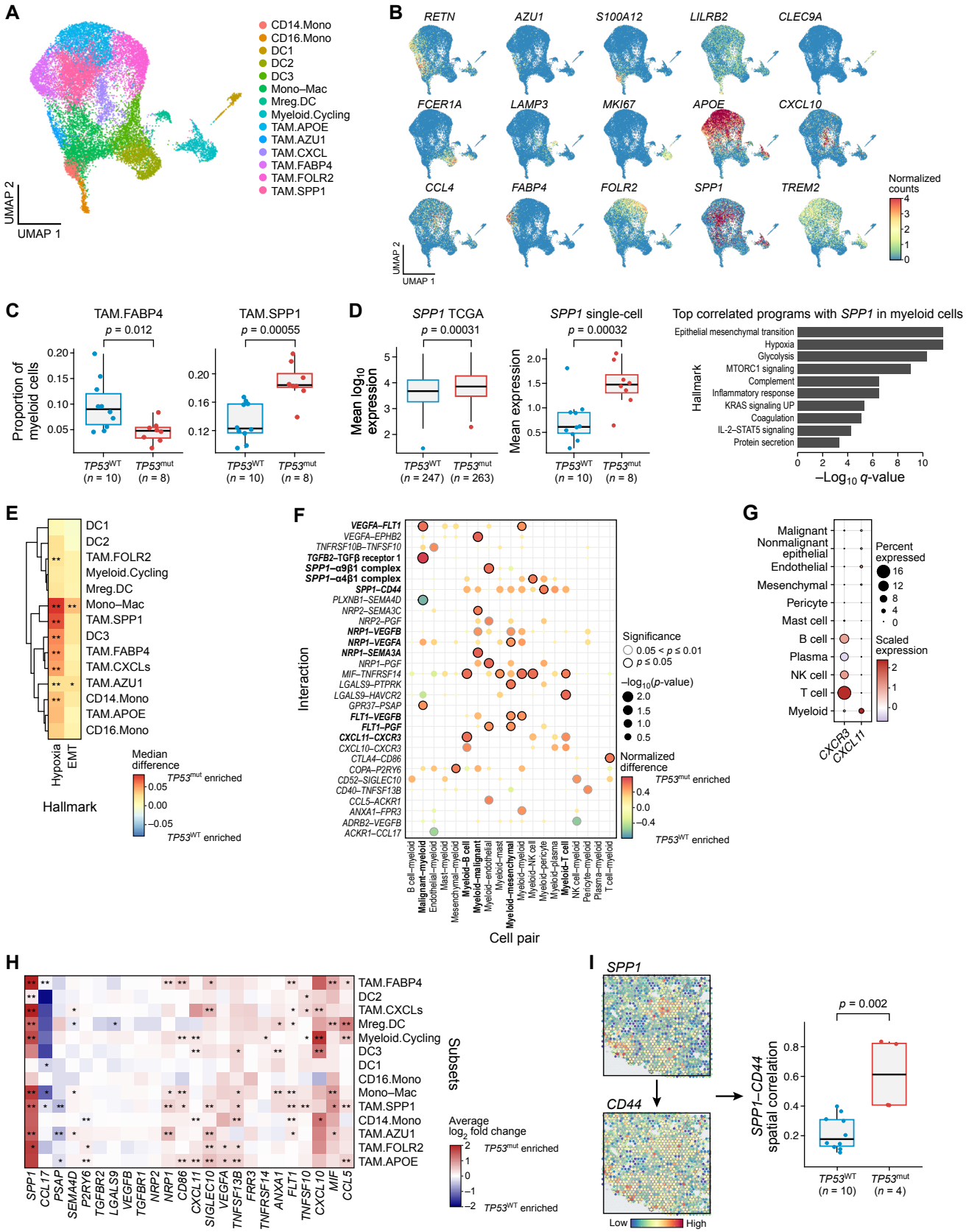

A
CD14.Mono
CD16.Mono
DC1
DC2
DC3
Mono–Mac
Mreg.DC
Myeloid.Cycling
TAM.APOE
TAM.AZU1
TAM.CXCL
TAM.FABP4
TAM.FOLR2
TAM.SPP1
UMAP 2
UMAP 1
B
RETN
AZU1
S100A12
LILRB2
CLEC9A
FCER1A
LAMP3
MKI67
APOE
CXCL10
CCL4
FABP4
FOLR2
SPP1
TREM2
Normalized counts
C
TAM.FABP4
p = 0.012
TAM.SPP1
p = 0.00055
Proportion of myeloid cells
TP53WT (n = 10)
TP53mut (n = 8)
D
SPP1 TCGA
p = 0.00031
Mean log10 expression
TP53WT (n = 247)
TP53mut (n = 263)
SPP1 single-cell
p = 0.00032
Mean expression
Top correlated programs with SPP1 in myeloid cells
Epithelial mesenchymal transition
Hypoxia
Glycolysis
MTORC1 signaling
Complement
Inflammatory response
KRAS signaling UP
Coagulation
IL-2–STAT5 signaling
Protein secretion
Hallmark
–Log10 q-value
E
DC1
DC2
TAM.FOLR2
Myeloid.Cycling
Mreg.DC
Mono–Mac
TAM.SPP1
DC3
TAM.FABP4
TAM.CXCLs
TAM.AZU1
CD14.Mono
TAM.APOE
CD16.Mono
Hypoxia
EMT
Hallmark
Median difference
TP53mut enriched
TP53WT enriched
F
VEGFA–FLT1
VEGFA–EPHB2
TNFRSF10B–TNFSF10
TGFB2–TGFβ receptor 1
SPP1–αvβ1 complex
SPP1–α4β1 complex
SPP1–CD44
PLXNB1–SEMA4D
NRP2–SEMA3C
NRP2–PGF
NRP1–VEGFB
NRP1–VEGFA
NRP1–SEMA3A
NRP1–PGF
MIF–TNFRSF14
LGALS9–PTPRK
LGALS9–HAVCR2
GPR37–PSAP
FLT1–VEGFB
FLT1–PGF
CXCL11–CXCR3
CXCL10–CXCR3
CTLA4–CD86
COPA–P2RY6
CD52–SIGLEC10
CD40–TNFSF13B
CCL5–ACKR1
ANXA1–FPR3
ADRB2–VEGFB
ACKR1–CCL17
Interaction
Significance
0.05 < p ≤ 0.01
p ≤ 0.05
–log10(p-value)
Normalized difference
TP53mut enriched
TP53WT enriched
B cell–myeloid
Malignant–myeloid
Endothelial–myeloid
Mast–myeloid
Mesenchymal–myeloid
Myeloid–B cell
Myeloid–malignant
Myeloid–endothelial
Myeloid–mast
Myeloid–mesenchymal
Myeloid–myeloid
Myeloid–NK cell
Myeloid–pericyte
Myeloid–plasma
Myeloid–T cell
NK cell–myeloid
Pericyte–myeloid
Plasma–myeloid
T cell–myeloid
Cell pair
G
Malignant
Nonmalignant epithelial
Endothelial
Mesenchymal
Pericyte
Mast cell
B cell
Plasma
NK cell
T cell
Myeloid
Percent expressed
Scaled expression
CXCR3
CXCL11
H
TAM.FABP4
DC2
TAM.CXCLs
Mreg.DC
Myeloid.Cycling
DC3
DC1
CD16.Mono
Mono–Mac
TAM.SPP1
CD14.Mono
TAM.AZU1
TAM.FOLR2
TAM.APOE
SPP1
CCL17
PSAP
SEMA4D
P2RY6
TGFBR2
LGALS9
VEGFB
TGFBR1
NRP2
NRP1
CD86
CXCL11
SIGLEC10
VEGFA
TNFSF13B
FRR3
TNFRSF14
ANXA1
FLT1
TNFSF10
CXCL10
MIF
CCL5
Subsets
Average log2 fold change
TP53mut enriched
TP53WT enriched
I
SPP1
CD44
Low
High
p = 0.002
SPP1–CD44 spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)

Figure 6

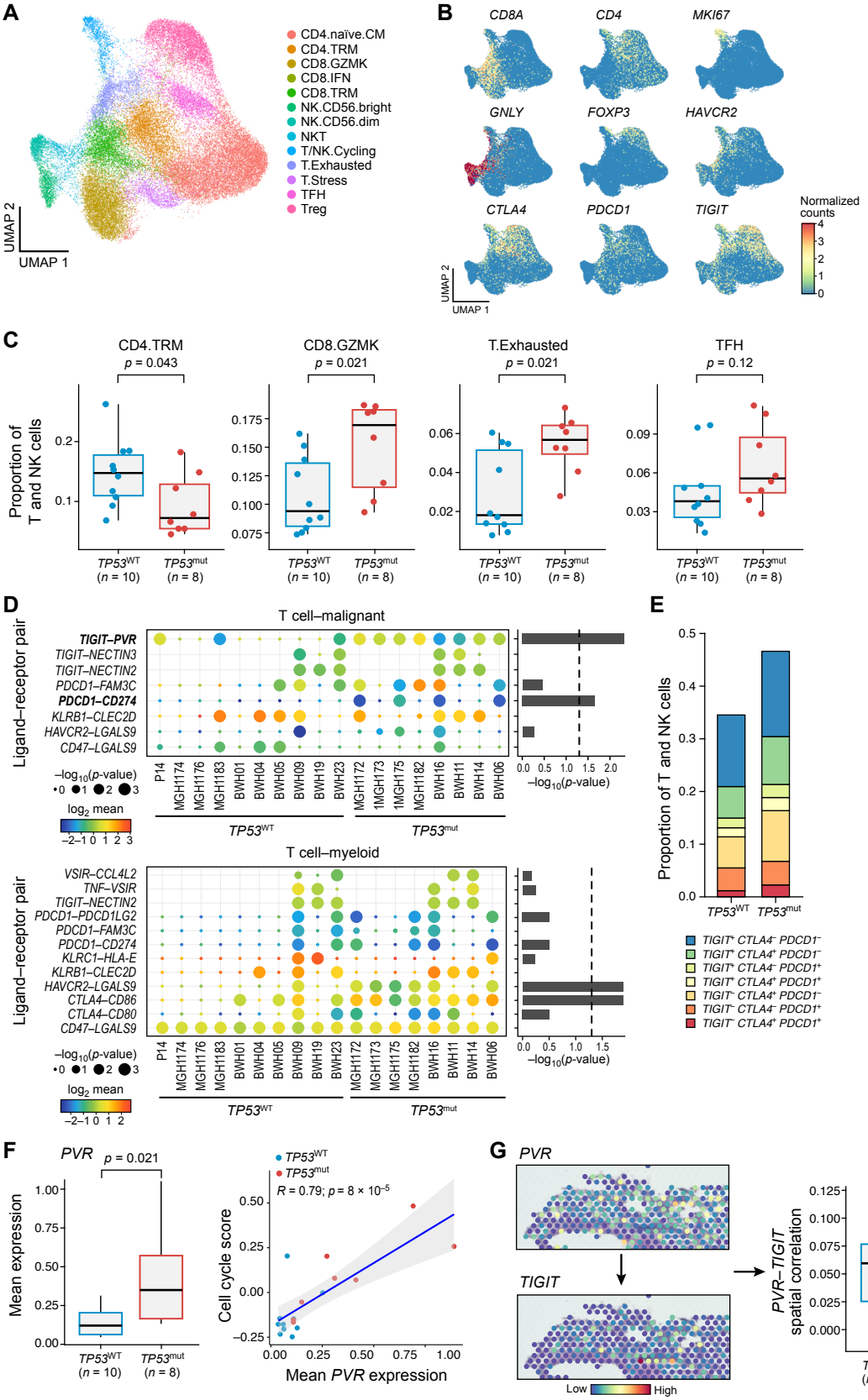
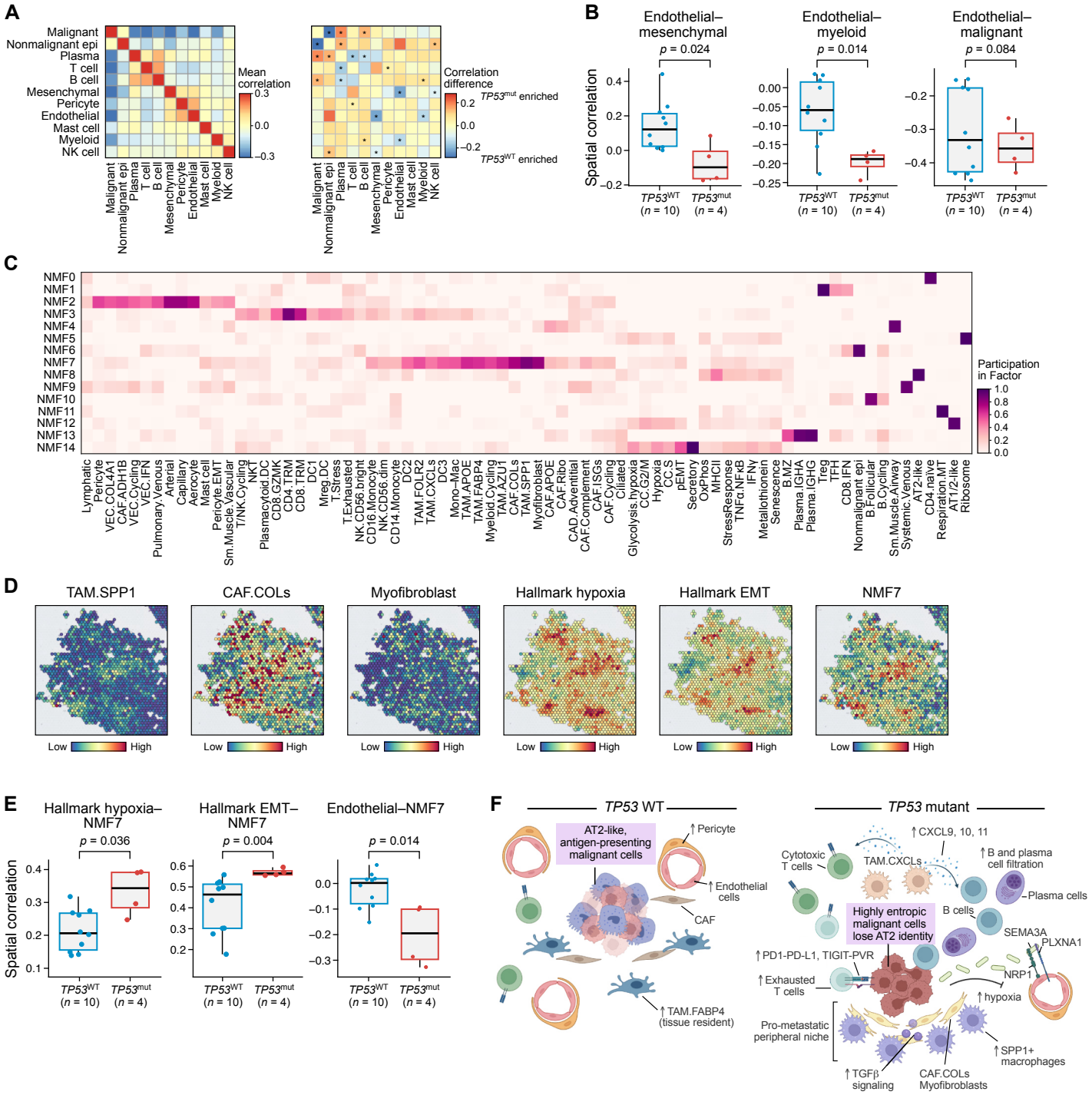
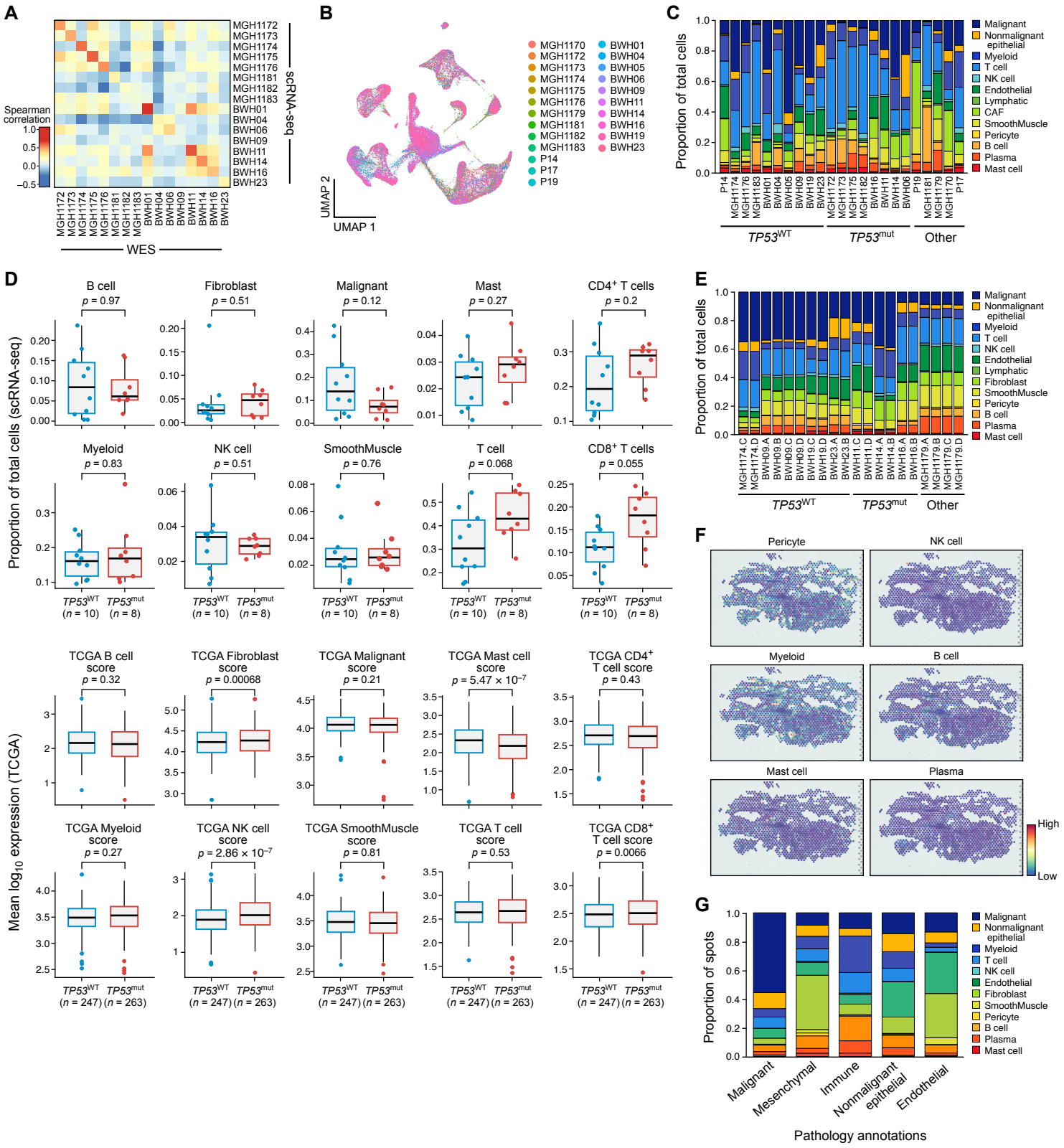

A
CD4.naïve.CM
CD4.TRM
CD8.GZMK
CD8.IFN
CD8.TRM
NK.CD56.bright
NK.CD56.dim
NKT
T/NK.Cycling
T.Exhausted
T.Stress
TFH
Treg
UMAP 2
UMAP 1
B
CD8A
CD4
MKI67
GNLY
FOXP3
HAVCR2
CTLA4
PDCD1
TIGIT
Normalized counts
C
CD4.TRM
p = 0.043
CD8.GZMK
p = 0.021
T.Exhausted
p = 0.021
TFH
p = 0.12
Proportion of T and NK cells
TP53WT (n = 10)
TP53mut (n = 8)
D
T cell–malignant
Ligand–receptor pair
TIGIT–PVR
TIGIT–NECTIN3
TIGIT–NECTIN2
PDCD1–FAM3C
PDCD1–CD274
KLRB1–CLEC2D
HAVCR2–LGALS9
CD47–LGALS9
–log10(p-value)
log2 mean
T cell–myeloid
VSIR–CCL4L2
TNF–VSIR
PDCD1–PDCD1LG2
KLRC1–HLA-E
CTLA4–CD86
CTLA4–CD80
E
Proportion of T and NK cells
TIGIT+ CTLA4− PDCD1−
TIGIT+ CTLA4+ PDCD1−
TIGIT+ CTLA4− PDCD1+
TIGIT+ CTLA4+ PDCD1+
TIGIT− CTLA4+ PDCD1−
TIGIT− CTLA4− PDCD1+
TIGIT− CTLA4+ PDCD1+
F
PVR
p = 0.021
Mean expression
TP53WT (n = 10)
TP53mut (n = 8)
R = 0.79; p = 8 × 10−5
Cell cycle score
Mean PVR expression
G
PVR
TIGIT
Low
High
p = 0.036
PVR–TIGIT spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)

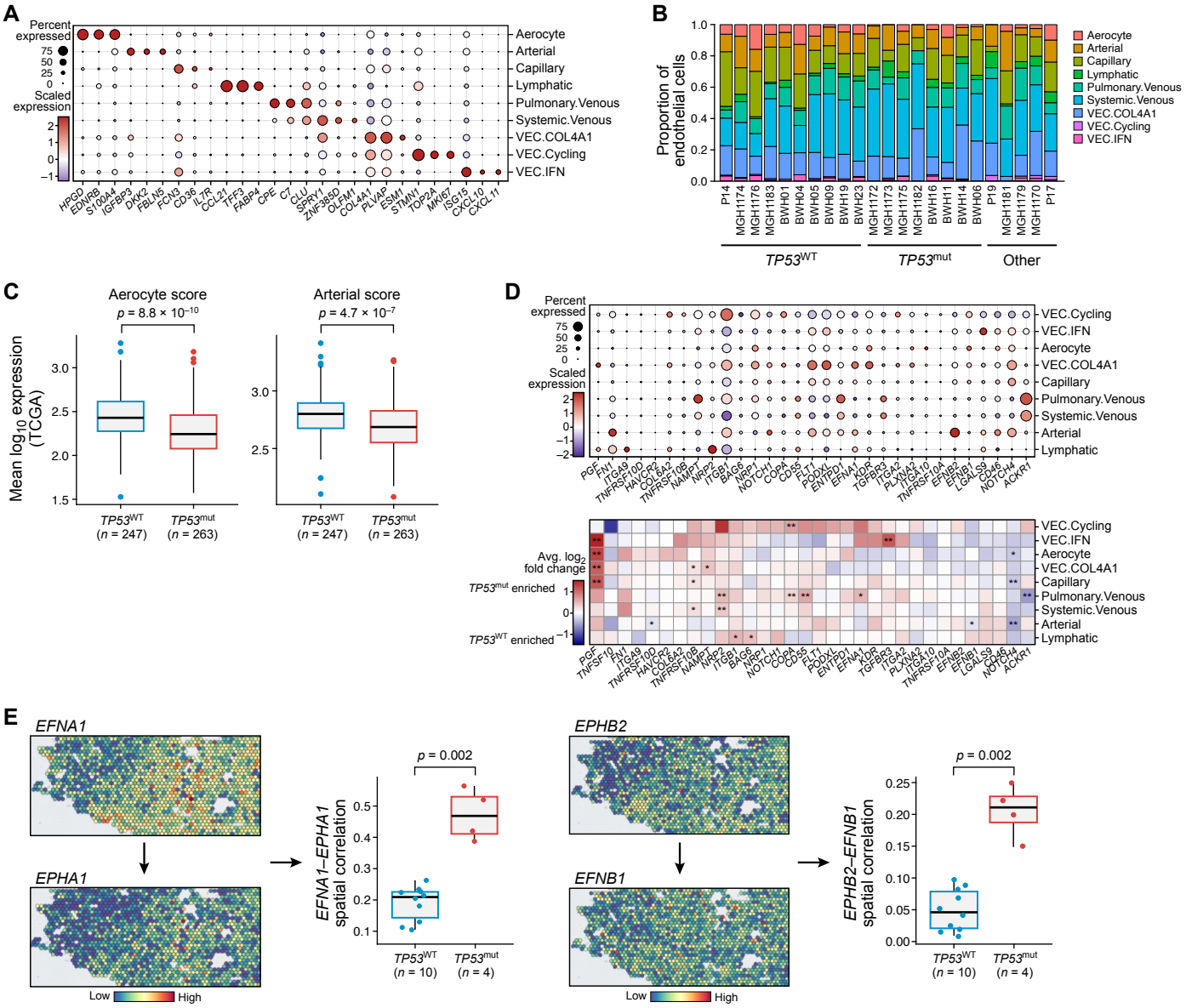
Figure 7


A
Malignant
Nonmalignant epi
Plasma
T cell
B cell
Mesenchymal
Pericyte
Endothelial
Mast cell
Myeloid
NK cell
Mean correlation
Correlation difference
TP53mut enriched
TP53WT enriched
B
Endothelial–mesenchymal
p = 0.024
Endothelial–myeloid
p = 0.014
Endothelial–malignant
p = 0.084
Spatial correlation
TP53WT (n = 10)
TP53mut (n = 4)
C
Participation in Factor
D
TAM.SPP1
CAF.COLs
Myofibroblast
Hallmark hypoxia
Hallmark EMT
NMF7
Low
High
E
Hallmark hypoxia–NMF7
p = 0.036
Hallmark EMT–NMF7
p = 0.004
Endothelial–NMF7
p = 0.014
Spatial correlation
F
TP53 WT
TP53 mutant
AT2-like, antigen-presenting malignant cells
Pericyte
Endothelial cells
CAF
TAM.FABP4 (tissue resident)
CXCL9, 10, 11
Cytotoxic T cells
TAM.CXCLs
B and plasma cell filtration
Plasma cells
B cells
Highly entropic malignant cells lose AT2 identity
SEMA3A
PLXNA1
PD1-PD-L1, TIGIT-PVR
NRP1
Exhausted T cells
hypoxia
Pro-metastatic peripheral niche
SPP1+ macrophages
TGFβ signaling
CAF.COLs Myofibroblasts

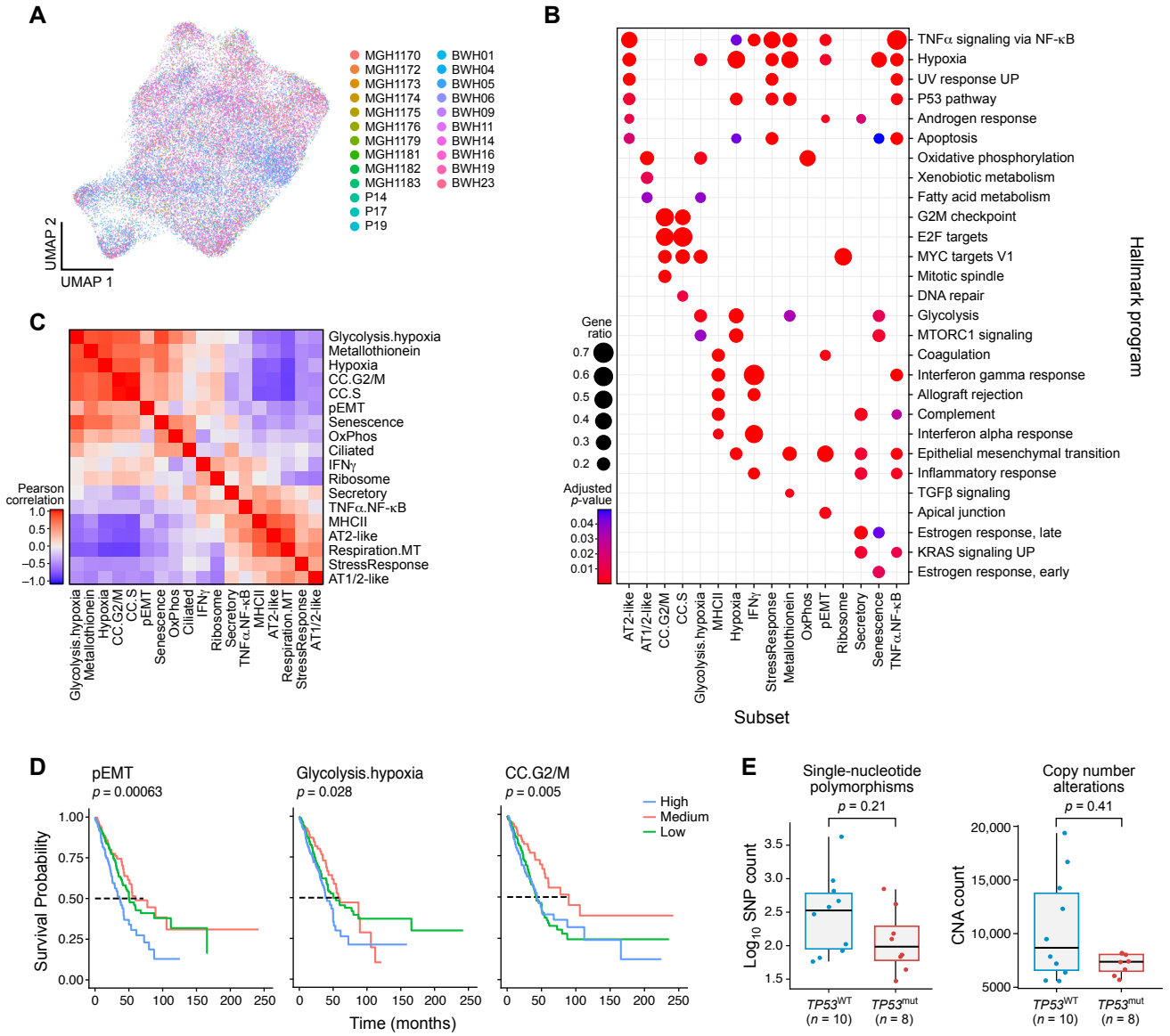
Supplementary Figure 1



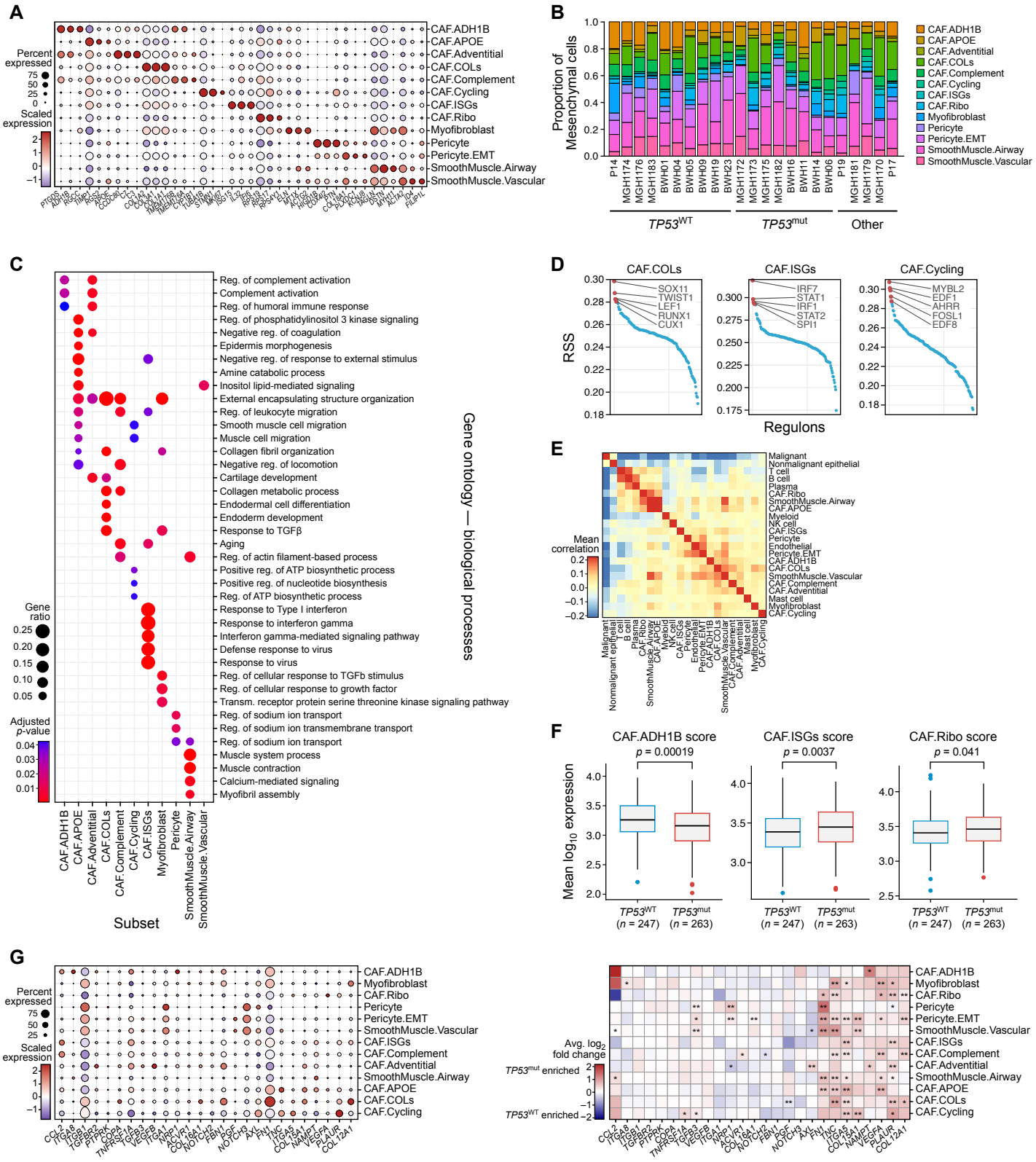
Supplementary Figure 2



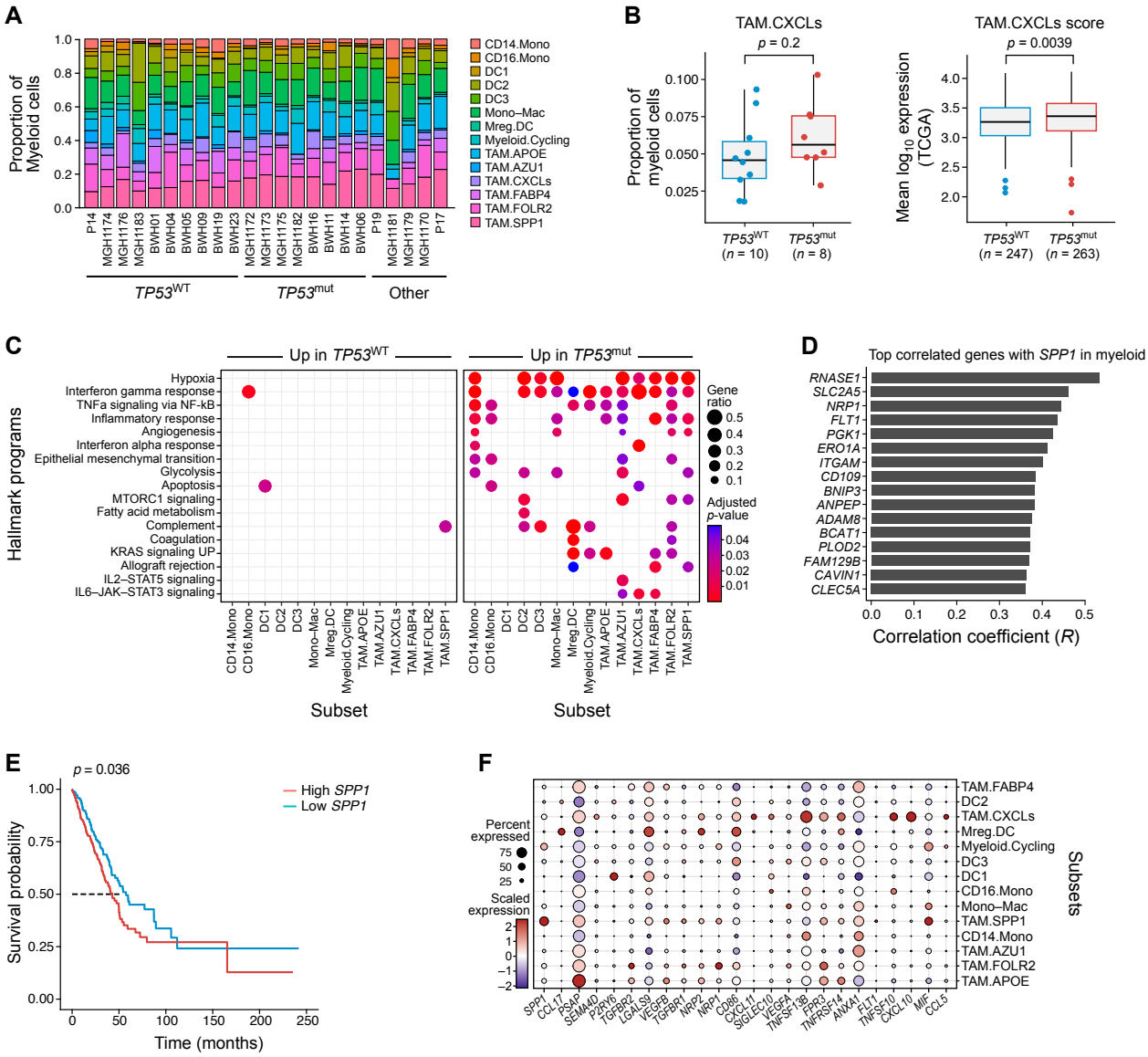
Supplemental Figure 3



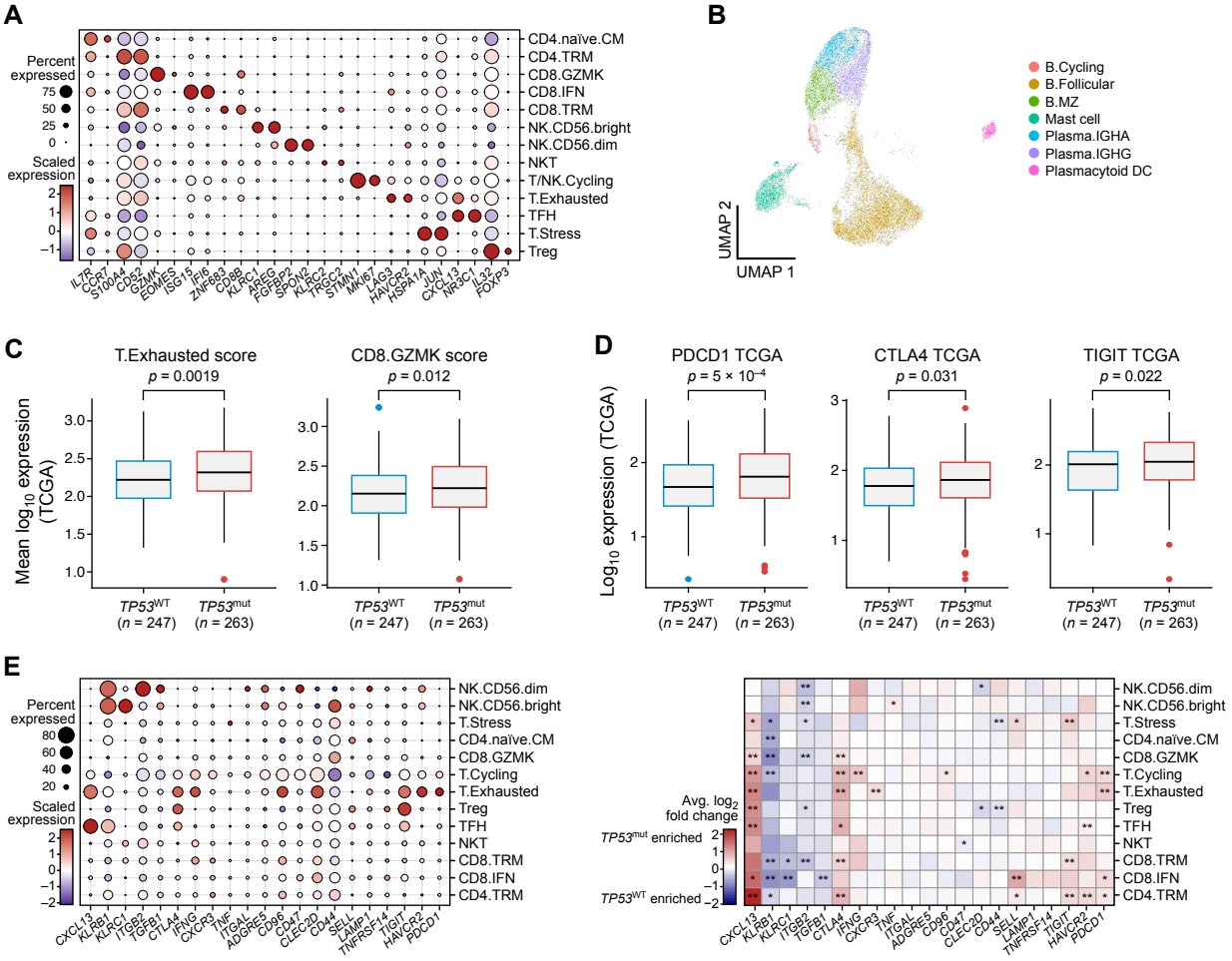
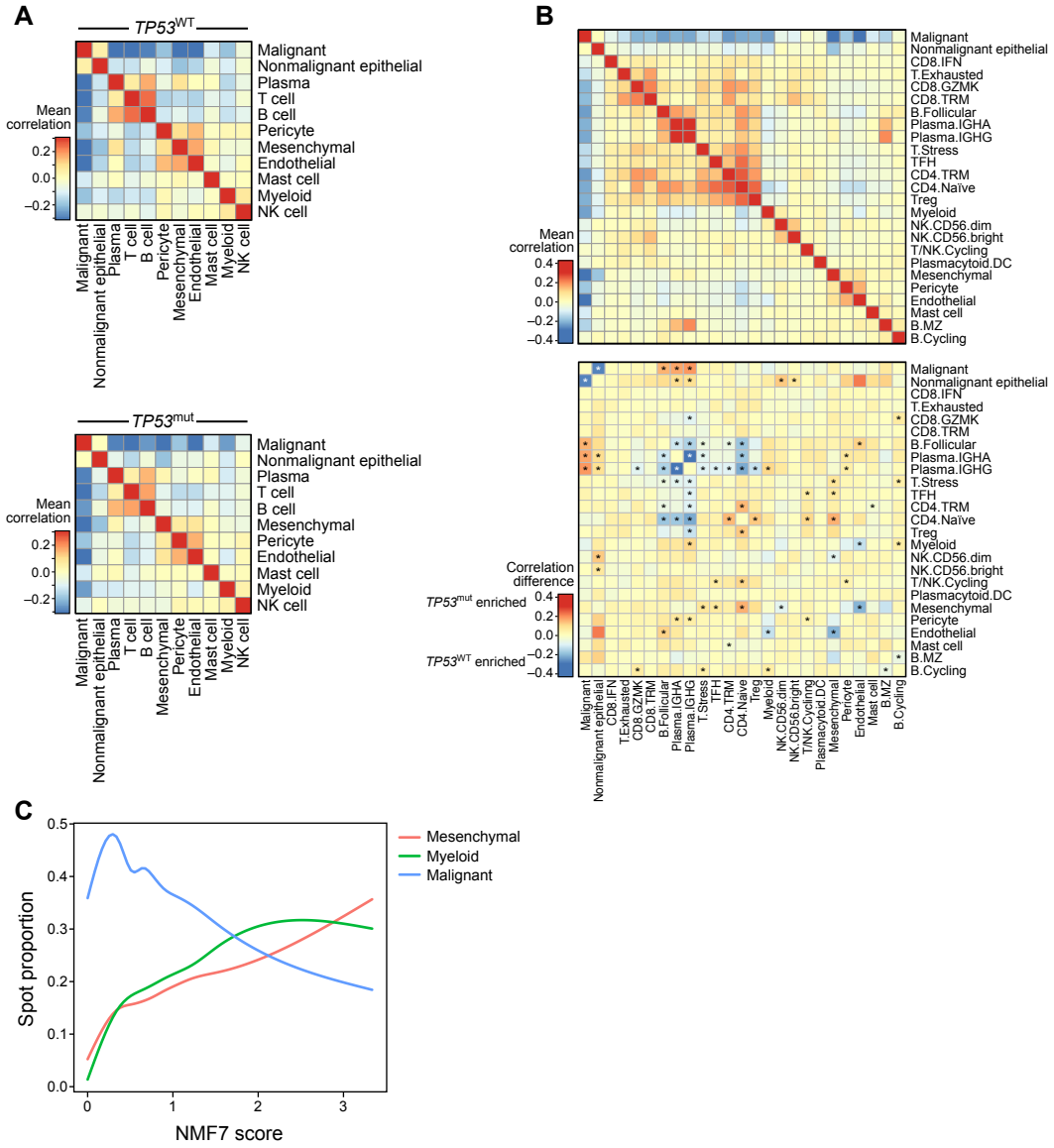
Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7